



Characterization of NosL, a novel Cu(I) protein of the nitrous oxide reductase gene cluster
by Michele Ann McGuirl

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in
Biochemistry

Montana State University

© Copyright by Michele Ann McGuirl (1999)

Abstract:

Metal chaperone structure and function is an emerging area of research in the field of bioinorganic chemistry. The assembly of the catalytic copper center in nitrous oxide reductase requires the assistance of auxiliary proteins. The isolation, and characterization of these ancillary proteins has not been previously reported. This investigation describes the cloning, recombinant expression, purification, and characterization of NosL from *Achromobacter cycloclastes*. NosL is an auxiliary protein of the nos (nitrous oxide reductase) gene cluster. Recombinant NosL has a high affinity Cu(I) binding site. Studies to elucidate the function of NosL suggest that it is not the electron-transfer partner to nitrous oxide reductase. In vivo investigations of the phenotype associated with nosL minus strains indicate that NosL is not strictly required for the production of active N₂OR, under the laboratory conditions employed in the study.

CHARACTERIZATION OF NOSL, A NOVEL CU(I) PROTEIN
OF THE NITROUS OXIDE REDUCTASE GENE CLUSTER

by

Michele Ann McGuirl

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Biochemistry

MONTANA STATE UNIVERSITY-BOZEMAN
Bozeman, Montana

October 1999

D378
M1792

APPROVAL

of a thesis submitted by

Michele Ann McGuirl

This thesis has been read by each member of the thesis committee and has been found to be satisfactory regarding content, English usage, format, citations, bibliographic style, and consistency, and is ready for submission to the College of Graduate Studies.

David M. Dooley

David M. Dooley
(Signature)

9/29/99
Date

Approved for the Department of Chemistry and Biochemistry

Paul A. Grieco

Paul A. Grieco
(Signature)

9-25-99
Date

Approved for the College of Graduate Studies

Bruce McLeod

Bruce R. McLeod
(Signature)

9-30-99
Date

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for a doctoral degree at Montana State University-Bozeman, I agree that the Library shall make it available to borrowers under rules of the Library. I further agree that copying of this thesis is allowable only for scholarly purposes, consistent with "fair use" as prescribed in the U.S. Copyright Law. Requests for extensive copying or reproduction of this thesis should be referred to University Microfilms International, 300 North Zeeb Road, Ann Arbor, Michigan 48106, to whom I have granted "the exclusive right to reproduce and distribute my dissertation in and from microform along with the non-exclusive right to reproduce and distribute my abstract in any format in whole or in part."

Signature

Michelle McGuff

Date

9/30/99

This thesis is dedicated to Erica Smith. May she, too, follow her dreams under God's guidance.

ACKNOWLEDGEMENTS

I would like to thank past and present members of the Dooley group, especially Doreen Brown, for their support and friendship during my stay in Bozeman. Dr. Martin Teintze also provided many helpful discussions and guidance. I would also like to thank Maureen McGuirl for her support during this period.

I wish to express my deepest appreciation to Dr. David Dooley for his many years of inspiration and guidance. He is the essence of the word "mentor."

TABLE OF CONTENTS

	Page
INTRODUCTION	1
Overview of Denitrification	1
Key Enzymes of Denitrification	4
Regulation of Denitrification	6
Structure of Nitrous Oxide Reductase	8
The <i>nos</i> (Nitrous Oxide Reductase) Gene Cluster	9
Metallochaperones	12
Research Goals	13
SEQUENCE ANALYSIS OF THE <i>NOS</i> GENE CLUSTER	15
Introduction	15
Experimental Methods	19
Isolation of <i>nos</i> DNA	19
DNA Sequencing and Analysis	20
Protein Characterization and Expression	20
Molecular Modeling	20
Results	21
The <i>nos</i> Gene Cluster	21
NosR	23
N ₂ OR (NosZ)	26
Model of the Cu _A Site in N ₂ OR	29
NosD, F, Y, and L	32
NosX	36
Discussion	37
Conclusions	45
CHARACTERIZATION OF NOSL, A NOVEL CU(I) PROTEIN	47
Introduction	47
Experimental Methods	51
Design of Recombinant Constructs	51
Expression of rNosD	52
Expression and Purification of rNosL	52
Analysis of rNosL	54
Results	55
Expression Trials of rNosD	55
Protein Characterization of rNosL	56
Metal Content of rNosL	57

Elucidation of the Ligands to Cu(I)-rNosL	61
Discussion	64
rNosD	64
rNosL	65
Conclusions	69
INVESTIGATIONS OF NOSL FUNCTION	70
Introduction	70
Experimental Methods	74
Insertion of the <i>nos</i> Gene Cluster into a Broad Host Range Plasmid	74
Construction of <i>nosL</i> Deletion Plasmids	74
Construction of a <i>nosZ</i> Deletion Plasmid	75
Construction of <i>nosL</i> Antisense Plasmids	75
Introduction of pML122-Based Plasmids into <i>A. cycloclastes</i>	76
Introduction of pML122-Based Plasmids into <i>S. meliloti</i>	77
GC Assay of N ₂ O Reductase Activity	77
Purification of N ₂ OR	78
Interactions of rNosL and apo-N ₂ OR	79
Reduction of N ₂ OR by rNosL	79
Oxidation of Cu(I)-rNosL During the Catalytic Turnover of N ₂ OR	80
Results	80
Purification and Preliminary Characterization of Apo- and Holo-N ₂ OR..	80
The Possible Function of rNosL as an Electron-transfer Protein	80
<i>In Vitro</i> Study of the Possible Function of rNosL as a Copper Chaperone to N ₂ OR	82
<i>In Vivo</i> Studies of the Possible Function of rNosL as a Copper Chaperone to N ₂ OR	84
Discussion	86
Conclusions	90
RESEARCH SIGNIFICANCE AND FUTURE DIRECTIONS	92
REFERENCES CITED	95
LIST OF APPENDICES	112
Appendix A - Comparison of 50 Amino Acid Sequences of N ₂ OR	113
Appendix B - Genbank File of the Sequence of the <i>nos</i> Cluster from <i>A. cycloclastes</i>	128
Appendix C - Genbank File of the Sequence of <i>nosR</i> from <i>S. meliloti</i>	135
Appendix D - XAS and EXAFS Data and Analysis for rNosL	137

LIST OF TABLES

Table	Page
1. Copper Analysis of rNosL and N ₂ OR	83
2. Rates of N ₂ O Reduction in Different Strains of <i>S. meliloti</i>	84
3. Rates of N ₂ O reduction at Different Concentrations of Copper in the Growth Media	86

LIST OF FIGURES

Figure	Page
1. The Global Nitrogen Cycle	1
2. The Denitrification Pathway	3
3. The Genes of the <i>nos</i> Gene Cluster	9
4. Possible Function of NosD, F, Y, and L	11
5. The Arrangement of Genes in the <i>nos</i> Gene Clusters of Four Denitrifiers	21
6. Comparison of the Deduced Amino Acid Sequences of NosR	24
7. Sequence Alignment of N ₂ OR	27
8. Ribbon Diagram of a Model of the Cu _A Domain	30
9. Surface Map of the N ₂ OR Cu _A Model	30
10. Close-up of the Modeled Cu _A Site	31
11. NosD Sequence Alignment	33
12. Sequence Alignment of NosF	34
13. Hydropathy Plots of NosY	34
14. Sequence Alignment of NosL	35
15. Sequence Alignment of NosX	36
16. The pET20b+ Expression Vector	49
17. The Lipobox Consensus Sequence of SpaseII	50
18. Chromatogram of the Purification of rNosL	53
19. SDS/PAGE Showing rNosD Expression Over Time	55

20. Native PAGE and Gel Filtration Estimate of rNosL MW	57
21. Spectroscopic Characterization of rNosL	59
22. The reaction of Biquinoline with rNosL	60
23. Oxidation of Cu(I)-rNosL	61
24. The Reaction of Various Forms of rNosL with the Cys reagent PDS	62
25. Analysis of the Copper Site in rNosL	63
26. Visible Absorption Spectrum of <i>A. cycloclastes</i> N ₂ OR	81
27. EPR Spectra of rNosL	82
28. Chromatogram of a mixture of rNosL and apo-N ₂ OR	83

ABSTRACT

Metal chaperone structure and function is an emerging area of research in the field of bioinorganic chemistry. The assembly of the catalytic copper center in nitrous oxide reductase requires the assistance of auxiliary proteins. The isolation, and characterization of these ancillary proteins has not been previously reported. This investigation describes the cloning, recombinant expression, purification, and characterization of NosL from *Achromobacter cycloclastes*. NosL is an auxiliary protein of the *nos* (nitrous oxide reductase) gene cluster. Recombinant NosL has a high affinity Cu(I) binding site. Studies to elucidate the function of NosL suggest that it is not the electron-transfer partner to nitrous oxide reductase. *In vivo* investigations of the phenotype associated with *nosL* minus strains indicate that NosL is not strictly required for the production of active N₂OR, under the laboratory conditions employed in the study.

INTRODUCTION

Overview of Denitrification

Denitrification is a key step of the global nitrogen cycle (Figure 1), providing the pathway by which fixed nitrogen is returned to the atmosphere (1,2). In the nitrogen cycle, atmospheric

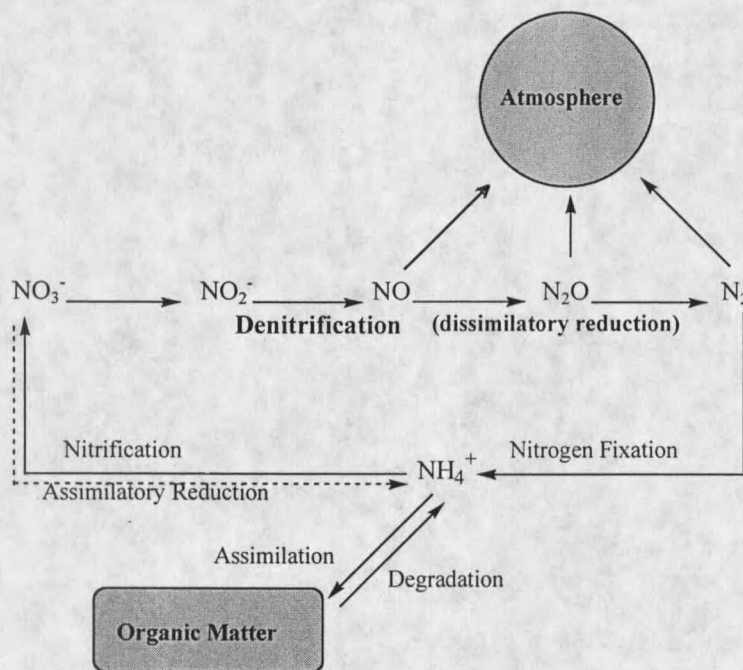


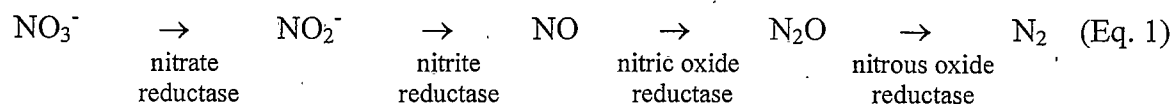
Figure 1. The global nitrogen cycle.

nitrogen is reduced to ammonia by a select group of bacteria. The ammonia is then assimilated by plants, animals and microbes. When returned to soil or water by the degradation of waste or decaying matter, released ammonia is either re-assimilated or converted to nitrate or nitrite through the

nitrification pathway. Nitrate thus formed experiences one of two fates. First, nitrate may be reduced back to ammonia via the assimilatory reduction pathway. This reaction occurs in plants, fungi, and microbes. Alternatively, a subfamily of bacteria known as denitrifiers may convert nitrate to nitrogen gas via the denitrification pathway and thus complete the nitrogen

cycle. Recently several fungi have also been shown to reduce nitrate to nitrous oxide, but lack the enzyme to complete the final step and produce N_2 (3,4).

Denitrification provides an alternative respiratory pathway for denitrifying bacteria under oxygen-limited conditions, with N-oxides acting as terminal electron acceptors in lieu of O_2 . Although denitrification has traditionally been considered an anaerobic process, several bacteria have been shown to be capable of aerobic denitrification. In either case, nitrate is reduced to dinitrogen by the action of four discrete metalloenzymes as shown below in Equation 1:



The electrons for the reaction are provided by small, soluble electron transfer proteins, which are in turn supplied by a membrane-bound bc_1 complex (5). In the bc_1 complex, electron transfer is coupled to H^+ pumping, which drives ATP synthesis (Figure 2).

Denitrification has been associated with several detrimental environmental effects. For example, denitrifying bacteria efficiently utilize nitrate that is applied as fertilizer to crops (1,6,7). Depending on soil conditions, up to 40% of the fertilizer may be lost to denitrifiers, which results in millions of dollars in wasted resources annually in the United States alone. Moreover, in fungi and a few bacteria, the denitrification pathway ends prematurely with the production of the greenhouse gas nitrous oxide (2,4,8). N_2O has a remarkably stable half-life of about 150 years (7,8). Studies show that N_2O reaches the stratosphere, where it is oxidized to NO by reaction with singlet oxygen, which is produced

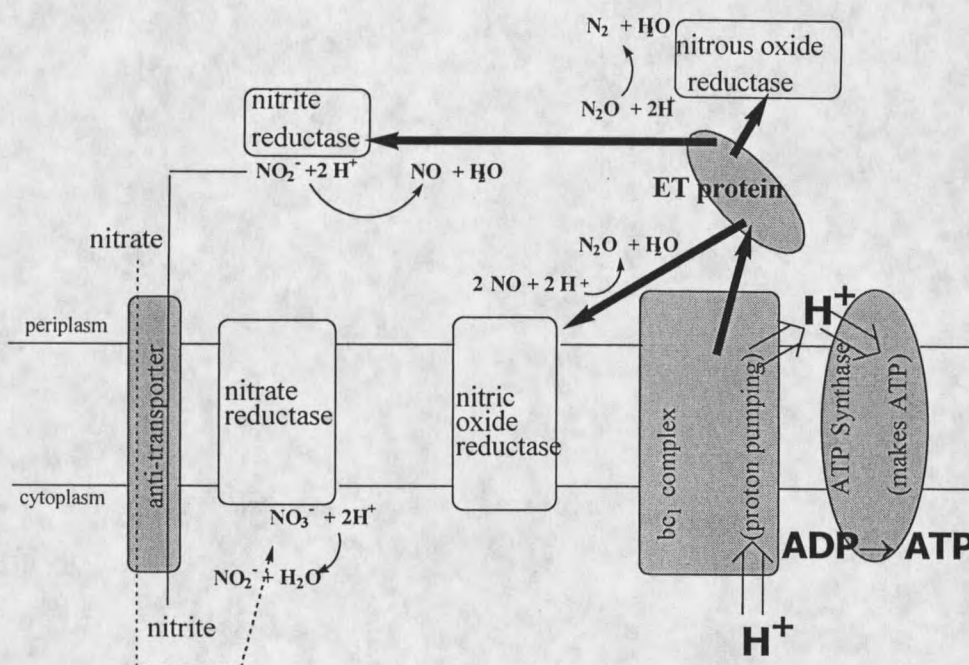


Figure 2. The Denitrification Pathway. $\blacksquare$ represents electron transfers between proteins. $\square$ represents the flow of protons used to drive ATP synthesis. For simplicity, only the membrane-bound form of nitrate reductase is shown.

by the photochemical decomposition of O_3 (9). NO continues reacting with O_3 , further depleting the earth's protective ozone layer.

While some aspects of denitrification are environmentally harmful, others are beneficial and have begun to be exploited commercially. For example, denitrifiers have been successfully used in bioremediation to remove nitrates from water supplies (10-12). It is perhaps ironic that the same organisms that mandate the use of high levels of fertilizer also serve to purify the groundwater that, as a result, becomes contaminated. Another area of environmental concern for which denitrification may provide answers is N_2O pollution. In addition to the N_2O produced from biological sources (1,8), significant amounts of N_2O are

released as a waste product during industrial nylon production (13) and biomass burning (14). Studying the biological mechanism of N_2O reduction to N_2 may aid the design of an inexpensive biomimetic catalyst. In addition to these anthropogenic applications, denitrification in marine sediments plays an important role in regulating the N-oxide flux in seawater, accounting for the removal of over 50% of N-oxides from the marine environment (15,16).

Key Enzymes of Denitrification

In the first step of denitrification, the enzyme nitrate reductase (NaR) catalyzes the 1-electron reduction of nitrate to nitrite (1,2,5). Both periplasmic and membrane-bound forms of NaR have been identified, and the role of the periplasmic NaR has long been debated (2,5). It now appears that the periplasmic enzyme plays different roles in different organisms, serving as the primary NaR during anaerobic denitrification in some bacteria (17,18), while in others it functions only during the aerobic process (19).

In addition to being the first enzyme of denitrification, NaR is also found in the assimilatory reduction pathway. In the latter, product nitrite is enzymatically converted to ammonia. Regardless of the pathway in which they function, NaRs typically contain the cofactor molybdopterin guanine dinucleotide (MGD), and an iron-sulfur electron-transfer center, in addition to heme centers (2,5). The structural homology among NaRs is not surprising, given that the enzymes catalyze the same reaction. Recently molybdenum-free membrane-bound and periplasmic NaRs have been isolated from the vanadate-reducing

bacteria *Pseudomonas isachenkovii*, but structural information on these variants is limited (20).

In the second step of denitrification, nitrite is reduced to nitric oxide by the enzyme nitrite reductase (NiR) (2). Two types of NiRs have been characterized from denitrifiers: a copper enzyme (CuNiR) (21-24) and a heme-containing protein (*cd*₁NiR) (25-30). While denitrifiers contain either CuNiR or *cd*₁NiR, but not both, it has been shown that CuNiR can be functionally produced in a *cd*₁NiR⁻ mutant strain (31). Both types of NiR have been extensively characterized and bear little resemblance to each other, or to the assimilatory NiR, which reduces nitrite to ammonia.

Once nitric oxide is released from NiR it is further reduced to nitrous oxide by the enzyme nitric oxide reductase (NOR). NOR, which is found in the cytoplasmic membrane, contains one type *c* and two type *b* cytochromes, in addition to a non-heme iron (32-37). Several structures for the active site of NOR have been proposed, including species in which the non-heme iron is linked by a μ -oxo or hydroxy bridge to one of the *b* hemes (37). Two possible mechanisms for NO reduction have been suggested. In the first, a single NO is bound and reduced to the nitroxyl anion (NO⁻), which is then non-enzymatically converted to N₂O (38). Alternatively, two molecules of NO could bind to the dinuclear center, setting up the formation of the N-N bond (37). Additional biophysical characterization of NOR is needed to fully understand the structure/function relationship of the metal sites in this intriguing enzyme.

Nitrous oxide reductase (N_2OR) is the terminal enzyme of denitrification and reduces N_2O by two electrons to N_2 . Although N_2OR s are typically soluble, periplasmic enzymes (39-44), at least one may be membrane-associated (45). To date, all N_2OR s have been found to be homodimers that contain two distinct dinuclear copper centers per subunit. The electron transfer copper site, called Cu_A , is analogous to the Cu_A site in cytochrome *c* oxidase (46-52), and the catalytic copper site, called Cu_{Cat} , is where N_2O is thought to bind and be reduced (53-55).

It is noteworthy that the terminal oxidases of aerobic respiration, *e.g.*, cytochrome oxidase and quinol oxidase, have structural components that closely resemble elements of NOR and N_2OR (2, 56-59). Additionally, denitrification genes have been identified in several halophilic and hyperthermophilic members of the archaea kingdom (2, 60-62). Thus, the study of the denitrification genes and their related proteins may provide clues to the evolution of enzymes involved in bioenergetic processes.

Regulation of Denitrification

The regulation of denitrification genes has been extensively studied in *Pseudomonas stutzeri*, *Paracoccus denitrificans*, *Pseudomonas aeruginosa*, and *Rhodobacter sphaeroides* (63-77). Regulation responds to both the oxygen tension and the type of N-oxide present. While in principle two transcriptional regulators (one for O_2 , one for N-oxides) could globally control the biosynthesis of all denitrifying enzymes, the data indicate that a far more complex situation exists. For example, in *P. stutzeri* both nitrate and nitrite induce the

transcription of all four denitrification enzymes, yet different transcriptional factors are used for each enzyme (63,64). Moreover, the response to exogenous signals differs among the denitrifiers studied, which hampers the efforts to identify and characterize regulatory proteins involved in denitrification. As previously mentioned, nitrite induces the production of NiR and NOR in *P. stutzeri*, while in *P. aeruginosa* and *R. sphaeroides* NO is the regulatory signal for the production of these enzymes (66,68,69). Despite this difference, regulation of NiR and NOR is tightly coupled in all organisms examined, perhaps to prevent the accumulation of NO, which may be toxic to cells.

The expression of denitrification proteins in response to oxygen is controlled by FNR-like proteins (2,65,67,70-77). FNR, the fumarate and nitrate reduction regulator protein discovered in *E. coli*, is related to CRP, the cyclic AMP receptor protein (78). Unlike CRP, FNR contains an N-terminal Cys-ligated iron-sulfur moiety that acts as an O₂ sensor. FNR recognizes a palindromic TTGAT-N₄-ATCAA sequence and both positively and negatively regulates transcription. FNR analogs from several denitrifiers have been identified and shown to be directly involved in denitrification: these include ANR, DNR, and NNR, as well as FnrA, FnrD, and FnrP. Some of these proteins possess the N-terminal Cys cluster indicative of the O₂-sensing Fe-S cluster, while others more closely resemble the FixK regulator in that they lack a Cys cluster. Evidence suggests that denitrification uses a cascade of multiple O₂-responsive regulatory proteins within the same organism.

Structure of Nitrous Oxide Reductase

The structural gene of N₂OR, *nosZ*, has been sequenced entirely for seven different denitrifiers (59,79-83) and partially for an additional 44 organisms (mostly uncultured marine bacteria) (84-87). Comparison of these translated amino acid sequences (Appendix A) reveals a C-terminal Cu_A binding motif that is similar to the Cu_A region of cytochrome *c* oxidase (59,79,87). In the latter enzyme, the two coppers of the dinuclear Cu_A center are bridged by two Cys residues, forming a near-coplanar geometry. Two His, one Met, and the peptide carbonyl oxygen of Glu provide the other ligands to the dinuclear site (51,52). In N₂OR, Glu is replaced by Trp in all known sequences except that from *Ralstonia eutropha*, where the corresponding amino acid is His.

In addition to the amino acids that are assigned as Cu_A ligands, there are several conserved amino acids (9 His, 3 Met) that are frequently coordinated to copper, and are therefore candidates for Cu_{Cat} ligands (59,79,80). Despite the lack of a strictly conserved Cys (other than the Cys associated with the Cu_A center), numerous spectroscopic studies have long been interpreted as evidence for a Cu-Cys thiolate bond in the Cu_{Cat} site of *P. stutzeri*. (53,54,88,89). Sequence analysis suggests that a Cys residue found in a variable position (within a 5-residue stretch) near the Cu_{Cat} domain might provide the thiolate bond (59). Alternatively, Farrar *et al* have proposed that all spectroscopic signatures previously thought to arise from a Cu_{Cat}-thiolate bond are in fact associated with a structural variant of the Cu_A

center (90). The authors postulate that the Cu_{Cat} site is a His-ligated, spectroscopically undetected (to date) Cu(II)-Cu(II) dinuclear center.

The *nos* (Nitrous Oxide Reductase) Gene Cluster

In addition to the structural *nosZ* gene, production of active N₂OR requires several ancillary *nos* genes (Figure 3). While none of these Nos proteins have been previously isolated or characterized, preliminary information about their structures has been gleaned from examining their translated amino acid sequences. *NosR*, located upstream of the structural *nosZ* gene in a separate transcriptional unit, codes for a regulatory integral membrane protein (59,91,92). Complete sequence information for *nosR* has been reported only for four organisms (59,83,92); however, the upstream flanking regions of other *nosZ* genes suggest that *nosR* is present in all organisms examined except *R. eutropha* (79).

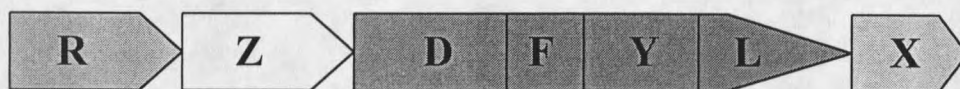


Figure 3. The genes of the *nos* gene cluster, showing the four transcriptional units *nosR*, *nosZ*, *nosDFYL*, and *nosX*.

The C-terminus of NosR contains two Cys clusters similar to the [4Fe-4S] binding motifs of several bacterial ferridoxins (59,92). Additionally, another potential metal site consisting of two highly conserved CysXaa₃CysPro motifs has been identified. These metal binding motifs have been noted in several other membrane proteins. For example, NapH and MauN have

been shown to have electron-transfer function and operate in conjunction with a periplasmic component (5).

In 1997 Chan and coworkers reported the existence of another *nos* gene, *nosX*, in *Sinorhizobium meliloti* (91). Tn5 mediated mutagenesis of this region indicates that this gene is located in a separate complementation group downstream from *nosDFYL* (Figure 3). Although preliminary studies suggest that NosX may interact with NosR, further work is needed to determine the precise role of *nosX* in the biosynthesis of N₂O reductase. While *nosX* has been identified in *Achromobacter cycloclastes* (59), and *B. japonicum* (83), there is no evidence for an analogous gene in the region downstream of *nosL* in *P. stutzeri* (91,93). Interestingly, *nosR/nosX* homologs have been noted in the gene clusters of other denitrification enzymes (94,95).

Another transcriptional unit, *nosDFYL*, is located downstream of *nosZ* and contains the genes necessary for proper copper incorporation into N₂OR (96). Thus, N₂OR belongs to a group of metalloproteins, e.g. urease, hydrogenase, and nitrogenase (97), which require other proteins for metal-site assembly or proper folding, or both. Analysis of the published *nosDFYL* sequences for *Sinorhizobium meliloti* (81,91), *P. stutzeri* (96,98) and *A. cycloclastes* indicates homology to multi-component ABC transporter systems (99). These systems, which are usually involved in the transport of small molecules and proteins across a membrane, consist of a cytoplasmic ATP/GTPase, a transmembrane protein, and one or two periplasmic components specific to the molecule to be transported. It is possible that for the bioassembly of N₂OR, the ABC system has been adapted for use as an energy transducer rather than as a transporter (59). In this hypothesis, the cytoplasmic hydrolysis of ATP or

GTP by NosF is coupled to the periplasmic incorporation of copper into N₂OR by NosD and NosL via the integral membrane protein NosY (Figure 4). NosD has been suggested to be a chaperone protein that keeps apo-N₂OR in a conformation amenable to copper insertion (96). The research described in this work shows that NosL is a Cu(I) protein that may function as a metallochaperone, providing a source of copper under limiting conditions and/or conferring metal specificity to the bioassembly process.

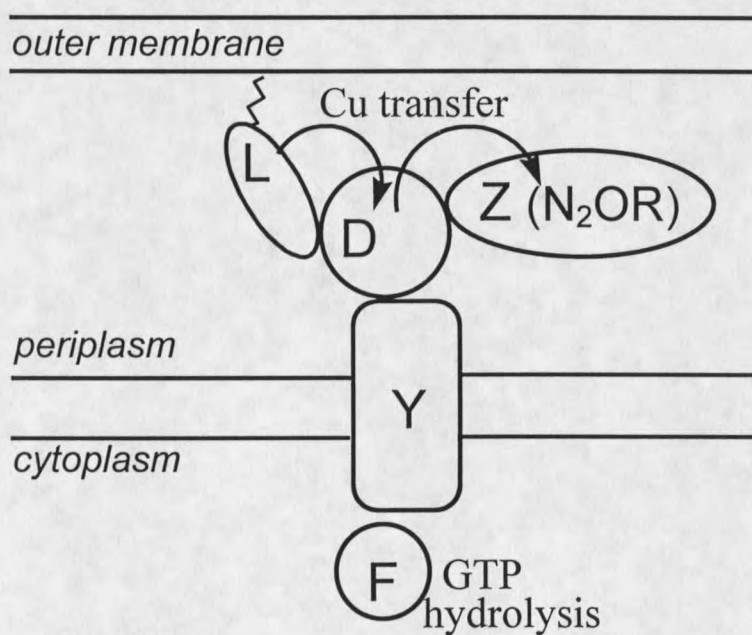


Figure 4. Possible function of NosD, F, Y, and L as part of a copper chaperone system, based on the ABC transporter motif.

Metallochaperones

Metalloproteins were once thought to acquire their metal cofactors by simple diffusion of the metals into pre-formed binding sites. This view was supported by a wealth of studies on enzymes for which the metallocofactors were readily removed and then reintroduced or replaced *in vitro* with no loss of activity (100,101). Indeed, metal substitution has proven to be a powerful tool in understanding the structure and function of many metalloenzymes. However, the discovery that several metalloproteins require the presence of additional proteins for metal site assembly (either *in vivo* or *in vitro*) has caused the diffusion hypothesis to be reassessed (97).

In the simplest case, one metallochaperone is believed to bind the appropriate metal and transfer it to the target protein. In an elegant study on yeast superoxide dismutase (SOD) and its copper chaperone CCS, Rae *et al* were able to show that CCS is required for copper incorporation *in vivo* (102). Moreover, the authors suggest a reason that metallochaperones are necessary in eukaryotic cells – the intracellular concentration of “free” copper is estimated to be less than one molecule per cell. This is probably due to the presence of metallothionein, a metal-sequestering protein found in most eukaryotes and a few archaea (103,104). Metallothionein is thought to protect cells from the potentially harmful effects of unbound, redox-active metals. The crystal structure of apo-CCS was recently determined (105). One domain resembles SOD itself, while another has a fold similar to the copper transport protein Atx1 (106). In both CCS and Atx1, copper is proposed to be ligated by two

Cys residues, although the potential involvement of other protein-derived ligands has not been ruled out (105,107).

For other metalloenzymes more than one auxiliary protein is required for metal site assembly (97). Some of these proteins are responsible for synthesizing bioinorganic cofactors such as hemes, corrins, and iron-sulfur clusters. In other cases, the roles of the accessory proteins are not yet understood. The best studied enzyme in this category is the nickel-containing enzyme urease (108). Urease requires 3-4 accessory proteins for nickel insertion and the post-translational carbamoylation of an active-site lysine residue. Complexes of urease with the auxiliary proteins UreD, UreF, and UreG have been isolated, but the roles of these proteins have not been elucidated. Remarkably, UreE, the protein identified as the nickel chaperone to urease, has not been observed to bind to any of the aforementioned proteins. Yet deletion of *ureE* results in the production of apo-urease, unless the intracellular nickel concentration is increased (109). This suggests that UreE may be needed only when nickel is not readily available from the environment.

Research Goals

In order to better understand the structure and function of the accessory Nos proteins, the *nos* gene cluster of *A. cycloclastes* was cloned and sequenced. At the onset of the project, the complete sequence of the *nos* cluster was available only for *P. stutzeri*, although four complete clusters are currently known. *A. cycloclastes* was an attractive organism for study because it contains a CuNiR, in contrast to the *cd₁NiR* found in *P. stutzeri*. Thus, it was

thought that *A. cycloclastes* might represent an evolutionary distant cousin of *P. stutzeri* among denitrifiers and would complement the pioneering work of Walter Zumft and coworkers on denitrification in the latter organism. A homologous over-expression system for N₂OR was developed, and the accessory proteins NosD and NosL were heterologously expressed in *E. coli*. Characterization of recombinant NosL has revealed that it is a Cu(I) protein. Moreover, NosL contains a novel type of biological copper site and may belong to the recently identified class of proteins called metallochaperones.

SEQUENCE ANALYSIS OF THE *NOS* GENE CLUSTERIntroduction

The development of nucleotide and protein databases has had an enormous impact on the field of protein biochemistry. The plethora of sequence information has led to the development of programs for predicting protein structure and function. These have become an indispensable part of the methodology used to investigate biochemical questions. While theoretical comparisons do not replace the need for experimental research, such predictions often provide a framework for considering a scientific problem, and may be used to guide experimental design.

The impact of sequence comparisons on the denitrification field has been substantial. Until 1998 only one *nos* cluster (from *P. stutzeri*) had been entirely sequenced (82,92,96,98). These initial data provided much information about the possible structures and functions of the Nos proteins. To enhance and refine the available information, the sequence of the *nos* cluster of *A. cycloclastes* will be presented along with the *nosR* sequence of *S. meliloti*. (The *S. meliloti* data completes a partial sequence published by Chan and coworkers (81).) In addition, *nos* sequence data from *B. japonicum* and *Rhodobacter sphaeroides* became available through Genbank over the past year. This will be included in the comparative analyses.

Analysis of the *P. stutzeri* *nos* cluster suggested that NosR is an integral membrane protein that contains six transmembrane α -helices (92). This led Zumft *et al* to construct *lacZ* and *phoA* fusion proteins, which were used to determine the orientation of the N- and C-terminus of NosR in the cytoplasmic membrane (77). The data indicated that a potential ATP/GTP binding site identified by homology studies (92) was located on the periplasmic side of the inner membrane, and consequently was probably not physiologically relevant. It will be shown that the topology of NosR is likely to be conserved among denitrifiers, but that the erroneous ATP/GTP binding site is not. This demonstrates the utility of obtaining the sequence of a protein from more than one source. While comparisons of one protein sequence against a database are certainly useful, homology studies within a protein family often help discern those motifs and topologies that are truly indicative of a specific structure or function, from those that are coincidental.

Comparison of the *P. stutzeri* NosR sequence with a database of known protein motifs suggested that NosR was a metalloprotein (92). Two Cys clusters located at the cytoplasmic, C-terminal end of NosR have the same spacing as found in several bacterial ferrioxins, where the cysteines serve as ligands to an iron-sulfur cluster. It will be shown that this iron-sulfur binding motif is conserved among the four known NosR sequences, along with another potential metal-binding motif.

Transposon mutational studies in *P. stutzeri* indicated that the proteins NosD, NosF, and NosY were necessary for the bioassembly of the N_2OR catalytic copper site (96), while

NosL was not (98). Again, homology studies proved useful in constructing hypotheses about how these proteins might act to incorporate copper into N_2OR . As a whole, the proteins resembled components of ABC transporter systems (99). These systems typically function in the transport of small molecules or proteins across a membrane by a mechanism that requires ATP or GTP hydrolysis. Further analysis of the individual protein sequences predicted that NosD and NosL are periplasmic proteins (96,98). Additionally, the signal sequence of NosL was predicted to be cleaved by a Type II signal peptidase (98). This peptidase recognizes a specific motif and cleaves upstream of a cysteine to which a glyceride-fatty acid is attached (110,111). The lipid moiety anchors the protein into the outer membrane. A CysXaaXaaCys motif is also present in *P. stutzeri* NosL, which suggested that NosL might function as a disulfide isomerase (98). It will be shown that the ABC transporter motif, the predictions of cellular location, and the lipid anchor motif for NosL are conserved among five known sequences. However, the disulfide isomerase motif noted in *P. stutzeri* NosL is not conserved. Additionally, the comparisons of five NosD and NosL sequences suggest that these proteins may function as metallochaperones. The ability of NosL to bind stoichiometric copper will be documented in the next chapter.

Prior to 1999, N_2OR was the only Nos protein that had been isolated and characterized. Biophysical investigations revealed that N_2OR is a homodimer containing two dinuclear copper centers (39-44). One copper site, Cu_A , is thought to function as the electron transfer center. This site has been spectroscopically characterized and is homologous to the

electron transfer center in cytochrome oxidase (COX) (46-52). In fact, comparative spectroscopic investigations of the two proteins led to the initial proposal that Cu_A was a dinuclear site bridged by one or more dithiolate (Cys) ligands (46-50). This model was confirmed when the crystal structures of two COX proteins were solved (51,52). At this point, sequence comparisons between and among N_2OR s and COXs proved especially insightful. Data will be presented that identify probable ligands to Cu_A in N_2OR . Moreover, the homology studies suggest that a His residue located in the Cu_A domain may be a ligand to the catalytic copper center. A model of the Cu_A domain in N_2OR based on the COX structure will be presented. While the overall sequence identity between the Cu_A domains of COX and N_2OR is low, two factors indicate that homology modeling is valid in this case. First, a comparison of the COX Cu_A structure with that of blue copper proteins, which also function in electron transfer, shows conservation of the protein fold (52,112). Yet the primary sequence homology between the COX Cu_A domain and blue copper proteins is quite low. This suggests that a specific protein fold has been evolutionarily conserved among copper-mediated electron transfer domains and is therefore likely to be present in N_2OR . Second, the spectroscopic signatures of metal sites in proteins are usually quite sensitive to minor perturbations in their coordination spheres (113). As the spectroscopic properties of Cu_A from COX and N_2OR are very similar, it is plausible that the protein environment of this dinuclear center in N_2OR will closely resemble that of COX.

Experimental Methods

Isolation of *nos* DNA. A 3 kbp *Eco*RI - *Sph*I region that contained *nosR* from *S. meliloti* was identified from the map of the previously described plasmid pWM4 (81). This region, which had been partially sequenced, was subcloned into pUC19 by digesting pWM4 with combinations of *Eco*RI, *Sph*I, *Ava*I, *Pst*I, or *Hinc*II restriction enzymes.

The *A. cycloclastes nos* cluster was isolated from genomic DNA provided by Dr. John Bollinger, who also provided plasmids pPP3 and pHH92, which contain part of *nosZ* and downstream genes (59). The remainder of the *nos* cluster of *A. cycloclastes* (*nosR* and the 5' end of *nosZ*) was cloned using a 540 base pair probe obtained by PCR. The PCR primers designed for this probe were based on the *nosZ* sequence determined from pPP3 (5'CATCAGTTCGCGCATCGAG) and on a 19-nucleotide stretch of *nosR* that is conserved in *P. stutzeri*, *P. aeruginosa*, and *Pa. denitrificans* (5'TGTTCGACTGGCTCAAGCG). Sequencing of the PCR product identified its relation to both *nosR* and *nosZ*. This product was used as a probe in isolating a 10 kb *Eco*RI fragment containing the *nos* cluster from genomic DNA, using standard Southern and colony hybridization techniques. The *Eco*RI fragment was subsequently cloned into pUC19 to produce plasmid pEE19. *E. coli* strain INV α F' (Invitrogen) was used as the host for all pUC19-derived plasmids and cultures were grown in LB medium containing either 60 μ g/mL ampicillin or 100 μ g/mL carbenicillin.

DNA Sequencing and Analysis. Sequencing was performed by the dideoxy chain-termination method (115) on pUC19-derived plasmids. (Some sequencing was provided by Ms. Laura Nelson using single-stranded DNA derived by subcloning into M13mp18 and M13mp19 with *E. coli* strain JM101 serving as the M13 host.) The end labeling-cycle sequencing method ($[\gamma^{33}\text{-P}]\text{ATP}$) with either the *fmol* or the Silver Sequencing kit (Promega), as well as the label-incorporation method ($[\text{S}^{35}]\text{dCTP}\alpha\text{S}$) with Sequenase 2.0 (Amersham) were used. Sequencing primers were either vector-derived (M13 forward and reverse) or designed from internal sequences in the genes of interest. Both coding and non-coding strands were sequenced at least once. Unless specified, sequences were assembled and compared using the GCG Sequence Analysis Package, V. 8 (Madison WI). The numbering of protein sequences is given with respect to the residue position in the multiple alignments shown in the figures.

Protein Characterization and Expression. The N-terminal amino acid sequence of mature N_2OR from *A. cycloclastes* was determined on an Applied Biosystems 475A Protein Sequencer by Dr. Gerhard Munske at Washington State University, Pullman WA.

Molecular Modeling. A model of the C-terminal, Cu_A domain of N_2OR from *A. cycloclastes* was constructed using Homology and Discover software (Biosym Technologies), using the Cu_A domain of bovine heart cytochrome oxidase (51) (PDB Accession #1OCC) as a template.

During energy minimization the positions of the assigned Cu_A ligands of N₂OR (H621, C656, C660, H664, M667, and the main chain of W658) were held fixed. The structure was refined to a RMS derivative of 0.002 kcal mol⁻¹ Å⁻¹, using the conjugate gradient algorithm with the Consistent Valence Forcefield (CVFF).

Results

The *nos* Gene Cluster. The organization of the *nos* cluster of *A. cycloclastes* (Figure 5) is identical to *S. meliloti* and *B. japonicum* but differs from *P. stutzeri* in the region that codes for *nosX*, which has not been found in the latter organism (91,93). Interestingly, a TBLASTN search of the *Neisseria gonorrhoea* genome (116) indicates the presence of a *nos* cluster containing *nosR*, *Z*, *D*, *F*, *Y*, and *L*. This organism also contains a *nosX* cognate on another contig, suggesting that other denitrifiers may have this non-contiguous arrangement of *nos* genes.

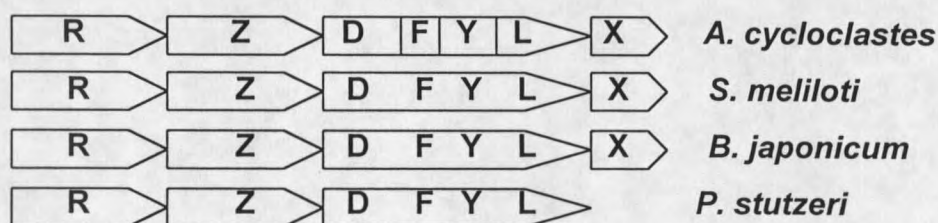


Figure 5. The arrangement of genes in the *nos* gene clusters of four denitrifiers.

Considerable variation is apparent in the sizes of the *nos* intergenic regions from *A. cycloclastes*, *S. meliloti*, *B. japonicum* and *P. stutzeri*. For example, in *B. japonicum* the stop codon of *nosR* overlaps the start codon of *nosZ*, but in the other organisms the genes are separated by an intergenic region ranging from 26 to 74 nucleotides. (See Appendices B and C for nucleotide sequences of *nosRZDFYLX* from *A. cycloclastes* and *nosR* from *S. meliloti*, respectively.) In *P. stutzeri* the *nosZ* and *nosD* genes, which are located in different transcriptional units, are separated by an intergenic region of 122 nucleotides. This distance decreases to 3 nucleotides in *A. cycloclastes* (Appendix B) and *B. japonicum*, while in *S. meliloti* the termination codon of *nosZ* and the initiation codon of *nosD* employ the same adenosine. There is also significant variation within each *nosDFYL* complementation group. Whereas in *S. meliloti* and *B. japonicum* the termination codon of each gene overlaps the start codon of the following gene, there is a 29-nucleotide intergenic region between *nosY* and *nosL* in *P. stutzeri*. While *nosD* and *nosF* in *A. cycloclastes* share 35 nucleotides, *nosF*, *nosY* and *nosL* overlap only at their initiation/termination codons. The *nosX* transcriptional unit begins 31 nucleotides downstream of *nosL* in *S. meliloti*, 12 nucleotides downstream in *B. japonicum*, and in *A. cycloclastes* this intergenic region is only 4 nucleotides in length.

Potential binding sites for the transcriptional regulator FNR (TTGATN₄ATCAA) are found about 30 - 90 nucleotides upstream of the coding sequences of *nosR* from both *A. cycloclastes* and *S. meliloti* (Appendices B and C). Also noted are several conserved FNR ½-boxes (ATCAA) in the promoter region of *nosZ* in *A. cycloclastes*. In *S. meliloti* there is a

degenerate σ^{54} binding motif (CTGGNAN₃₋₈TGCAA) upstream of the *nosR* initiation codon, although its associated transcriptional start site falls within an FNR box. Degenerate σ^{54} motifs are also evident in the promoter regions of the *nosR*, *nosDFYL*, and *nosX* transcriptional units in *A. cycloclastes*. The structural stem-loop features associated with rho-independent transcriptional termination are also discernible. The data are consistent with the *nos* gene cluster being divided into four transcriptional units, *nosR*, *nosZ*, *nosDFYL*, and *nosX*, as determined for *S. meliloti* (81,109).

NosR. The predicted amino acid sequences of NosR from *P. stutzeri*, *S. meliloti*, *B. japonicum*, and *A. cycloclastes* are similar in size, amino acid content, and secondary structure, as indicated in Figure 6. Although there are three potential initiation codons located within a 60-nucleotide span in *A. cycloclastes*, their respective ribosome binding sites suggest that either the second or third codon is the most likely start of NosR (Appendix B). Of these, translation beginning at the second ATG produces a protein sequence that is most consistent with the other deduced sequences, especially considering the length of the N-terminus before the first transmembrane helix. In NosR from *P. stutzeri* a second possible initiation site has been noted 123 nucleotides upstream of the start codon used in the sequence represented in Figure 6 (77,92). Homology with the other NosR proteins indicates that this alternative start site is improbable, as it increases the size of the periplasmic N-terminus by 41 amino acids.

1 I 70
 BJ .MLQORTHVVPDFLTQAARALFALAFLCVLSSSENTASAADINCYLSKIAPADFFAGADRFPGPPQGEPPPIIP
 SM ..MYQAAIHKNLFKRLRLGLVALCIAMLLAGPSMA..AGQISNYLQKVQPQQIFPGATRFGEVTGDPPIAP
 AC .MLRQRFAGKHFGMKLFRLVLLCL.SLWGAVSASAASVIOEFLPTTTASDLVPGADGFGAVRDDVAAAP
 PS MASREIWSLAGTFSTWVKGFFILLILSV.....CASSIOAKEYAAEQQRDKFFPGASLSAAEGDYQVRT
 C -----I-----
 71 140
 BJ AYRGDQLQGFVYLNDSDFANAVGYSGKPIQMLVGIDPKGVLTLGLKLVHKEPIVLVGIPEKRIEAVNKLI
 SM VFRGESLLGYAYLNSDVTSSVGYSKPIHIVVGIDPKGVVRLGLVDHKEPIVLIGIPEAKVVASVNALI
 AC VLKGGQTVAVVFTSDFVGTGYSKPIHTLVAVDKDAKIVGVRLVKHSEPIVLIGIPDAKVKALVAGYA
 PS ITKGDEVLYAFQSIIRVTDMPAYSGBKPINMQILLDEGVIVDAYMLEHHEPIVLIGIPEQKVHDFNANYS
 C ---G-----YSGKPI-----D-----H-EPIVL-GIP-----
 141 210
 BJ GAD...LGSVARGAAPAPQVDIVSGATVTVLVIGDSIVRSATKLIRSGRLGAQGSFVAAAGPRAAKAID
 SM GKD...LGRVSAGVEGPPQVDIVSGATVTVLVMGDSIVRSALKLIRGNRFGADA..VAPEQSSEIRKVD
 AC GLD...LVAAEKSGGTTHEVDIISGATVTVMVIDDSIVRSGLKVARA..LGLGGLTPEADNAGPRFEID
 PS GIHVDQRRVVGRSSDKSAVTVDVATGATVTVMVINEIVMRAAHTVAVDLGLVEAGATARPKPALVREDV.
 C G-----VD---GATVTV-V-----R-----
 211 280
 BJ AAKSEVRDQWTLVGDSVRRLTSLVADINQAFEKSGNAVAAARPEEGAPDDSFIDLYADVAVPTIGRSL
 SM LLRTGTSDWETLVGDGVSRLTLVGEVTEAFRQAGQPAADRPETHSPGDRFIDLYIAPVSVPIGRSL
 AC PAATAPADWVEAEGDGLRLSLDVAQVNAAFAGHPDERARERAIQAAPETTFIEMQAALVSVPSIGRAF
 PS ...FQPTSWTELVGNGAIRRMHLTRGOVDDAF.KGTEAEGVDVAAAEQRDETFFIDLYATHLNPTIGRNL
 C -----W---G-G--R---L-----AF-----FI-----P-IG---
 281 350
 BJ LGDESYGRLAGRLKPGQAFVAVAGAGRYSFKGAAYVRGGIFDRIELIQETNSIRFRDRDTRLGSLAADG
 SM LGDSRYEQMRALKPGEEAIVVAGDGAYSFKGSGYVRGGIFDRIELIQDQGLRFRDRYHTRLPTLAASG
 AC LGAAEEKNLQAWLGEGEHAIAVVGRLYSFKGSGYVRGGIFDRIELIQDDVSVFRDSDHRLGALISLGA
 PS LGERQYADLMANLKPGEHAFVLANGEYSFKGSGYVRGGIFDRIELIQDQGLRFRDRYHTRLPTLAASG
 C LG-----L--G--A--V---G-YSFKG--YVRGGIFDR--L-Q-----FRD---RL-----G
 351 420
 BJ APDFPEIALFIVPADFAFDATEPWALQLLVQVRVVGARDKAWVTFDLGYTLPETYLKREAPVRPAAAAAAP
 SM APRLREIALFVVPADFGFDIAAPWELQLLVQRTAVGRDKAVLPYNLSYVLP.....DAHVTVEAAAAAPS
 AC APEYNEMDLFRIKAGLGFDPKPFRIQLLAQREVGPVEKVFHTFDLGYQLPQKYLR.....AVAP
 PS MPEFFEMAIFTAREQYRFDPGSPWNLELLVRRQVGPVESIFTSFEMPYVMPEEYIERVPLTAEELAAIEE
 C -P---E---F-----FD---P---LL--R-----Y--P-----A-----
 421 II 490
 BJ TAKSSPAAPEARVPTTDEDPLWMRIWRADALKITGAMLALGALTLIFFQNFVRRPAVYTWVRRFYLL
 SM PVEPSIAEQATPAEPIPEGEALWVKMWEMNRVSAVVTMAAVLVLTLIFFQDNLVKRPALFAVVRVYLL
 AC PPVKDVAADVADRDEETAQSD.LWKRIWDDSKLKIIVLGAMLVLTGVFFFQTYAVRNARAFYIFRIIFLT
 PS A.....NRPLWVNIWYQKSFOVGVLVALALLTVILFLQDKFTQHPNFKRLRHGYLV
 C -----LW---W-----LT---F-Q-----R---L---
 491 III IV 560
 BJ FILVWLGWYANAQLSVVNVLTFTNSLLTGFSWEYFLSAPLIFLLWASIAAGLLFWGRGPFCEGWLCPFGAL
 SM FTLVWLGWYANAQLSVVNVLTFFNSLVTGFHWEFFLSAPLVFLLWGSVAAALLFWGRGPFCEGWLCPFGAL
 AC VTLVFLGWYANAQLSVVNLMAFLGSLVNGFSWQAFLLDPLTFILWFVAAALLFWGRGAYCGWLCPFGAL
 PS FTVFIGWYALGQLSVVNVLTFFVHALVQDFRWELFLTDPIVIFILWVFTAASILLWGRGVFCGWLCPFGAL
 C ---V---GWYA---QLSVN-----L---F-W---FL---P---F-LW---AA---L-WGRG---CGWLCPFGAL
 561 V 630
 BJ QELLNNAAKALKVPQYRVPWGLHERLWPIKYIIFLGLFGMSLYSTAFAEQLAEVEPFKTSIVLKFAREWP
 SM QELTNNVARWLKVQYRVPWGLHERLWPIKYIIFLGLFGLSLYSLALAEEMFAEVEPFKTAIILKFAREWP
 AC QELTNQIARKLRIPQWTLPWGLHERLWPIKYMIFLGLFGVSLMSVEQAELHAEIEPFKTAIILKFIRAWP
 PS QELINEAARKLRIPQYDLFPVSVHERLWAIKYIVLLVLFGISLESMMMAEKAAEIEPFKTAITLKFDRQWW
 C QEL-N--A--L--PQ---P---HERLW-IKY---L-LFG-SL-S---AE---AE-EPFKT-I-LKF-R-W-

continued on the next page

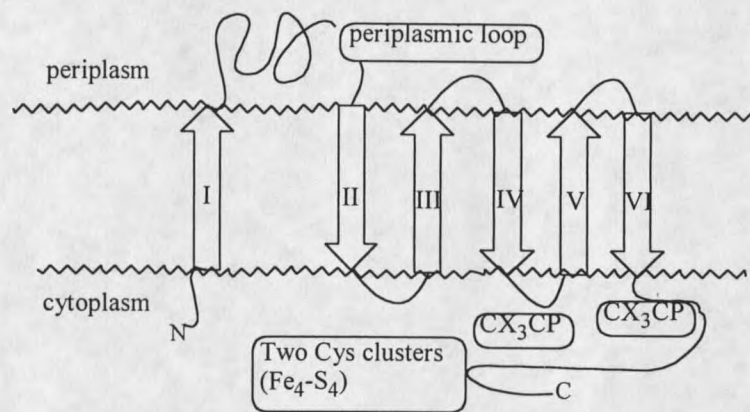
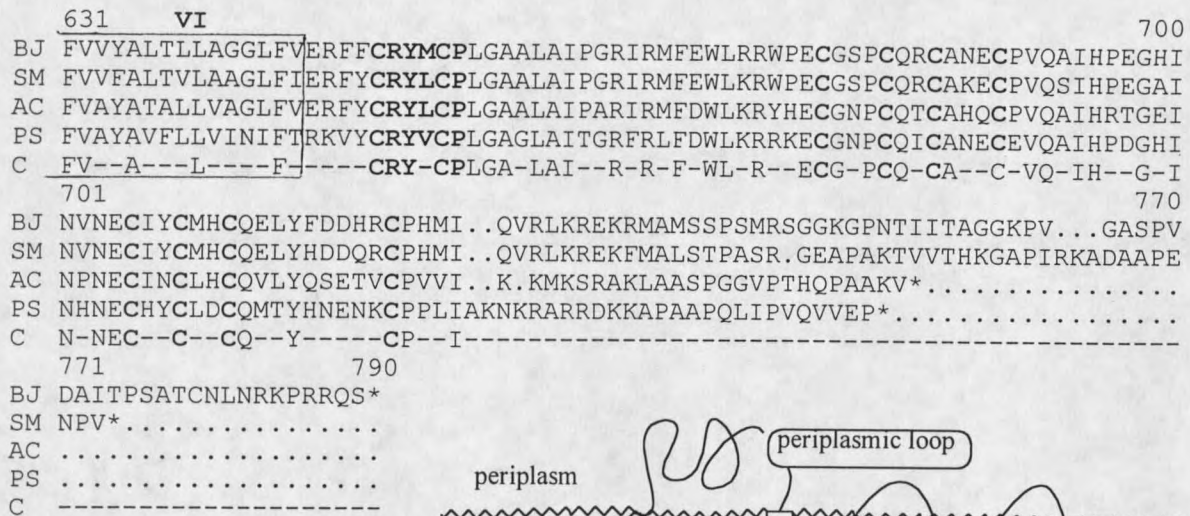


Figure 6. Comparison of the deduced amino acid sequences of NosR from *B. japonicum* (BJ), *S. meliloti* (SM), *A. cycloclastes* (AC), and *P. stutzeri* (PS). Predicted transmembrane helices are boxed and numbered; potential metal-binding sites are indicated in bold; C, consensus sequence. Inset: a cartoon representation of the predicted structure of NosR.

Transmembrane helix prediction (TMPred (117)) of the four NosR sequences, when combined with the *phoA* and *lacZ* fusion experiments performed on *P. stutzeri* NosR (77), indicate the presence of 6 conserved transmembrane helices with a large periplasmic loop between the first and second helices, as depicted in Figure 6 (inset). NosR is predicted to be an integral protein of the cytoplasmic membrane. Interestingly, there is only ~30% similarity among the four organisms in the periplasmic loop portion of the protein, which displays three

short, homologous regions. In contrast, 129 of 296 amino acids are conserved in the C-terminal helix domains, including several potential metal-binding motifs. One such motif consists of two CX₃CP clusters, one located at the C-terminal end of Helix IV and the other after Helix VI. Topological analysis implies that both are on the cytoplasmic side of the membrane. In addition to this potential metal site, there are two [Fe₄S₄]-cluster binding motifs in the cytoplasmic C-terminal tail. The pattern of spacing between Cys residues in this region is similar to that found in some bacterial ferridoxins (92).

N₂OR (NosZ). The alignment of the deduced amino acid sequence of N₂OR from *A. cycloclastes* with six published sequences is shown in Figure 7. N₂ORs have been predicted to have unusually long periplasmic signal sequences and this has recently been confirmed for *P. stutzeri* (118). The start of the mature *A. cycloclastes* protein was determined by sequencing the N-terminus and indicates a signal sequence of 46 amino acids, confirming the cleavage site predicted by the (-3), (-1) method of Von Heijne (119). The mature N₂ORs range from 65,000 to 67,350 Da per monomer, with an average pI of 5.9.

N₂ORs contain 9 - 11 Cys residues, depending on the source. Of these, C656 and C660 presumably serve as Cu_A ligands. A Cys residue at position 190 that had been previously considered a likely Cu_{Cat} ligand (79) is not conserved in *A. cycloclastes*, which has Thr at this position. Although no other non-Cu_A Cys residues are strictly conserved, both *A. cycloclastes* and *Pa. denitrificans* have Cys at position 185, while C225 is found in all

115
AC MESKE-----HK-GLSRRALFSATAGSAILAG-----TVGPAALSIGAAGLATPARA---ATGADGSVAPGKLDYYGFWSSGQTGEMRILGIPSMRELMRVPVFNRC
PD MESKQ-----EK-GLSRRALLGATAGGAAGAVAGFGGRLALGPAALGLGTAGVATVAGSGAALAASGDGSVAPGQLDDYYGFWSSGQSGEMRILGIPSMRELMRVPVFNRC
BJ MSDSD-----NIKGVSRRTLLGTTAAAGVGLAGGAVVTKDGAGFVSTADAQTKSAAPKAPPARPAVQKTEVAPGELDEYYVFFSSGQSGEMRIIGLPSMRELMRVPVFNRC
RM MSNEE-----TKMRLNRRQMLGTTAFMAA-AGAVG-----AGGALTLS-GGTATPARAQETSGSSYE--VKPGELDEYYVFFSSGQSGEIRIVGAPSMREMMRIPVFNRC
PA MSDKQTDKD---EKTGLSRRGFLGASAI,TRS-----AVAASGLVGGVMTRDSWAAAKEAQKIHVAPGELDEYYGFWSSGGHQGEVRLVGVPSMRELMRIPVFNVD
PS MSDKDSKNTPOVPEKLGLSRRGFLGASAVTGA-----AVAATLGGAVMTRESWAQAVKESKQKIHVGPGLDDYYGFWSSGGHQGEVRLVGVPSMRELMRIPVFNVD
AE MSKEK---ASIGNPGGGIGRRQFLGTAALAGLVVACTDKGAAPAAAAVGAPASGAHG---AAHGAGASVHLKPGELDTYYGLWSSGGHTGDMRVLGLPSGREILRIPCFVNDAL
C M.... RR...:A | : : : : : PG:LD.YY .|S:G..G|.R|:G PS RE:|R|P.F :

116 230
AC TGWGQTNESIRIHQRTMTEKTKKQLAANGKKIHDNGDLHHVHMSFTDGKYDGRYLFMNDKANTRVARVRCDVMKTDAILIIPNAKGIHGMRPQKWPR-----SNYVFCNGEDE
PD TGWGQTNESVRIHERMTSERTKKFLAANGKRIHDNGDLHHVHMSFTEGKYDGRFLFMNDKANTRVARVRCDVMKCDAILIIPNAKGIHGLRPQKWPR-----SNYVFCNGEDE
BJ TGWGQTNESLKVLTGEGMPATREFLKNRGG-TFMNGDLHHPHVSFTDGTGGRYAFMNDKANSRVARVRCDVMKCDKIIELPNQHTVHGLRLQKYP-----TGIVFCNGEDG
RM TGWGQTNESRKVMTEGLLPETVEFLKDQGG-LYLNGDLHHPHPSFTDGTGGRYLYANDKSNRSRVCRIRLDVMKCDKIIELPNQHTVHGLRLQKYP-----TGIVFCNGEDA
PA TGWGLTNESRHILGDTAK-----FLNGDCHHPHISMTDGKYDGKYLFINDKANSRVARIRLDIMKCDKITTPNVQAIHGLRLQKVPH-----TKYVFCNAEFI
PS TGWGLTNESRHIMGDSAK-----FLNGDCHHPHISMTDGKYDGKYLFINDKANSRVARIRLDIMKCDKITTPNVQAIHGLRLQKVPH-----TKYVFANAEFI
AE VGWGITNESKKVMGA-----RPDGTLLRYTVADTHHTHASYKDGNYDGRYAWVNDKINSRIARIRLDYFICDKITELPNVQGFHGFIPDKRDPVDTKINYTTRVFCGGEFG
C GWG TNES .| |D HH H S..|G.YDG|| |NDK N|R| R|R.D.. D : ..PN ..HG| .K | VF .|E

231 345
AC APLVNDGS--TMTDVATYVNIFTAVDADKWEVAVQVKSGLNLDNCADYEGKWAFSTSYNSEMGMTLEEMTKSEMDHVVFNIAEIEKAIKAGQYEEING--VKVVDGRK--EAK
PD TPLVNDGT--NMEDVANYVNVFTAVDADKWEVAVQVLVSGNLDNCADYEGKWAFSTSYNSEKGMTLEEMTKSEMDHVVFNIAEIEKAIKAGQYEEING--VKVVDGRK--EAS
BJ VPLPNDGK--VLDNPKEYHSIFTALDGDGMKVAWQVMVSGNLDNDVSDYQGYCFSTCYNAEEGVTLAEMTANEQDWVIFNLKRIEEAVKKGDFEKEMNG--VPVIDGRK--G--
RM VVPVNDGK--TMGDKNSYQAIFTAVDGATMEVAVQVMVDGNLDNDVADYQGYCFATCYNSEEGFTLADMMAEQDWVIFNLKRIEEAVKAGDYKEIGG--VPVLDGRK--G--
PA IPHPNDGKVFDLEDENSF-TMYNAVDAETMEVAFQVIVDGNLDNTDADYTGRTAATCYNSEKAFDLGGMNRNERDWVVFDDISAVEKEIKAGRFTITLGDSKVPVVDGRKKDGD
PS IPHPNDGKVFDLQDENSY-TMYNAIDAETMEMAFQVIVDGNLDNTDADYTGRTAATCYNSEKAFDLGGMNRNERDWVVFDDIHAVEAAVKAGDFITLGDSKTVPVLDGRKKDGD
AE IPLPSAPT---EDAGKYRSFTCVDAETMAVRWQVLIDGNCDLVATSYDGKLAATNQYNTENGAFEDMMSAERDACVFFNIARIEAAVQAGKFYTYGDSKVPVVDGTQAANKD
C P . : . | . | : | . |D| . : : |QV |.GN.D : .Y.G|. : . YN:E | . . M E D .V.F| : |E | G. | . . V:DG.:

346 460
AC SL--FTRYIPIANNPHGCMAPDRKHLCAVAGKLSPTVTVDVTKFDALFYDN-AEPRSAVVAEPELGLGPLHTAFDGRGNAYTSLFLDSQVVKWNIDEAIRAYAGEK INPIKDK
PD SL--FTRYIPIANNPHGCMAPDKHLCAVAGKLSPTATVLDVTRFDAVFYEN-ADPRSAVVAEPELGLGPLHTAFDGRGNAYTSLFLDSQVVKWNIDEAIRAYAGEK VDPIDDK
BJ SP--YTRYVPVSNPHGMNTAPDGIHIVAAGKLSPTVTVMVRLFNQLFDDK-IKPRDVVAEPELGLGPLHTAYDGKGNAYTTLFLDSQVVKWNIDLAKRAFKGEK VDPITQK
RM SP--YTRYVPVPSPHGINTAPDGIHVANGKLSPTVTVFDVRKFDDLFDDKI-QARDTVVAEPELGLGPLHTAYDGKGNAYTTLFLDSQVCKWNIDEAKRAYAGEK VDPIRHK
PA SV--VTRYIPVKNPHGLNTSTDGKYFIANGKLSPTCSMIAIDLLPDLFAGKLKDPDRDVVGEPELGLGPLHTTFDGRGNAYTTLFLDSQVVKWNMEEARRAYKGEK VNYIKQK
PS SK--FTRYVPVKNPHGCMNTSSDGKYFIAAGKLSPTCSMIAIDKLPLFAGKLADPRDIVGEPELGLGPLHTTFDGRGNAYTTLFLDSQVVKWNMEEAVRAYKGEK VNYIKQK
AE PKTALTAYSVKPNPHGVNASPDQYFICAGKLSPTATVIELSRVLGWFDGKQEKLLDDAIVAEVELGLGPLHTAFDGRGNAYTTLFLDSQVVKWNLDAAIKFHKGDKNAYVVD
C : .T Y| : |...PHG.N : :D : . . GKLSP | : . : . . F . . . |V|E ELGLGPLHT : |DG|GNAYT|LF:DSQ:KWN:| A | : G|K . | . |

continued on the next page

461 575

```

AC LDVQYQPGHLKTMGETLDAANDWLVLCKFSKDRFLNVGPLKPENQDLIDISGDKMVLVHDGPTFAEPHDAIAVSPSILPNIRSVWDRKDPLWAETRKAQAEADVIDEWTEAV
PD LDVHYQPGHLKTMGETLDATNDWLVLCKFSKDRFLNVGPLKPENQDLIDISGDKMVLVHDGPTFAEPHDAIAVHPSILSDIKSVWDRNDPMWAETRAQAEADGVIDDNWTEEV
BJ LDVHYQPGHNHSSMGQTKEADGKWLISLNKFSKDRFLNVGPLKPENQDLIDISGDKMVLVHDGPTFAEPHDAIVHRSKINPIS-IWDRADPMFADAVKQAKADGINLEA-DSKI
RM LDVHYQPGHNHTSMGQTKEADGKWLISLNKFSKDRFLNVGPLKPENQDLIDISGDKMVLVHDGPTFAEPHDCIMARRDQI-KTRKIWDRNDPFFAPTAMAKKDGINLEE-DNKV
PA LDVHYQPGHLHASLCETSEADGKWLVALSKFSKDRFLPTGPLHPENDQDLIDISGDKMVLVHDGPTFAEPHDCIMARRDQI-KTKKIWDRNDPFFAPTAMAKKDGINLDT-DNKV
PS LDVHYQPGHLHASLCETNEADGKWLVALSKFSKDRFLPVGPLHPENDQDLIDISGDKMVLVHDGPTFAEPHDCIMARRDQI-KTKKIWDRNDPFFAPTAMAKKDGINLDT-DNKV
AE LDLQYQPGHVNASQSETVAADGKYLAVGCKFSKDRFLPVGPLHPENEQLIDISGDKMVLVHDGPTFAEPHDCIMARRDQI-KTKKIWDRNDPFFAPTAMAKKDGINLDT-DNKV
C LD:YQPGH.:|T A .|L KFSKDR|L GPL.PEN|QLIDISG. M L: D P |EPHD . . .: .|| D .A . . V
576 679
AC IRDGNKVRVYMTSVAPSFQSPSTVKEGDEVTVIVTNLDEIDDLTHGFTMGNGHVAMEVGPQQTSSVTFVAANPGVYWYYCQWFCHALHMEMRGRMFVEPKGA*
PD IRDGNKVRVYMSSVAPSFQSPSTVKEGDEVTVIVTNLDEIDDLTHGFTMGNGYVAMEIGPQMTSSVTFVAANPGVYWYYCQWFCHALHMEMRGRMLVEPKGA*
BJ IRDGNKVRVYMTSTAPAFGLEQFQVKGQDQVTYITNIDAVEDLTHGFCIVNYGIQMEVAPMATASVSFSADKAGVYWYYCSWFCHAMHMEMKGRMFVEPKSV*
RM IRDGNKVRVYMTSAAPAFGLDDFTVKQGEDEVTVIVTNLDEIDDLTHGFCIVNYGINMEVAPQATASVTFKASRPGVYWYYCTWFCHAMHMEMKGRMLVEAQGA*
PA IRDGNKVRVYMTSMAPAYGLTEFKVKQGEDEVTVIVTNMDQIEDVSHGFVVMVNHGVSMEISPQQTSSITFIADKPGLHWYYCSWFCHALHMEMVGRMMVEPA*--
PS IRDGNKVRVYMTSMAPAFGVQEFQVKGQDEVTVITNIDQIEDVSHGFVVMVNHGVSMEISPQQTSSITFVADKPGLHWYYCSWFCHALHMEMVGRMMVEPA*--
AE FRNGRKVTVKITSQAPAFSLREFKLKGEDEVTLILTNLDKIEDLTHGFAIPKYNVNFIVNPQETASVTFVADKPGVFWCYCTHFCALHLEMRTRMIVEA*---
C .R|G.KV.V:|S AP:|: .F.:K:G||VT: :TN:D ||D:|HGF: .:|. | P T:S||F A.:G::W.YC.:FCHA|H|EM:RM.VE:
Cox sfdsymiptselkpgelrllevdnryvlpmentimrlvsseDvlHswavpslgktdaiPgrlnqtTlmssrPGlyyggCseiCgsnHsfMpivlelvplkyfekwsasml*

```

28

Figure 7. Sequence alignment of N₂OR from *A. cycloclastes* (AC), *Pa. Denitrificans* (PD), *B. japonicum* (BJ), *S. meliloti* (SM), *P. aeruginosa* (PA), *P. stutzeri* (PS), and *R. eutropha* (RA). An arrow marks the start of the mature protein from *A. cycloclastes*. Cu_A ligands are underlined; potential Cu_{cat} ligands are in bold, with a possible semi-conserved Cys indicated by ∇. An alignment of the consensus sequence (C) Cu_A domain with that from COX is also shown, with conserved residues in upper case.

organisms except *P. stutzeri*. An examination of published partial NosZ sequences (Appendix A) shows one or both of Cys185 and Cys190 are present in all known NosZ sequences. Other potential ligands to the catalytic Cu_{Cat} site are 9 conserved His, (especially the HHXH motif at position 155), 3 Met, and 8 Tyr. Two of these, H661 and M671, are located within the C-terminal Cu_A domain identified in other N₂ORs (79,80).

Model of the Cu_A Site in N₂OR. A model of the C-terminal domain of *A. cycloclastes* N₂OR, based on the crystal structure of the Cu_A domain from Subunit II of bovine heart cytochrome oxidase (51), is shown in Figure 8. Although within this domain only 8 of ~125 amino acids are conserved between cytochrome oxidase and N₂ORs (including five of the six copper ligands, bottom line of Figure 7), little steric hindrance between side chains was observed in the model, even in the absence of energy minimization. Whereas the peptide carbonyl of a Glu residue serves as a Cu_A ligand in cytochrome oxidase, this amino acid is replaced by Trp (His in *R. eutropha*) in N₂ORs. Other conserved residues include a proline that is located at a turn in the secondary structure and an aspartate that is thought to shield Cu_A from solvent (51,52). A surface patch of negatively charged residues (Figure 8, white) is present in both cytochrome oxidase and N₂OR: in the former enzyme this region has been proposed to be the docking site for cytochrome *c* (52,120). Two residues, H661 and M671 (Figure 8, hatched), are conserved among N₂ORs but not cytochrome oxidases and are potential Cu_{Cat} ligands. A surface map of the model is also shown in Figure 9.

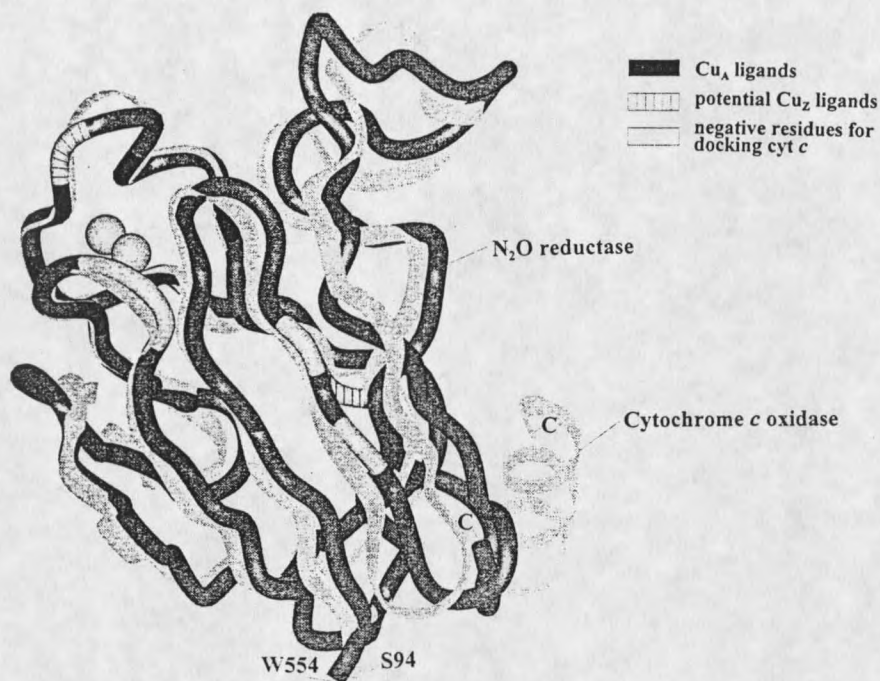


Figure 8. Ribbon diagram showing a model of the Cu_A domain in N_2OR (dark) based on the Cu_A domain of COX Subunit II (light) as template.

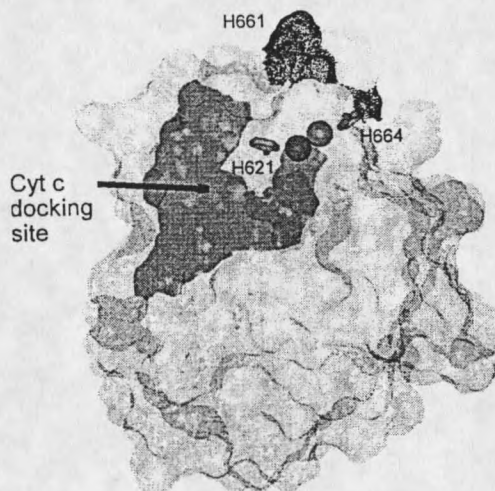


Figure 9. A surface map of the N_2OR Cu_A model, showing the spatial relationship of the purported Cyt *c* docking site (negatively charged patch) to the two Cu_A ligands H621 and H664, and to the proposed Cu_{cat} ligand H661.

A closeup of the modeled Cu_A site in N_2OR is shown in Figure 10. The Cu_A ligands (black) are H621, C656, the peptide carbonyl of W658, C660, H664, and M667. Examination of the positions of the potential Cu_{cat} ligands (hatched) M671 and H661 shows that only the histidine is exposed to the domain surface: M671 is buried within the protein. The side chain of H661 is roughly *trans* to that of C660 and its ring nitrogens are within 9.5 Å of Cu_A . In cytochrome oxidase, the site corresponding to the position of H661 is located at the interface of the Cu_A domain of Subunit II with Subunit I, which contains the catalytic, Fe-Cu metallocenter.

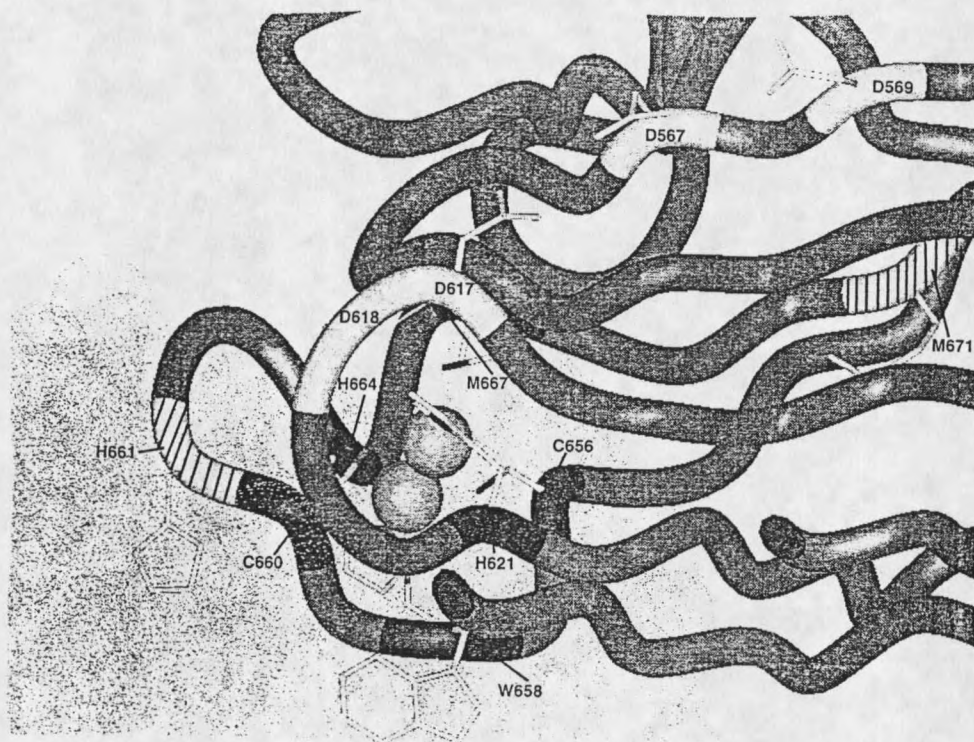


Figure 10. Closeup of the modeled Cu_A site. White, surface Asp residues; black, proposed Cu_A ligands; spheres, copper ions. The position of COX subunit I (surface dots) with Cu_A is also shown.

NosD, F, Y, and L. Of the Nos proteins for which sequences are available, NosD shows the most variation among organisms. As seen in Figure 11, the C-terminal half of NosD from *S. meliloti* and *B. japonicum* contains a 29 amino-acid stretch that is shorter in *P. stutzeri* and absent in both *A. cycloclastes* and *R. sphaeroides*. However, the C-termini of NosD from the latter two organisms are longer than in other organisms, so that, despite the deleted loop, the five NosD proteins are similar in length. Overall the NosD sequences display only 22% identity. Alignment of the five deduced sequences illustrates the conservation of 2 His, 2 Met, and 2 Tyr, which might be metal ligands, and 4 C-terminal Trp residues. For *A. cycloclastes*, a signal sequence of 18 amino acids is predicted (119), resulting in the production of a periplasmic protein of 45,900 Da, with a pI of 6.4.

Analysis of NosF indicates that this protein is highly conserved among *S. meliloti*, *P. stutzeri*, *B. japonicum*, *R. sphaeroides*, and *A. cycloclastes*. NosF is predicted to be a cytoplasmic protein that contains both an ATP/GTP binding site and an ABC transporter motif (Figure 12). NosY appears to be an integral membrane protein (~28 kDa) composed of six transmembrane helices. The hydropathy patterns (Figure 13) of NosYs from the five organisms is highly consistent, although the sequence identity is only 24%.

The homology among NosLs from *A. cycloclastes*, *S. meliloti*, *B. japonicum*, *R. sphaeroides*, and *P. stutzeri* is displayed in Figure 14. Like the previously described protein sequences (91,98), NosL from *A. cycloclastes* contains a Type-II signal peptidase recognition sequence. When cleaved, resulting mature protein contains an N-terminal Cys that is

↓

```

1                               100
AC .....MRLSVLLIGFLLASPLLAERAVMPGAGSLAAAIAGAQP GDVLVLT DGAYDGPVTIDRPLTITGPRGAVVNGLKQGSVVTIAASDVTLTGF
RS .....MMRCLVLALCLLAGPLCAAERTVPAEAGALKAAVAGAAPGDVLILAPGRHEGPVLLDRPLTLLGREGAVVDGGGRGSLVLTVPDVRAGL
BJ .....MLPAALVAAGLSGFASAETL.TAVPGQPLQALLDRARDGDVVELAPGDYNGSIRIDRPLALTGRPGAVLDGGGSGNVVTVTAPDVTVRGV
SM ..MSRPNISAFGMAALAAVILACPVSAATIRKSADGLPLQPVLDRASPGDVIVL.QGEHQGPVTIDKTLTLEGEPGA LVMGNGKGSVITVKA PQSIVRGL
PS MFKAQATFSRYSAAVSLLLLFSGAAQAA.....PQSITTLPLQPDGENRWRLPAGEYQGGFTIEQPMQLRCEPGA VIQSQGGSSLLISAPDVLVEGC
C -----A-----L-G--G-----GA-----G-----A-----G-----
101                               200
AC TVTGSGRINQDL DAGVKIVQGADRARVEKLLVTDNMHGIDVHGGRTIVSGNEIIGTRSARMNERNGNIYVWNSPGTLLQDNIIRYGRDGI FSNASADSI
RS EIRGSGSSHEEIDAGVRLKGA KR PVVEDNRLFGNLVGVDIHGATDARVSRNLIVGREDARMNARGNGIYVWNAPGLIVEENDISLGRDGI FVNVSDRNI
BJ TIRGSGRDLQAMNSGVFLQKTAERATIENNRLVGNLFGVYVHGATGSRVANNLIEGLRGHRLSEAGNGVSLWNAPEVTIADNTFRYGRDGI FTISSRKDR
SM EVRSGSGKDLYGMDSGIFVAQTASGARVEKNTIIGNLVGIYLHGARDSWALGNRIIGLREGRISEAGDGISVWNAPGARVVDNDVSYGRDGI FSKTSKRN
PS TLYEWGSDLTAMDSAVFILPAAERAQISNNRMRGPGFGVFVDGTRDVQVIGNEIDGDAGVRSQDRNGIHLFAVSGARVLHNVHVRNARDGIYIDTSNGNH
C -----G-----A-----G-----G-----N-I-G---R---G-G-----N-----RDGI---S-----
201                               300
AC YRRNIMRDLRF AVHFMYTRNTEVSDNISIGNHLGF AIMFSNR AKILNNLSLGDREHGLMLNYANNADVSGNLV.....
RS FRNNRMRNLRF AIHYMYAHSEVTGNISIGNHLGYALMYSHHLTVTDNLSLRDRDHGVMNLNYVNNAEVARNLV.....
BJ FVGNRF EQVRFAVHYMYTNDSEVSGNVSIGNHVG YAIMYSNRLIIRDNISDRDRDHGLLFNYANYGRDRNRVTGGL.....LDTCPAGSEEGPDDER
SM FRGNRFRELRF AVHYMYTNDSEISDNVSTGN AVGYAIMYSDRLKIKGNRSDGDRDHGLLLN YANN SRITGNIVVGR LQPADRWL KARSSGHGV PKTDEEN
PS LEGNVIEDVRYGVHYMFANENSLIDNVTRTRTGYALMQSRKLTVTGNRSEQDQNYGILMNYITYSTITGNFVSD.....V
C ---N---R---H-M-----N-----G-A-M-S-----N-S--D---G---NY-----N-V-----
301                               400
AC .....RGGTKKCLFIYNAHKNLVWGNRFESC GIGIHFTAGSEKNVLTGNAFIANREQVKYVGTRNMEWSHEGRGNFWS DHPAFDLNGDGVADSF
RS .....RGGTGKCTFIYNAHRNVIHDNRFEGCGTGIHFTAGSERNMLTGNAFVGNRTQVKHVGTKLV EWSHEGRGNFWS DHPAYDL DGDGIADAT
BJ MIPDTKREQSLRSGPEKCVFIYNTNHNRF RNNWFERCAIGVHFTAGSEGNEITGNAFVGNRNQVKYVGTRNLDWSKGGRGNYWSDNPAFDLDGDGIADTA
SM QTAGADR....RLGPEKCVFIYNANKNRFRD NVFEGCAIGIHFTAGSEGNLISSNSFNNRNQVKYVGTRHLDWSSEGQGN YWSDNPAFDLDGDGIGDNP
PS QRGDTGGDSMISGGEGKALFIYNSLFNTIENNHFEKSSLG IHLTAGSEDNRISGNAFVGNQQQVKYVASRTQEWSVDGRGN YWSDYLGWDRNNDGLGDIA
C -----G--K--FIYN---N---N-FE---G-H-TAGSE-N---N-F--N--QVK-V-----WS--G-GN-WSD---D---DG--D---
401                               490
AC YRPNDLMDQILWSQPAASLLTGSPAVQIVRWSQRDFPATLPGGVRDSAPLMRPLTIPVPLEILAYEAEAAGRWTEGNYDDTDADNLQAH*
RS FRPNDLMDHILWSQPAAALLVGSPAVQLIRWSQASFPATLPGGVRDSHPLMAPIDIPVDDVARLEA.AVPPWPQG..ETSDADTLAGH*
BJ YRPNDLVDRVLWTAPSAKVLINSPAVQVLRWAQAQFPALYPGGVIDSHPLIAPPPRPSGKEGGR*.....
SM YRPNDLIDKVLWTSPOAKLLTTS PAVQVIRWAQAQFPAILPGGVVDSRPLMPVAGRVA VQ*.....
PS YEPNDNVDRLLWLYPQVRLLMNSPSIEVLRWVQRAFPVIKSPGVQDSHPLMKLPTEKLLTEKQEPTS*.....
C --PND--D--LW--P---L--SP-----RW-Q--FP-----GV-DS-PL-----

```

33

Figure 11. NosD sequence alignment from *A. cycloclastes* (AC), *R. sphaeroides* (RS), *B. japonicum* (BJ), *S. meliloti* (SM), and *P. stutzeri* (PS). C, conserved amino acids; bold, potential copper ligands; ↓, the predicted start of mature *A. cycloclastes* NosD.

```

1
BJ MSATVSVTG VVKNYRTVRALRDVSFDLAPGRLSALVGHNGAGKTTLIKMLGLIRPDQGSIRVLDEDPAGEFSARRQLGYLPENVAFNAAALTGRETLTF
SM MSGTVEIAGVSKCYGDSTVVRDISFGLGAAETVALVGHNGAGKTTLIKMLGLIRPTKGLVRLVGENPATGDFAVRQRLGYLPESVSFNMALTGRETLRF
AC MPTTLTISR LTKRFQSVEALKSVSLALEPGRRALLGHNGAGKSTMMKIVLGLIPFDSGEVSV..CGSAPGSPAARTQVAYLPENVAFHFPALTGEEQLRH
RS ...MLTLSQVTRTFGGLTALDAVSFAVAPGERVALLGHNGAGKSTLMKLVGLIAADGGRI SV..GGHAPGSHAARALTAYLPENAAFHFPALTGAETLRY
PS .MNAVEIQGV SQRYGSMTVLHDLN LNLGEGEV LGLFGHNGAGKTTSMKLILGLLSPSEGQVKV L..GRAPNDPQVRRQLGYLPENVTFYPQLSGRETLRH
C -----L-GHNGAGK-T--K--LGL-----G--V-----A-----R-----YLPE---F---L-G-E-L--
                                     ATP/GTP binding site
101
BJ YARLKHVKPSKAWAFLDRVGLT. DAADRRVGTYSKGMQRQLGLAQALLGRPRVLLLDEPTTGLDPALROTIFYEILNELRDDGATVLISSHALNELEDRAE
SM YARLKQVDGAATGDLFERVGLAQEAVDRPVRTYSKGMQRQLGLAQALLGMPRIILLDEPTSGLDPALRRNFYELITELRAKGTTVLLSSHALTELEGRAD
AC YLALRGENPKQATELLARVGLG. HAARRRIGTYSKGMQRQVGLAQTLIGRPRLVLDEPTSGLDPVSRDFYDLLDGLAAEGTAILLSSHVLTEVEARTD
RS YLRLRGEDPAKAGALLERVGLG. AAAQRRIGTYSKGMQRQVGLAQALIGHPRLLVLDEPTSGLDPVSRDFYALLAELAADGAAILLSSHALTEVEARTD
PS FARLKGAALTQVDELLEQVGLA. HAADRRVKTYSKGMQRQLGLAQALLGEPRLLLLDEPTVGLDPIATQDLYLLIDRLRQRGTSIILCSHVLPGVEAHIN
C ---L-----VGL---A--R---TYSKGMQRQ-GLAQ-L-G-PR-L-LDEPT-GLDP-----Y-----L---G-----SH-L---E---
                                     ABC motif
201
BJ HVLIMNRGLLAAQGTLSLSELRISOLPIRVSV D.....VDAEVSSDRKIALLLHDITSDAALLSGLTITEPTLDDLYAH
SM RVIIIVNKGVKIADGTLEQLRRIARLPTRISLKL SQAGATPAWLNGGMKWCRGPDGA...LRLRGSQSEKLPLLARIAALGSAIADLDVIPPSLEDIYSH
AC SILILSGGELVAEGTLADLRTRAAPVAFSVR. PAPGHAPALAATFPEGVEGEDGL...LRLRGSQSEKLPLLARIAALGSAIADLDVIPPSLEDIYSH
RS RIAILSGGRLMANDTLANLRQAALPIRLQVS. ARPGAADVAHRL. OGGRNNGVL...VDLSCRPEKLDRLGAILTALGPMVEDVEILPPSLEDLYTH
PS RAAILAKGCLQAVGSLSQLRAEAGLPVR..IRASGISERDSWLQRWTDAGHSARGLSSESSIEVAVNGHKLVLRLQLLGESEP. EDIEIHQPSLEDLYRY
C ---I---G---A---L---LR---LP-----
301
BJ FLNGGVTK*.....
SM FSRDGO*.....
AC FSRDGO*.....
RS YSRPS*.....
PS YMERAGDVRAQEGRL*
C -----

```

Figure 12. Sequence alignment of NosF from *A. cycloclastes* (AC), *R. sphaeroides* (RS), *B. japonicum* (BJ), *S. meliloti* (SM), and *P. stutzeri* (PS). C, conserved amino acids. The conserved ATP/GTP binding site and ABC transporter motif are marked in bold.

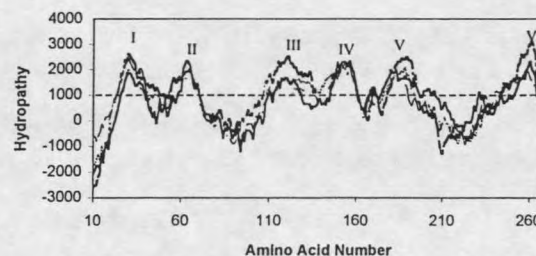


Figure 13. Hydropathy plots of NosY from *A. cycloclastes*, *R. sphaeroides*, *B. japonicum*, *S. meliloti*, and *P. stutzeri*. The dashed line depicts the cutoff score for a transmembrane helix.

presumed to bind lipid, which anchors NosL into the outer membrane. A putative disulfide isomerase motif (CysXaaXaaCys) found in *P. stutzeri* is not conserved in any of the other sequences, nor is a reported His-rich region in the C-terminus. There are 3 conserved Met, 2 Cys, 1 Tyr, and 1 Trp, but no His, among NosLs from these organisms.

NosX. Immediately downstream of *nosL* in *A. cycloclastes* is an open reading frame (ORF) that encodes a protein of 303 amino acids. As seen in Figure 15, this ORF displays 60% similarity and 41% identity to NosX of *S. meliloti*, and contains a possible transmembrane helix at positions 278 - 301, comparable to that found in other NosX sequences. Except for a potential signal sequence of 37 amino acids, no other structural motifs are evident in the *A. cycloclastes* NosX sequence.

```

      1                               ↓                               50
BJ MAPANLTRRRMITIAAATAGSAFLSGGGHA.SSTGAVRWGRSALGAQVSI
SM .....MAAAAGLPLLDLRGRAEGAVAAVTWRGRALGAPATL
AC ...MLLTRRLIAISAVLATAPAM....AHANAATRLWTGQALGARASI
C -----A--A-----A-----W-G-ALGA-----
      51                               100
BJ DIFHPDREEAERIIQDCLTEVRRLEGQFSLYRADSAICALNRSGILVAPE
SM IILNLESQADAAGLVDRVVAEVARLERVFSLYQRDSALAE LNRTGAIAAPP
AC RIAHP...EAGTIVSRVMAEIDRLNLSLYRANSALCRLNRDGHLDQDPP
C -----A-----E--RLE---SLY---SA---LNR-G---P-
      101                               150
BJ PDMVALLKASLQFSDLTGGTFDPTVQPLWQLYAAHFSSNPDPGGPPAGK
SM PDLVNILLEASRDFWETTGAFDPSVQPLWTLYAEHFAAGDADPAGPPEAA
AC FELLECLSLADAVHLASGGRFDPTVQPLWTLWAQ..SAANGTRPGPADVD
C -----L-----GG-FDP-VQPLW-L-A-----GP-----
      151                               200
BJ LAEALAKVGSSGLRVSEKLVALLRHGAITLNGIAQGYATDRVVDVLRAR
SM KRRALSRVGFDFKLEFNDRVVFARPGMALTNGIAQGYITDRIVGLLKDA
AC RVRG..NVGWGKVKLGAAAITL.HPGMALTNGIGQGYVADRVAAMLEVE
C -----VG-----G-A-TLNGI-QGY--DR----L-----

```

Continued on the next page

```

      201                                     250
BJ GLSTTLVNMGEIRALGARADGTPWRVGLSDPDRAGALTETID3LIDRAVAT
SM GIANSLVSMGETRAIGSQHDGRPWRVGLATREDASTPDSVLNLVNRAVAT
AC GLTDILIDTGEFRALGGRPGGGGWVRL3ETGER.....LPLRQRAVAT
C  G----L---GE-RA-G----G--W-V-L-----L--RAVAT
      251                                     300
BJ TAGAGFRFDPAGRFTHLFDPTG.RSPALYRTVSVIAPTATEADALSTAF
SM SSPDGFRFD3DSGRFGHILDPLSG.RAPRLRRRVSVVAPTATAADAFSTAF
AC SAPLGTTFDQAARDGHILDPLTGMPAPANWRSVSVSASSAAIADALATAA
C  ----G--FD---R--H--DP--G---P---R-VSV-A--A--ADA--TA-
      301                                     350
BJ SLMPASQIGDIAATRPDVASSGLINQWMEICAFNGRLRGGVAIRSGGESG
SM SLMGSSAV.RIACEHH....SELTVDMISTSGAHERFGRAA*.....
AC CLLPDAG.....SIVAMNARFDGVRLEAASV*..
C  -L-----R-----
      351                                     369
BJ IRNPRYGFP3AHTLSKRAP*
SM .....
AC .....
C  -----

```

Figure 15. Sequence alignment of NosX from *A. cycloclastes* (AC), *B. japonicum* (BJ), and *S. meliloti* (SM). A potential transmembrane helix is boxed, a possible signal sequence cleavage site is indicated by an arrow.

Discussion

While the *nos* genes in *S. meliloti* are localized on a megaplasmid (120), their location in *A. cycloclastes* has not been determined. *A. cycloclastes* is closely related to *S. meliloti*, as evidenced by comparison of its 16S rRNA sequence with other gram-negative bacteria (59) and suggested by J.M. Tiedje (121). Nevertheless, the two organisms differ in that *A. cycloclastes* does not nodulate alfalfa (122) and is not known to fix N₂. Thus, despite the close evolutionary relatedness of *A. cycloclastes* to *S. meliloti*, the two should remain in separate taxa.

The intergenic spacing within the *nos* clusters examined is variable among organisms. In *P. stutzeri*, which has six unique *nosZ* transcriptional start sites, the 74-nucleotides between *nosR* and *nosZ* contain the two promoters that are most highly expressed under anaerobic and aerobic conditions (77). The decrease in the size of this and other intergenic regions in other organisms may mean that transcription is less complex in these organisms. For example, in *S. meliloti* and *B. japonicum* the initiation codon of the *nosDFYL* transcriptional unit overlaps the termination codon of the preceding *nosZ* unit, and in *A. cycloclastes* only three nucleotides separate *nosZ* from *nosDFYL*. This is in stark contrast to the 122 nucleotide *nosZ-nosD* intergenic region found in *P. stutzeri*. Variability among organisms within *nosDFYL* is also evident. For instance, in *A. cycloclastes* *nosD* and *nosF* overlap by 35 nucleotides, while in the other organisms these genes share only initiation and termination codons. The large overlap in *A. cycloclastes* is unusual, but not unprecedented. For example, *trpE* and *trpG* in *Methanobacterium thermoautotrophicum* overlap by 26 nucleotides, although this arrangement is not observed in the *trp* operons of other organisms (123). If the *nos* gene group represents a type of multi-component ABC transporter system, then the compression of these genes in *A. cycloclastes* might represent an evolutionary step toward the fusion of the components, making the system more like the multifunctional, single-gene ABC transporters found in eukaryotes and some prokaryotes (99).

It is well established that Nos proteins are expressed under conditions of low oxygen and high N-oxides (4). In *E. coli*, the protein FNR regulates the expression of many

anaerobic proteins (78), and several FNR-analogs have been found that regulate the denitrification pathway (2,65,67,70-77). While FNR and degenerate σ^{54} (RpoN, nitrogen-responsive sigma factor) binding sites have been detected within the promoter regions of the genes sequenced in this investigation, their significance remains unclear because the location of these sites relative to the start of transcription is unknown.

The high similarity found in the C-termini of NosR from three sources suggests that this region is of functional importance. In fact, a search of gene databases indicates that the overall C-terminal structure of "helix-CX₃CP-helix-helix-CX₃CP-FeS-FeS" (Figure 6, inset) is found in a variety of membrane proteins. At least two of these, NapH and MauN, are thought to have electron-transfer function and operate in conjunction with a periplasmic component, NapG or MauM (5). Hence, this structure may represent a universal redox-sensing domain that produces a transmembrane response to changes in the cellular redox potential. The remainder of the NosR protein is characterized by a large periplasmic loop that shows only limited homology among the organisms examined. TFASTA searches of the three conserved regions of this loop produced no significant matches that might suggest functionality.

The unusually long leader sequence predicted for N₂ORs (NosZ) has been confirmed for *P. stutzeri* (118) and now *A. cycloclastes*. This presequence contains the double Arg motif that is found in all N₂ORs as well as a number of proteins that require cofactor insertion or processing (118,124). Proteins with these signal sequences are transported by the TAT (twin

arginine transport) system in lieu of the Sec-dependent system (124,125). This finding is significant when considering the homology of the *nosDFYL* operon with ABC transporters. The TAT leader sequence in N₂OR makes it highly unlikely that NosD, F, Y, and L function to transport N₂OR into the periplasm. Since copper incorporation into N₂OR is believed to occur in the periplasm (2), it is also unlikely that the *nosDFYL* code for proteins involved in copper transport.

The folding of the Cu_A domain in cytochrome oxidase is similar to that of other electron-transfer copper proteins, such as plastocyanin, despite the low homology suggested by primary sequence comparisons (52,112). Additionally, resonance Raman spectroscopy has shown that the Cu_A site is highly conserved between N₂OR and cytochrome oxidase (126). Hence, a model of the N₂OR C-terminal Cu_A domain was constructed, based on the crystallographically-determined Cu_A site of bovine heart cytochrome oxidase (51). The model suggests that of the two potential Cu_{Cat} ligands located in the Cu_A domain, only H661 is available for metal binding; the other, M671, is buried within the protein. If H661 is a Cu_{Cat} ligand, then the organization of the copper sites in N₂OR would be similar to those found in the multi-copper enzymes nitrite reductase and ascorbate oxidase (112). In the latter proteins, a Cys-His motif provides the link between the electron-transfer copper that is coordinated to the Cys (corresponding to C660 in N₂OR), and the catalytic copper, which is ligated by the His residue (His661 in N₂OR). In this context, it is noteworthy that superimposing the N₂OR model on the entire cytochrome oxidase structure implies that H661 would be located at the

interface of the Cu_A domain with another subunit. Remarkably, the catalytic coppers in nitrite reductase and ascorbate oxidase (Type II and the trinuclear cluster, respectively) are located at the interface of two domains in a structurally homologous fashion (112). Sequence comparison of the trinuclear copper-binding region in ascorbate oxidase with the Type II-binding region in nitrite reductase reveals that the two share 50% similarity and 25% identity. Since the presumed Cu_{Cat} domain of N₂OR, residues 150 – 212 in Figure 7, also displays 50% similarity and 25% identity to the ascorbate oxidase trinuclear domain, it is plausible that the Cu_{Cat} site is also found at a domain interface.

Optical and EPR spectroscopies suggest that the Cu_{Cat} sites in N₂ORs from *A. cycloclastes* and *P. stutzeri* are remarkably similar (127). Moreover, examination of the enzyme from *P. stutzeri* by magnetic circular dichroism, infrared, and resonance Raman (53,54,88,89) has provided evidence for a Cu_{Cat}-thiolate bond originating from a Cys residue. However, the only conserved Cys (outside of the presumed Cu_A ligands) noted in previous comparisons of N₂ORs is T190 in *A. cycloclastes*. Accordingly, if Cys is a Cu_{Cat} ligand, then the residue must occupy a variable position among N₂ORs. Positional variation of Cys metal-ligands is observed among ferridoxins (128), in which one Cys ligand to the iron-sulfur cluster may be separated from other ligands by 2 to 8 amino acids. Interestingly, N₂OR from *A. cycloclastes* does contain a Cys at position 185, only 5 residues away from C190, and both are followed by aspartate in their respective sequences. The function of C190 as a Cu_{Cat} ligand was tested by mutation (C190G in Figure 7, corresponding to C165 in the *P. stutzeri*

sequence) of this residue in N₂OR from *P. stutzeri* (98). The copper content and enzyme activity were unaffected by the mutation (although protein stability was changed), leading the authors to conclude that C190 is not a Cu_{Cat} ligand.

An alternative explanation is possible, however, and is preceded by mutational studies on the Cys ligands of a Fe₂-S₂ ferredoxin (129). In these experiments, the site-directed mutagenesis (Cys→Ser) of a Cys ligand in *Azotobacter vinelandii* ferredoxin resulted in another Cys residue of the same flexible, hydrophilic loop becoming coordinated to the metal, in preference to serine coordination. The spectroscopic characteristics of the metal site in the mutant protein were barely distinguishable from the wild type. Invoking a similar type of "ligand swapping" argument for the C190G mutant N₂OR may explain the observation that mutagenesis did not appear to significantly perturb the Cu_{Cat} site. If, in the absence of C190, another Cys became coordinated to Cu_{Cat}, then only minor changes in copper content and catalytic activity are feasible. One possible source of a replacement Cys could be the CHHPH sequence found at position 153 (Figure 7) in *P. stutzeri* N₂OR. If the His residues of this sequence are Cu_{Cat} ligands (HXH is a common copper binding motif (112)), then C153 might be sufficiently close to coordinate to copper in the absence of the normal ligand.

Characterization of the *A. cycloclastes* mutant protein (C185G) might provide a good test of this hypothesis, since the enzyme from this organism does not have a Cys residue near the conserved HHXH motif, and the closest Cys residue in the primary sequence is 40 residues away at position 225. Hence, ligand swapping might be less probable in *A. cycloclastes*.

In 1998 Farrar *et al* presented a radical new hypothesis about the nature of the dinuclear copper centers in N_2OR (130). The model is based on EPR and MCD quantitation of the N_2OR copper centers in several oxidation states. The authors propose that the spectroscopic evidence for a Cys-ligated catalytic copper site should be attributed to a variant form of the Cu_A center (called Cu_Z). Furthermore, the catalytic copper center is suggested to be a His-ligated, $Cu(II)$ - $Cu(II)$ site that has yet to be observed spectroscopically. This interpretation is satisfactory in that the semi-conservation of a Cys ligand and/or ligand swapping is no longer required to explain previously published results (*vide infra*). However, the model assumes that no interconversion between sites occurs during the various stages of reduction, which may not be valid for a system designed to transfer electrons from one site to another. Elucidating the nature of Cu_{cat} will clearly require further research.

It has been proposed that NosD acts as a copper insertase or chaperone in an energy-requiring reaction with N_2OR . The transmembrane protein NosY is the energy transducer that couples ATP/GTP hydrolysis by the cytoplasmic NosF with this reaction (59,96). The conservation of these structural motifs among NosD, F, and Y from three organisms supports this hypothesis, which is largely based on the conserved ABC transporter motif and NTP binding site in NosF. In NosD there are several conserved His, Met, and Tyr residues that might be copper ligands, but other than a HXM sequence, no typical copper-binding motifs are apparent. The lack of readily discernible metal-binding sites in NosD is reminiscent of the nickel insertase component (UreE) of urease bioassembly. In UreE the C-terminus was

truncated to eliminate a His-rich region, and yet remained competent at binding nickel and delivering the metal to apo-urease (109). Hence, the characterization of NosD will be critical to understanding the structure and function of this protein in Cu_{Cat} biogenesis.

The role of NosL in formation of the Cu_{Cat} site remains unclear. Although transposon interruption of *nosL* in *P. stutzeri* did not produce a Nos minus phenotype (98), the Tn5 insertion site was near the C-terminus, leaving open the possibility that a sufficiently large fraction of NosL was transcribed to allow the truncated protein to function. NosL might be an optimizer of Cu_{Cat} assembly, perhaps acting as a membrane-anchored scaffold upon which NosD and N₂OR are gathered for reaction, and thus, not be strictly required for metal incorporation. Alternatively, NosL may be a metallochaperone that is required only when the environmental concentration of copper is low. A recent study of the domains of CCS, the copper chaperone to superoxide dismutase, suggests that its Atx1-like, copper transport domain is not required for metallochaperone activity except under copper-limited growth conditions (131). The characterization of recombinant NosL from *A. cycloclastes* is described in the next chapter.

Another gene, *nosX*, immediately follows *nosL* and has been identified in *S. meliloti* as being essential for N₂OR activity (91). Comparison of the NosX sequences from *A. cycloclastes*, *B. japonicum*, and *S. meliloti* shows the conservation of potential signal cleavage sites and hydrophobicity/hydrophilicity patterns. Surprisingly, the sequence downstream of *nosL* in *P. stutzeri* shows no homology to *nosX*. However, it is possible that

for the latter organism, *nosX* resides at another location. Examples of this are found for several denitrification genes. For instance, in the denitrifier *R. eutropha*, *nosZ* is located on a megaplasmid but is not flanked by either *nosR* or *nosD* (79). Moreover, the nitrite reductase (*nir*) genes of this denitrifier are evidently partitioned between megaplasmid and chromosomal DNA (132).

Conclusions

The comparison of *nos* sequences from different organisms has confirmed several initial observations made for *P. stutzeri*, and invalidated other hypotheses based on the limited sequence data previously available. Based on sequence analysis, it appears likely that NosR and NosY are integral membrane proteins. NosR is predicted to be a metalloprotein that contains two Fe_4S_4 clusters and possibly another metal center. The homology of NosDFYL to ABC transporters is also supported by the addition of four sequences to the database. Unfortunately, the meaning this similarity in terms of copper-site bioassembly is still unclear. NosL is likely to be anchored into the outer membrane by a covalently attached lipid, but does not function as a disulfide isomerase. The discovery of NosX in organisms other than *S. meliloti* and the recognition of *nosX* cognates in other denitrification gene clusters suggests that this protein merits further study.

The construction of a model for the Cu_A domain of N₂OR provides insight into the relationship of protein structure and function. For example, the conservation of an acidic surface patch in both N₂OR and COX suggests that the former enzyme interacts with its electron transfer partner by electrostatic recognition, as does COX. Interestingly, the model was also useful in providing clues to the possible structure of Cu_{Cat}. Of two possible Cu_{Cat} ligands located in the Cu_A domain, only H661 is positioned in a manner amenable to metal-binding. When the structures of other multi-copper proteins are considered, it seems likely that His661 is a Cu_{Cat} ligand. The effects of mutating this residue should be informative about the structure of the catalytic center.

While sequence analysis is a useful tool, the isolation and characterization of Nos proteins is essential to elucidating their structure and function. Moreover, understanding the mechanism of copper site assembly in N₂OR could be informative for other metalloproteins, given the emergence of metallochaperone chemistry as a general phenomenon. Since copper incorporation into N₂OR occurs in the periplasm, the periplasmic proteins NosD and NosL should be considered prime targets for further study.

CHARACTERIZATION OF NOSL, A NOVEL CU(I) PROTEIN

Introduction

Elucidation of the mechanism of copper incorporation in N₂OR is a substantial, but not intractable, undertaking. The analysis of deduced amino acid sequences has provided some information about possible structural and functional features of the Nos proteins. However, the isolation and characterization of these Nos proteins is essential to understanding their specific roles in copper site bioassembly. The recombinant expression of NosD and NosL will be presented, and the characterization of recombinant NosL (rNosL) will be described.

Recombinant protein expression has become a major tool of modern biochemical research. Many choices of vectors exist for expression in *E. coli*. A perusal of the pET manual (Novagen, Inc.) shows the wide variety of options for expression that are commercially available. Expression vectors typically supply an inducible promoter, ribosome binding site, multiple cloning site, transcriptional terminator, and selection marker. Additionally, many vectors include periplasmic leader presequences and/or fusion (protein or peptide) tags. The fusion tags may serve two purposes: (1) the rapid detection of the recombinant protein using antibodies to the tag; and (2) its purification via interaction of the tag with an affinity adsorbent. These factors have made fusion tags a highly desirable characteristic in expression vectors. Furthermore, the discovery of broad host range plasmids has made expression possible in organisms other than *E. coli* (133).

Recombinant proteins are sometimes produced in only in small quantities or in inactive forms. In some instances, the addition of even a small fusion tag has been known to cause large changes in the catalytic activity of a recombinant protein (134). Occasionally, inactive proteins may be converted to a useful state by post-purification treatments. Considering these potential problems, it is important to verify, when possible, that the structure and function of the recombinant protein is similar to the native protein. In the case of the Nos proteins, only N₂OR has a functional assay with which to determine if a recombinant protein has retained its activity. No method for comparing recombinant vs. native protein exists for the other Nos proteins, as none have been previously isolated or characterized. Additionally, the auxiliary Nos proteins are expressed only in small quantities in their native hosts, effectively eliminating the possibility of purifying the proteins from wild type hosts. Therefore, despite the inherent uncertainties of recombinant expression, the potential advantages of producing copious amounts of pure protein invite its use for this research project.

The deduced amino acid sequences of NosD and NosL are available for five organisms (59,81,83,84,91,96,98). Both proteins are predicted to be soluble proteins of the periplasm. Based on this prediction, the *E. coli* expression vector pET20b+ (Novagen) was chosen as a vector for the recombinant expression of NosD and NosL. This vector contains the periplasmic leader sequence of the *E. coli* PelB protein just upstream of its multiple cloning site (Figure 16). The localization of rNosD and rNosL to the periplasm would allow osmotic shock to be used in the purification of the proteins, eliminating the need for a fusion tag.

The periplasmic leader of NosD is predicted to belong to the same class of signal sequences as PelB, being cleaved by Signal Peptidase I. In contrast, NosL displays a

presequence that is thought to be cleaved by Signal Peptidase II (SpaseII). This enzyme is also known as Prolipoprotein Signal Peptidase because it only cleaves certain presequences to which

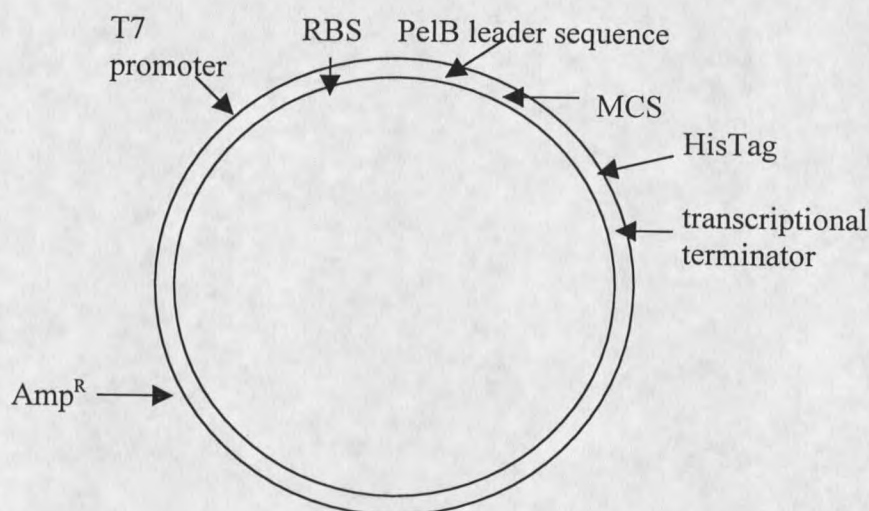


Figure 16. The pET20b+ expression vector.

a lipid moiety has been covalently attached (110,111). SPaseII recognizes a highly specific signal sequence consisting of at least one Lys or Arg residue in the first seven positions, followed by a stretch of uncharged and hydrophobic residues, and ending in a sequence called the lipobox. The lipobox consensus sequence, Figure 17, was derived from examination of over 130 bacterial lipoproteins sequences (111) and exhaustive mutational studies of the signal sequence in murein lipoprotein (135). Lipid modification occurs at the +1 Cys. The identities of the +2 to +5 residues are thought to determine the localization of the lipoprotein to either the inner or outer membrane.

Position	-5	-4	-3	-2	-1	+1
Amino	L, I, V, A, S,	L, A, V, T, I,	L, I, V, F,	A, S, I, T, V	G, A, S	C
Acid	G, M, F, T, N	M, F, G, S, H	A, S, T	G, C, Q, M		

Figure 17. The lipobox consensus sequence of SPaseII.

SpaseII activity has been documented for many species of bacteria (including several denitrifiers) and appears to be a general prokaryotic pathway (111). The lipoproteins associated with SpaseII perform many different functions, including penicillinase, lysis proteins, electron transfer proteins (cytochromes and blue copper proteins), and secretion factors. The nature of the fatty acid moieties may be determined by chemical (amino acid sequencing, fatty acid analysis) and biophysical (mass spectrometry) methods. An assay for SPaseII activity has been developed (110). Additionally, the inhibitor globomycin has been used to isolate prolipoproteins and elucidate the mechanism of post-translational modification (136).

Lipid modification is a multi-step process that has been best characterized for murein lipoprotein (111). Initially, glyceryl transferase creates a thioether link between a glycerol molecule and the +1 Cys residue. O-acyl transferases then acylate the C₁ and C₂ positions of the glycerol moiety. The resulting diglyceride prolipoprotein serves as the substrate for SPaseII. The peptidase cleaves the leader sequence between the -1 residue and the invariant +1 cysteine. Additional acylation of the newly created N-terminal Cys occurs in murein lipoprotein but is absent in other proteins (111,137). Literature indicates that both the lipobox sequence and lipid modification are required for SPaseII activity. Moreover, the selectivity of the lipobox sequence is demonstrated by the ability of *E. coli* cells to modify the lipobox even when it is artificially inserted into an internal position in a protein sequence (138).

Of the proteins encoded by the *nosDFYL* operon, one or both of the periplasmic proteins would be expected to function as a metallochaperone. It will be shown that NosL binds Cu(I) stoichiometrically, and that Cys, Met, and His are the probable ligands to the copper ion.

Experimental Methods

Design of Recombinant Constructs. Recombinant NosD (rNosD) and recombinant NosL (rNosL) were expressed in *E. coli* strains DE3(BL21)pLysS and DE3(BL21) (Stratagene), using the pET20b+ vector (Novagen). This vector carries the coding sequence for the PelB periplasmic leader under the control of the T7 promoter. Inserts containing the predicted mature coding sequences of each gene were produced by PCR. Primers were designed to introduce a 3' *XhoI* site (*nosL*) or *BamHI* site (*nosD*) into the PCR products. Specifically, nucleotides 4465-5739 (see Appendix B) were amplified for *nosD*, and 7479-8000 for *nosL*. After digestion with either *XhoI* or *BamHI*, each insert was then ligated into pET20b+ that had been prepared by digestion with *EcoRV* and either *XhoI* or *BamHI*. The integrity of each construct (pD21, *nosD*; pL46, *nosL*) was verified by DNA cycle sequencing as described in the previous chapter. The pL46 construct produces a periplasmic protein in which the N-terminal Cys of mature NosL has been replaced with Met-Asp. Control strains used were DE3(BL21)pET20b+ and DE3(BL21)pLysS/pET20b+.

Expression of rNosD. Strain DE3(BL21)/pLysS/pD21 was grown to an $OD_{600} = 0.4$ at 37°C in Luria-Bertani (LB) media containing 100 µg/mL carbenicillin, followed by induction with 1 mM isopropyl-β-D-thiogalactoside (IPTG) at 22°C. Protein content was monitored over time via SDS/PAGE (8-25% gels) on a Pharmacia PhastSystem.

Expression and Purification of rNosL. Recombinant NosL was expressed using strain DE3(BL21)/pL46. Initially, rNosL was expressed in LB media containing 100 µg/mL carbenicillin. However, when preliminary metal-binding studies on purified rNosL indicated the uptake of stoichiometric copper, the growth media was supplemented with $Cu(NO_3)_2$ to allow for the direct isolation of metallated protein. Four 1-L cultures of DE3(BL21)/pL46 were grown to $OD_{600} = 0.8$ at 37°C and then transferred to 20°C prior to induction. (For copper-supplemented media, 50 µM $Cu(NO_3)_2$ was added at the time of the last inoculation.) The cultures were induced by the addition of 0.5 mM IPTG.

During the purification of copper-containing rNosL, all solutions were made anaerobic by Argon purge. Cells were harvested 10 hours after induction. Whereas longer induction periods resulted in cell lysis, shorter periods yielded significant amounts of unprocessed, cytoplasmic pre-protein, as judged by SDS/PAGE. Thus, the 10-hour induction period represented an attempt to balance cell lysis with periplasmic processing. The cells were collected by centrifugation and washed with 10 mM MES, pH 6.5 (Buffer A). Next, the cells were subjected to osmotic shock according to the method of Ausubel *et al* (139), except that EDTA was omitted from the hypertonic solution. This step separated any remaining

unprocessed rNosL from the mature protein. The periplasmic fraction was mixed for 30 minutes with 30 mL DEAE-FastFlow (Pharmacia) equilibrated in Buffer A. The slurry was packed into a 5x15 cm column, washed with 300 mL Buffer A, and eluted with Buffer A containing 1.0 M NaCl. The eluent was concentrated to 10 mL in an Amicon filtration apparatus using a YM10 membrane and equilibrated into Buffer A either by anaerobic dialysis or successive concentration/dilution steps. rNosL was purified to homogeneity (Figure 18) by gradient elution on a MonoQ HR10/10 column (Pharmacia) at a yield of 4 mg/L culture. Occasionally a gel filtration (S-200 HR, Pharmacia) column was employed as a final purification step.

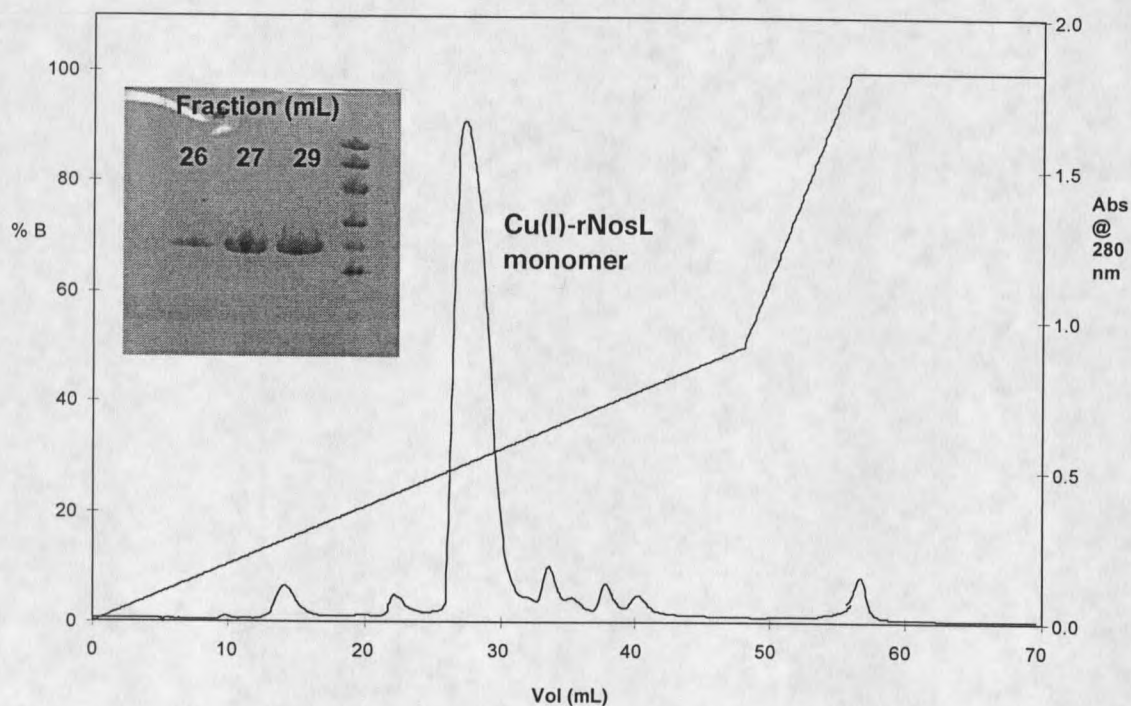


Figure 18. Chromatogram of the purification of rNosL by ion exchange chromatography (MonoQ column) Inset: SDS/PAGE of the rNosL column fractions.

Analysis of rNosL. A Pharmacia PhastSystem with 8-25% gradient gels was used for SDS and Native PAGE analyses. For isoelectric focusing, the PhastSystem was used with Pharmacia Broad pH (3-9) Range gels. Amino-terminal sequencing was performed by Gerhard Munske, Washington State University, Pullman, WA. LC-ES/MS (liquid chromatography-electrospray mass spectrometry) was performed by Lowell Ericsson, University of Washington, Seattle, WA. The metal content of NosL was measured by flame atomic absorption spectroscopy (AAS) on a Buck Scientific instrument (copper, cobalt), or by inductively coupled plasma-atomic emission spectroscopy (ICP/AES) (manganese) by Little Bear Labs, Inc., Boulder, CO. UV/Vis spectra were recorded on either a Hewlett Packard 8452A instrument or a Varian Cary-219 spectrophotometer interfaced to an OLIS controller. Magnetic circular dichroism (MCD) measurements were performed on a Jasco J-710 instrument equipped with a 1 T magnet. X-ray absorption spectroscopy (XAS) and extended X-ray absorption fine structure (EXAFS) data were collected in collaboration with Dr. Robert Scott and Mr. Nathaniel Cosper, University of Georgia, Athens. The oligomeric states of rNosL were determined by gel filtration on a 1.6x50 cm Sephacryl S200 HR column. 10 mM KPO_4 pH 7.0 + 0.5 M NaCl buffer was used at a flow rate of 0.15 mL/min. Fluorescence spectra were recorded on a Spex fluorimeter interfaced with Grams software (Galactic Industries).

Electron paramagnetic resonance (EPR) measurements were obtained at 77K on a Bruker 220D-SRC Spectrophotometer. For data in Figures 21 and 23, conditions were 9.435 GHz frequency, 4 G modulation, 0.2 s time constant. Power settings were 20 mW (Figure 21, Figure 23,A) and 0.5 mW (Figure 23, B and C).

Results

Expression Trials of rNosD. When expressed in DE3(BL21) cells, rNosD appears to have a toxic effect on cell growth, as judged by the lower rate of exponential growth compared with a control strain of DE3(BL21)/pET20b+. Therefore, the strain DE3(BL21)/pLysS was selected as a host, because pLysS is a specific inhibitor of the T7 RNA polymerase used by the pET system for induction. Strains containing pLysS exhibit significantly less protein expression prior to induction.

As seen in Figure 19, induction of DE3(BL23)/pLysS/pD21 produces a protein of ~ 43 kDa. This is close to the calculated MW of 46 kDa for rNosD. Attempts to increase the expression level of rNosD by changing the temperature or amount of IPTG during induction were unsuccessful. Two other *E. coli* strains, C41(BL21) and C43(BL21) (140) were also used

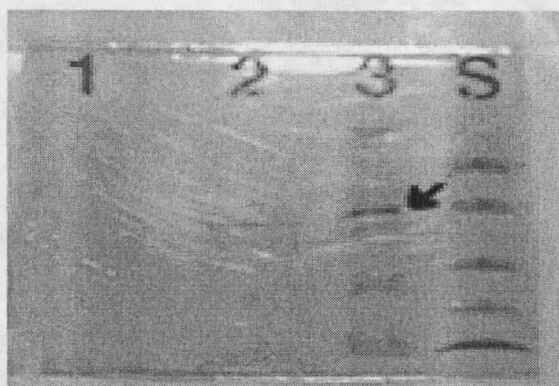


Figure 19. SDS/PAGE showing rNosD expression over time. An arrow marks the position of rNosD. Lane 1, 0 h; lane 2, 8 h; lane 3, soluble cell fraction after 8 h induction; S, MW standards 94, 67, 43, 30, 20.1 and 14.4 kDa.

as hosts for pD21, but these showed no improvement in expression over the original strains. The addition of metals (10 μ M Cu(II), Fe(III), and Mg(II)) to the growth media did not reduce the apparent toxicity of rNosD. No attempts were made to purify the recombinant protein.

The cytoplasmic expression of mature rNosD was also attempted using the pProEX-

HTa vector (Gibco BRL). While high levels of expression were achieved with this construct, NosD from this system formed inclusion bodies that could not be re-solubilized.

Protein Characterization of rNosL. NosL is predicted to be a periplasmic lipoprotein, anchored to the outer membrane via a thioether bond between a lipid and the N-terminal Cys of the mature protein (59,91,96,98). Additionally, NosL is not produced in high yields in wild type *A. cycloclastes*. Therefore, to obtain sufficient quantities of NosL for characterization, an expression system in *E. coli* was used (59). While rNosL is also a periplasmic protein, its N-terminal Cys has been replaced with Met-Asp. This substitution avoids the potential problem of lipoprotein aggregation during purification (111) and allows the use of osmotic shock as a purification procedure.

When grown in media supplemented with $\text{Cu}(\text{NO}_3)_2$ and purified anaerobically, rNosL was found to be almost exclusively monomeric, as judged by native PAGE and gel filtration on a standardized column. However, aerobic purification from metal-depleted media produced a 50/50 mixture of monomeric and dimeric apo-rNosL (Figure 20). The apparent molecular weights of the monomer (26.9 kDa) and dimer (48.3 kDa) calculated from the retention times on the gel filtration column are higher than expected. This suggests that the tertiary structure of rNosL deviates from a spherical form. When either the apo- or Cu(I) monomer forms of rNosL were aerobically stored at 4°C for several days, partial conversion to the dimer form was noted. The monomeric molecular weight of rNosL was estimated to be $18,543 \pm 0.05\%$ by LC-ES/MS, in agreement with the mass calculated from the protein sequence (18,538 Da). An additional peak at 17,570 mass units was also detected and probably represents a proteolytic

fragment in which ten C-terminal amino acid residues have been removed. The identity of rNosL was confirmed by N-terminal sequencing of the purified protein. The pI of rNosL was determined to be ~ 4 .

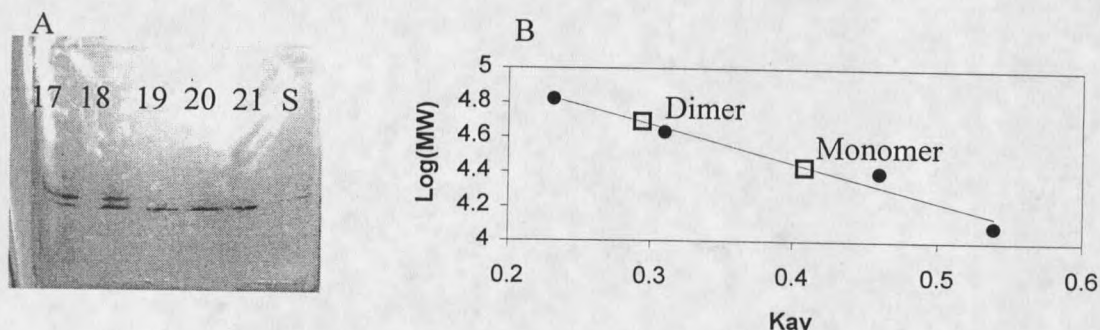


Figure 20. (A) Native PAGE of gel filtration rNosL fractions isolated from metal-depleted media. Each fraction was $> 98\%$ pure by SDS/PAGE and displayed a single band at 18,500 Da. S, SDS/PAGE standards described in Figure 19. (B) Determination of the MWs of dimeric and monomeric rNosL from tubes 17 (dimer) and 20 (monomer) in (A) by gel filtration on a standardized column. Proteins standards were bovine serum albumin, 66kDa; ovalbumin, 43 kDa, trypsinogen, 24 kDa; and cytochrome c, 12.4 kDa.

In order to determine an extinction coefficient @ 278 nm ($\lambda_{\max}$) for rNosL, the molar ellipticity ($\Delta\epsilon_M$) of a protein sample with UV absorbance $\Delta A_{278} = 0.905$ was measured by MCD. Using the published value for Trp of $\Delta\epsilon_M = 777 \text{ deg cm}^2 \text{ decimol}^{-1} \text{ kG}^{-1}$ (141) and the molar ratio of 2 Trp per monomer given by the rNosL amino acid sequence (59), the protein concentration of the sample was determined to be 0.84 mg/mL. Thus, ϵ_{278} was calculated to be 10.8 cm^{-1} for a 1% solution, or $20.0 \text{ mM}^{-1} \text{ cm}^{-1}$.

Metal Content of rNosL. When isolated from metal-depleted cultures, rNosL does not contain copper. However, the uptake of one copper per monomer (as determined by AAS) was detected

when monomeric apo-NosL was dialyzed first vs. 1 mM Cu(I) or Cu(II), then vs. buffer to remove excess copper. No EPR signal was detectable from copper-rNosL samples (< 4% based on integrations using Cu(II)triethanolamine, CuTEA, as a standard), Figure 21. Additionally, no differences in Trp fluorescence were noted between apo- and copper-rNosL, and no features were apparent in the visible absorbance spectra of these samples (Figure 21). Collectively, these data indicate that copper is present in the protein as Cu(I). This was confirmed by reacting copper-rNosL with the Cu(I)-selective reagent 2,2'-biquinoline (BQ), which forms a purple complex with Cu(I) absorbing at 546 nm under acidic conditions (142). Using the published extinction coefficient for Cu(I)-BQ of $6.3 \text{ mM}^{-1} \text{ cm}^{-1}$, the stoichiometry of Cu(I):monomer was confirmed to be 1:1 (Figure 22A). No reaction was observed for apo-rNosL. Similar analyses of rNosL anaerobically isolated from Cu(II)-supplemented media showed that rNosL contains Cu(I) in a 1:1 stoichiometry. However, upon aerobic storage at 4°C for a week, both EPR and BQ-titration studies indicated that 20 - 25% of the copper was oxidized to Cu(II) (Figure 22B). Dialysis vs. EDTA selectively removed the cupric ion from the protein. While Cu(I)-rNosL could not be oxidized by H_2O_2 , ferricyanide and hexachloroiridate successfully oxidized the metal to the cupric state. Rapid concentration/dilution of the oxidized sample resulted in the complete loss of copper from the sample. Furthermore, the EPR spectrum of oxidized rNosL was found to depend upon the buffer used in the experiments. Whereas phosphate buffer gave a broad signal similar to aqueous cupric sulfate, in Good's buffers the EPR signal resembled Type II copper (Figure 23).

Metal substitution of monomeric apo-rNosL, prepared by treatment with N,N-diethyldithiocarbamate (143), was attempted by anaerobically dialyzing the protein overnight

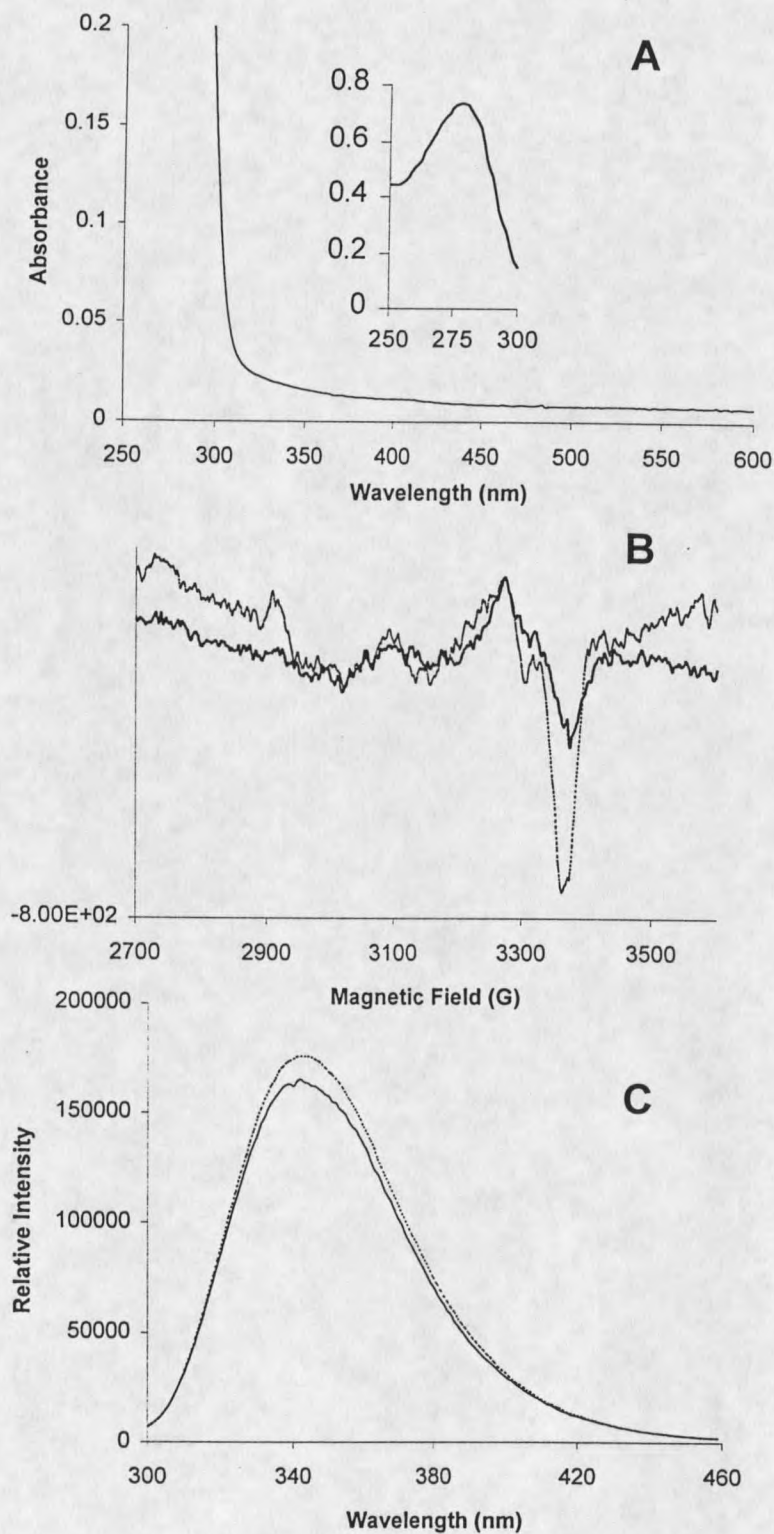


Figure 21. Spectroscopic characterization of rNosL. (A) UV/Vis of 37 μM Cu(I)-rNosL. (B) EPR spectra of 490 μM Cu(I)-rNosL (—) and 50 μM CuTEA (---). (C) Emission spectra (290 nm excitation of tryptophan) of 19 μM apo-rNosL (—) and Cu(I)-rNosL (---).

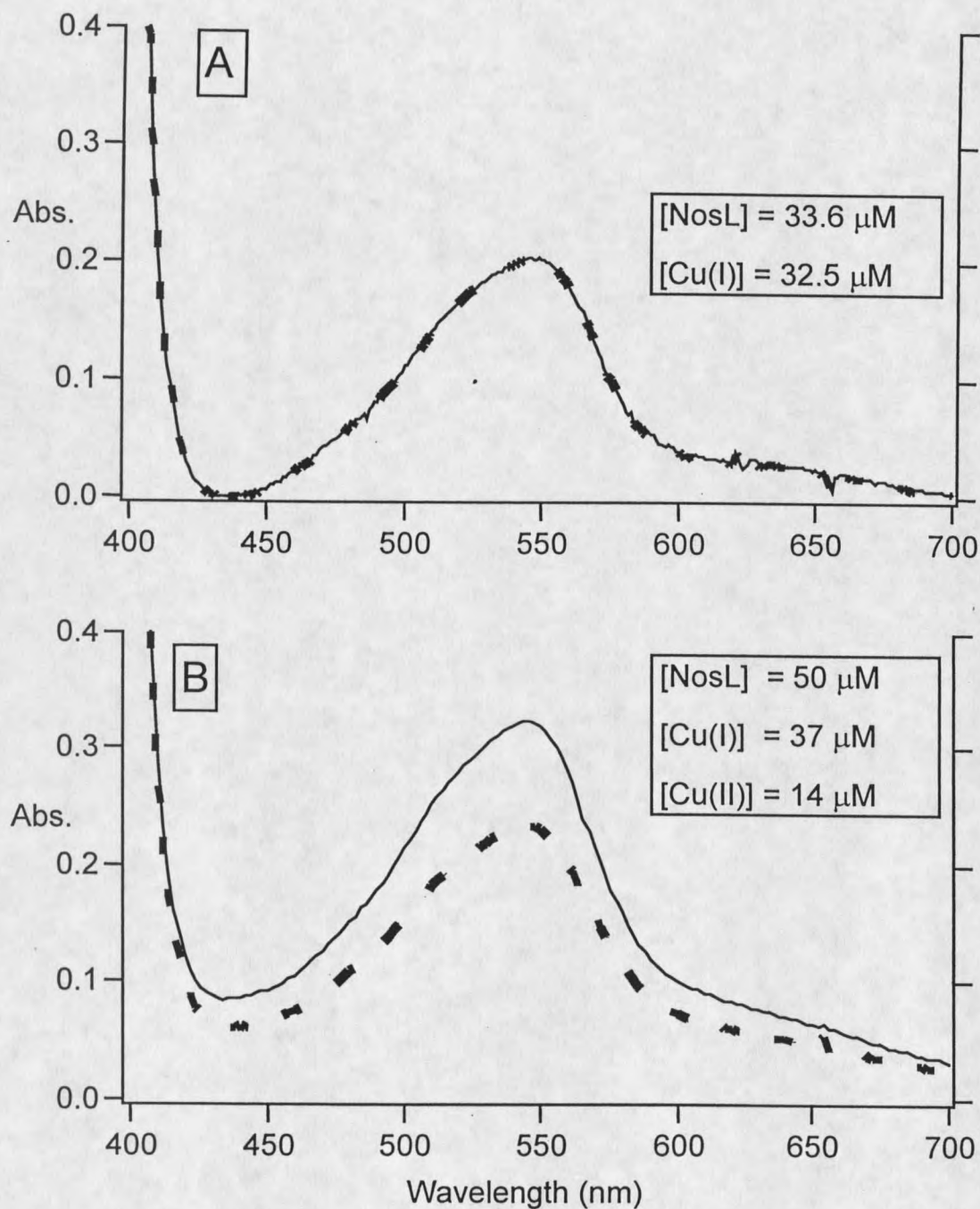


Figure 22. The reaction of Biquinoline with rNosL before (A) and after (B) aerobic storage at 4°C. (---) protein as isolated; (—) after reduction with ascorbate. The amount of Cu(II) in (B) was determined by subtracting the amount of Cu(I) detected in the protein as isolated from the total amount detected after reduction.

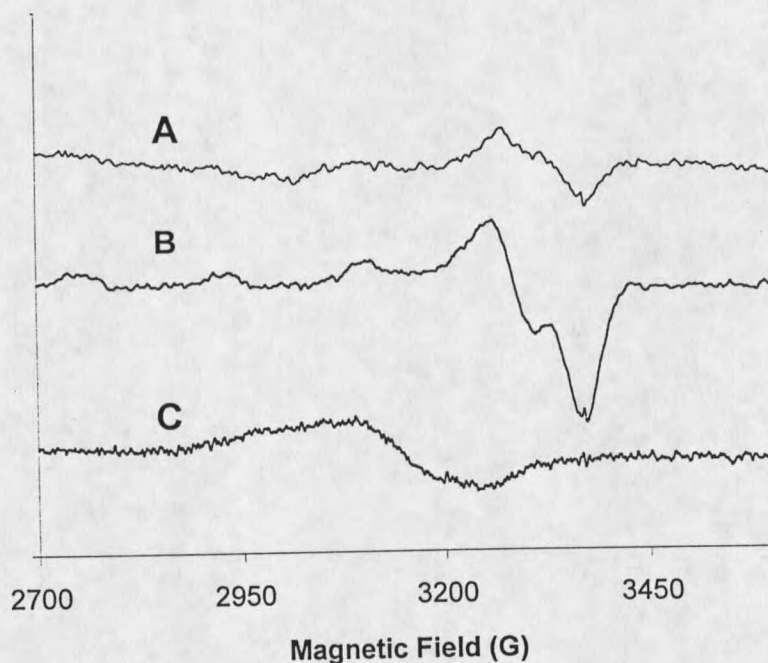


Figure 23. Oxidation of Cu(I)-rNosL (A) by ferricyanide in 10 mM HEPES pH 7.5 (B) and 10 mM KPO₄ pH 6.5 (C). Integrations using CuTEA standards indicate the presence of < 5% Cu(II) in (A), and 100% Cu(II) in (B) and (C). Protein concentrations were 0.49 mM, (A) and (B), and 0.17 mM (C).

vs. 1 mM of the appropriate metal chloride in 20 mM TrisCl pH 7.5, followed by extensive dialysis vs. buffer to remove any extraneously bound metals. Reconstitution was attempted with Mn(II), Co(II), and Co(III). No binding of these metal ions to apo-rNosL was detected. Apparently the copper-binding site of rNosL is highly selective for Cu(I).

Elucidation of the Ligands to Cu(I)-rNosL. The potential role of Cys as a Cu(I) ligand was investigated by reaction of rNosL with the reagents 4,4'-pyridine disulfide (PDS) and 5,5'-

dithiobis(2-nitrobenzoic acid) (DTNB) according to the method of Ramachandran *et al* (144). These reagents react with free Cys, but not disulfide or metal-bound Cys (145,146). Whereas the monomeric apo-protein reacted with PDS to form the 324-nm absorbing species characteristic of the thiopyridine anion, no reaction was seen with Cu(I)-rNosL (Figure 24). Additionally, the dimeric form of apo-rNosL did not react with PDS even under denaturing conditions (4 M GuanidineHCl). Similar results were obtained when the various forms of rNosL were reacted with DTNB.

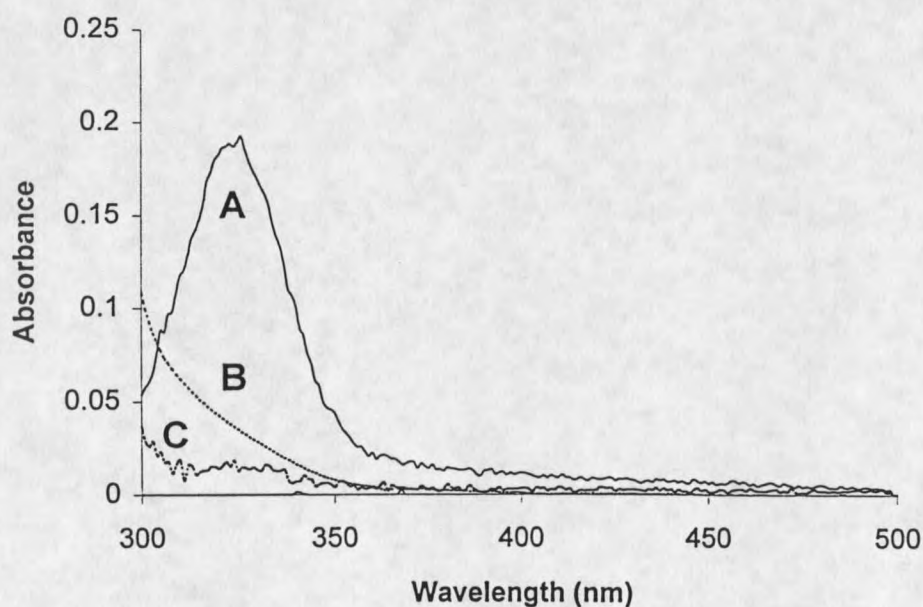


Figure 24. The reaction of various forms of 15 μ M rNosL with the Cys reagent PDS. A, apo-monomer; B, apo-dimer in 4M guanidine HCl; C, Cu(I)-monomer. In A, 100% free cysteine was detected relative to protein concentration.

Recombinant copper-NosL was studied by XAS and EXAFS techniques. XAS confirmed the presence of Cu(I) in rNosL (Figure 25). Fits to the EXAFS data (Figure 25 and Appendix D) suggest that the cuprous ion is three-coordinate, with two sulfurs at an average distance of 2.32 Å, and one N/O donor at 1.99 Å from the copper. The EXAFS data suggest that a His residue is the likely N/O donor.

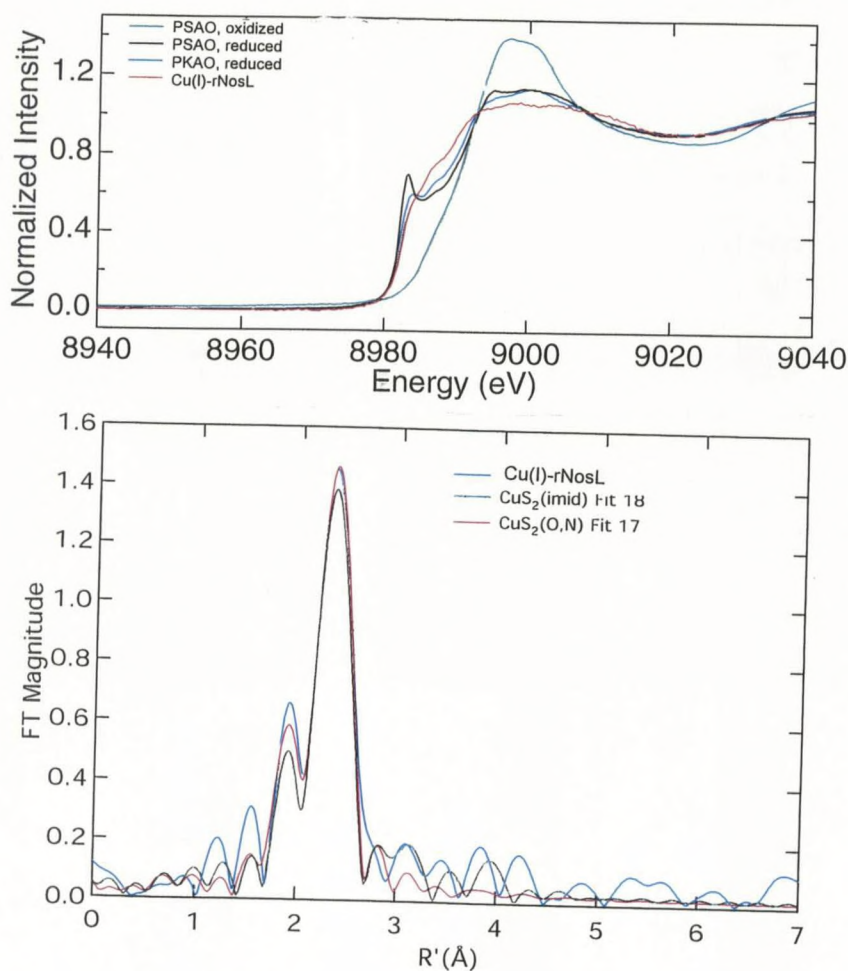


Figure 25. Analysis of the copper site in rNosL. (A) XAS of rNosL compared with the His-ligated copper centers in amine oxidases from pea seedling (PSAO) and porcine kidney (PKAO). (B) Curve fitting results for EXAFS data of Cu(I)-rNosL, compared with CuS₂imidazole and CuS₂(O,N) models. Complete data are available in Appendix D.

Discussion

rNosD. Although the heterologous expression of rNosD was achieved in *E. coli*, the success of the experiments was limited. Periplasmic expression resulted in the production of a recombinant protein that was toxic to the host cells. The use of a pLysS host strain was helpful in allowing small quantities of rNosD to be expressed. Unfortunately, the primary function of LysS is as a lysozyme. Its presence weakens the outer membrane such that osmotic shock is no longer useful as a purification method. Under these conditions, purification would be extremely difficult without the addition of a fusion tag. As an alternative, cytoplasmic expression was attempted, which produced large quantities of rNosD as inclusion bodies. Attempts to resolubilize the protein from this system were unsuccessful.

In recombinant bacterial systems, cell toxicity has been demonstrated to have two main causes (140). The protein itself may be toxic to the host, in which case either a different host should be selected or a low yield of recombinant protein be accepted as the best possible outcome. If the latter scenario is chosen, care must be taken to ensure that the host does not mutate the gene of interest over time in an effort to alleviate its toxicity. Alternatively, the mRNA may cause the toxic effects. For the pET and other high over-expression systems, the rate of transcription is often six-seven times faster than translation. This creates a large steady-state pool of mRNA. For reasons that are unclear, some mRNAs interfere with the normal growth cycle of the host. Miroux and Walker have described two *E. coli* DE3(BL211) derivative strains (C41 and C43) that reportedly have more balanced transcription/translation rates (140). Expression of pD21 in these hosts did not remedy the toxicity problems, however,

suggesting that it is the rNosD protein itself that is responsible for the early death of its host.

The characterization of rNosD was not pursued further because it was anticipated that purification would be very laborious under the available expression conditions. The inability to use osmotic shock to isolate the periplasmic proteins from the large pool of cytoplasmic soluble proteins was considered a major disadvantage. Moreover, SDS/PAGE results on a control strain suggested that the host produced several proteins of similar size to rNosD. These proteins would complicate the identification of rNosD during purification, in the absence of a suitable detection method.

Several alternatives for expressing NosD are possible. Plasmid pD21 could be modified to include a fusion tag, which would allow the rapid purification of the small amounts of rNosD noted for the periplasmic constructs. Alternatively, homologous recombinant expression of NosD could be attempted. An expression system in *A. cycloclastes* has recently been developed and was used successfully for the overproduction of N₂OR, albeit in an inactive form (147). This system uses the constitutive neomycin promoter carried on a broad host range plasmid (133). Another possible strategy is the use of immunoblotting as a detection method. Antibodies to rNosD could be obtained from denatured protein purified from the inclusion bodies produced by the cytoplasmic constructs.

rNosL. The heterologous expression and purification of rNosL has resulted in the first reported characterization of an auxiliary Nos protein. The relative ease of isolation of rNosL from the periplasmic fraction compensates for its relatively low (~ 4 mg/L culture) level of expression and eliminates the need for a fusion tag. The expression construct pL46 replaces the naturally

occurring NosL leader sequence with that of PelB. This design avoids the potential problems associated with the purification of a lipoprotein and allows the use of osmotic shock as a purification method. The assignment of NosL as a periplasmic lipoprotein with a leader sequence recognized by SPaseII appears plausible, considering the large number of bacterial proteins that have been demonstrated to be SpaseII-type lipoproteins (111). The breadth of functionality associated with these proteins and the high specificity of the lipobox sequence for lipid modification also validate this assumption. Rarely has the lipid moiety been required for function in lipoproteins of this type, and these examples involve proteins that facilitate cell surface adhesion (148) or the selective release of other proteins (149). NosL is predicted to be a charged ($pI \sim 4$), soluble protein. Hence, deletion of the lipid moiety and replacement of the N-terminal Cys with Met-Asp are not expected to significantly alter NosL structural characteristics that are related to function. Although NosL has been depicted as being anchored into the outer membrane (2), sequence analysis of the +2 to +5 region does not suggest a particular localization of NosL to either the inner or outer membrane. The tentative designation of NosL as a protein of the outer membrane merely reflects the ratio of known outer membrane to inner membrane proteins, which is roughly 2:1 (111).

Recombinant NosL is a monomer of 18,540 MW that contains one Cu(I) ion, despite being found in the relatively oxidizing environment of the periplasm. Analysis of Cu(I)-rNosL by LC-ES/MS demonstrated the site-specific cleavage of ten amino acids from the C-terminus. The origin of this cleavage has not been investigated, but may be due to the presence of a catalytic amount of protease in the purified protein. Oxidation of the copper site eventually leads to loss of copper and protein dimerization. An apo-dimer form of rNosL is also isolated

when rNosL is aerobically purified from metal-depleted media. In the apo-dimer, the Cys residues (one per monomer) are either oxidized or involved in an intersubunit disulfide bond, as indicated by their lack of reactivity with PDS.

Recombinant NosL appears to have a high specificity for copper, as it does not bind Mn(II), Co(II), or Co(III). The results of PDS and DTNB titrations indicate Cys as a likely Cu(I) ligand. This is not unexpected, since cysteines have also been implicated as Cu(I) ligands in the chaperone proteins Atx1 and CCS (105,107). Comparison of the five available NosL sequences indicates conservation of the single Cys present in the recombinant protein used in this study. This is in addition to the N-terminal Cys found in the wild type protein that is believed to serve as a lipid anchor (Figure 14). The identification of Cys as a copper ligand was also suggested by preliminary analysis of EXAFS data, which indicate a $\text{Cu(I)S}_2(\text{N/O})$ metal center in rNosL, with His as the likely N/O donor. Three methionine residues, including a CysXaaMet motif located near the N-terminus of NosL, are conserved and are candidates for the second S ligand. While no His residues are strictly conserved, each NosL sequence contains at least one His within 5 residues of the CysXaaMet motif. One of these semi-conserved His residues may be a metal ligand. The non-conservation of His ligands in metallochaperones has been previously noted for UreE, the metallochaperone of Ni(II)-urease (108,109).

In contrast with blue copper electron-transfer proteins, and in accord with known copper transfer proteins (102,107), the NosL Cu(I) site is stable to oxygen. The addition of the inner sphere, 2-electron oxidant H_2O_2 had no effect on the oxidation state of copper in rNosL, whereas the outer sphere, 1-electron oxidants ferricyanide and hexachloroiridate were able to oxidize the metal. Interestingly, once copper has been oxidized it has little affinity for rNosL

and is readily removed from the protein. The dependence of the EPR spectrum of oxidized rNosL upon the choice of buffer suggests that, upon oxidation, copper is released from the protein metal-binding site. The ability to release copper may be indicative of the metal-ion transfer property of a metallochaperone.

Comparative sequence analysis shows no homology between NosL and known blue copper electron-transfer proteins. Additionally, its low pI (~ 4) means that NosL is unlikely to be able to bind to the electron-transfer docking site of N_2OR , which is thought to consist of an acidic surface patch in the Cu_A domain (59). Moreover, NosL is produced only in small quantities in its native organism, unlike many electron-transfer proteins. These factors argue against NosL functioning as an electron-transfer protein.

NosL is encoded on the same transcript as three other gene products (NosD, NosF, and NosY) that have been shown to be required for assembly of the active-site dinuclear copper center of N_2OR (96). Thus, it is plausible that NosL is involved in copper incorporation into the active site of N_2OR , perhaps serving as a metallochaperone, similar to CCS, the metallochaperone for superoxide dismutase (102,105), or UreE, the chaperone for urease (108). In *P. stutzeri*, interruption of *nosL* did not result in the inability of the organism to reduce N_2O (98). However, since the interruption occurred late in the protein sequence (after the conserved Cys and Met residues), it is possible that a truncated but functional NosL protein was produced. Alternatively, NosL may not be strictly required for metal incorporation into the catalytic site of N_2OR , but may be necessary only when environmental conditions are copper-limited. For example, the deleterious effects of removing the urease Ni(II)-metallochaperone UreE from the urease gene cluster may be overcome by increasing the cellular nickel.

concentration (109). In the case of CCS, the Atx1-like domain, which is expected to bind Cu(I), is only required for metallation of superoxide dismutase when the concentration of environmental copper is low (131).

Conclusions

The isolation and characterization of rNosL represents a significant step toward understanding the mechanism of copper bioassembly in N₂OR. The finding that rNosL selectively binds stoichiometric Cu(I) ion was unexpected, as no known copper-binding sequence motifs were found by sequence homology studies. This, along with EXAFS data, suggests that rNosL contains a novel type of biological copper center. Assuming that the assignment of NosL as a lipoprotein is valid, conclusions drawn from the investigation of the recombinant protein will be informative for native NosL.

N₂OR from *A. cycloclastes* has been isolated and characterized and may be prepared in both its active, copper-loaded form and an inactive apo-form. This will enable the investigation of potential rNosL/N₂OR interactions to begin. However, the ultimate goal of developing an *in vitro* copper-reconstitution system may require, at minimum, the presence of NosD. Thus, the isolation and characterization of NosD remains an objective for the study of metallocenter assembly in N₂OR.

INVESTIGATIONS OF NOSL FUNCTION

Introduction

The elucidation of protein function is frequently pursued through *in vivo* studies. In fact, the initial assignment of proteins to a particular pathway is often accomplished by random mutagenesis of the organism. Strains that exhibit a change in phenotype are isolated and genes that have been mutated are identified. It was by this method that the *nos* genes were first detected (82,92,96). Once genetic data are available, the manipulation of specific genes is feasible. Typical strategies include site-directed mutagenesis, individual or multiple gene knockouts, and targeted gene interruption. When using these techniques, care must be taken to ensure that changes made in one gene do not cause changes in the control regions of a neighboring gene. This is particularly true for bacteria, where gene regulons often overlap and polycistronic mRNA is frequently encountered.

Knowledge of the functions of the auxiliary Nos proteins has been enhanced by *in vivo* studies. Characterization of N₂OR produced by strains in which *nosD*, *nosF*, or *nosY* were interrupted by Tn5 mutagenesis has been especially insightful. In these variants, neither the transcription of N₂OR, its translation, nor its localization to the periplasm was affected (96). Rather, the mutations caused the production of an inactive form of N₂OR. The inactivation was attributed to the absence of copper in the catalytic site, which could not be reconstituted *in vitro*.

In vivo studies require knowledge of the microbiology of an organism so that it may be genetically manipulated. The microbiology of *A. cycloclastes* has not been widely researched, despite the fact that the first Cu-NiR structure was obtained from this organism (21). No viable plasmids had been previously identified for this host and no methods for gene transfer were documented. Additionally, efforts to create *nos* deletion strains of *A. cycloclastes* have been unsuccessful (147). The discovery that *A. cycloclastes* and *S. meliloti* are closely related (59) suggests several potential alternatives for the genetic manipulation of *A. cycloclastes nos* genes. (*S. meliloti* has been extensively studied because of its important role in alfalfa nodulation.)

The relatedness of *A. cycloclastes* and *S. meliloti* implies that genes from one might be functionally expressed in the other. Additionally, pML-type, broad host range plasmids are known to replicate in high numbers in *S. meliloti* (133). It will be demonstrated that placement of the *A. cycloclastes nos* gene cluster into a pML-based plasmid allows for high homologous over-expression of active N₂OR. Moreover, when this plasmid is transformed into *S. meliloti*, the resulting strain functionally expresses the *A. cycloclastes nos* genes. These observations led to the development of two strategies aimed at circumventing the lack of *A. cycloclastes nos* knockout strains. First, the phenotypes of non-denitrifying *S. meliloti* strains carrying plasmid-borne mutations in *A. cycloclastes nos* genes will be reported. Both *nosL* and *nosZ* deletion strains will be investigated. Second, the results of RNA interference experiments will be presented. In these studies, pML-based plasmids were constructed that carry the codons for *antisense-NosL* under the control of a neomycin promoter. The expression of antisense mRNA has been shown to interfere with the translation of the sense strand and thus prevent protein

expression (150).

Another route to defining protein function is to investigate possible interactions between two or more proteins. These types of *in vitro* studies require the isolation and characterization of the individual proteins. Once the components are available, experiments are designed to detect interactions between the proteins. Such inquiries might involve detecting a multiprotein complex, monitoring enzymatic activity or cofactor transfer, or assessing structural modifications related to the interaction. The purification and characterization of Cu(I)-rNosL was described in the previous chapter. The isolation of both the active and metal-depleted forms of N₂OR has been achieved and will be reported, along with studies of their interactions with rNosL.

Several functions for NosL have been proposed since its initial identification in 1996. A CysXaaXaaCys motif was noted in the first published sequence from *P. stutzeri* (98). This motif is known to be present in the active site of protein disulfide isomerases (151). It was therefore suggested that NosL facilitated disulfide formation in N₂OR, which typically contains 9 – 11 Cys residues per subunit. In the same report, *nosL* was interrupted by insertional mutagenesis with a gentamycin-resistance cassette. The mutant strain showed no impairment in its ability to reduce N₂O. This led the authors to conclude that NosL is not necessary for copper incorporation into N₂OR, and that other protein disulfide isomerases might substitute for NosL in its absence.

Comparison of five NosL sequences demonstrates that only one cysteine of the disulfide isomerase motif is conserved (Figure 14). Therefore, it is unlikely that NosL functions in this capacity. Other than the lipoprotein signal sequence (111), no homology to

proteins of known function is apparent in the lineup. The presence of a lipid anchor suggests that NosL might act as a scaffold, assisting interactions between N₂OR and NosD (59). Hence, NosL would not be strictly required for copper incorporation.

The finding that rNosL selectively binds Cu(I) (*vide infra*) broadens the scope of possible functions for the protein. Copper-containing proteins assume a wide variety of roles in biochemistry. The relatively small size of NosL and the ability of rNosL to bind copper suggest that NosL might be the electron-transfer partner to N₂OR. On the other hand, the acidic nature of NosL (pI ~ 4) argues against this role. The electron-transfer docking site in N₂OR has been predicted to be composed of surface Asp residues (59), making it unlikely that NosL would bind. Moreover, electron-transfer proteins retain their metal cofactors upon oxidation, in contrast to the behavior observed for rNosL. Data will be presented to indicate that rNosL does not function as an electron-transfer partner to N₂OR.

Small Cu(I) proteins have recently been shown to be involved in copper trafficking between proteins (102, 105). The hypothesis that NosL acts as a copper chaperone is consistent with the requirement of auxiliary proteins for copper site assembly in N₂OR. Yet the insertional mutation study in *P. stutzeri* (98) suggests that NosL is not necessary for the production of active N₂OR. However, in that experiment the insertion occurred late in the *nosL* sequence. It is possible that a truncated but functional NosL protein was produced in the mutant strain. This prospect indicates the need for more rigorous phenotype studies. The effects of deleting *nosL* in the heterologous *S. meliloti* expression system (discussed earlier in this chapter) will be reported.

Experimental Methods

Insertion of the *nos* Gene Cluster into a Broad Host Range Plasmid. The construction of a broad host range plasmid bearing the *A. cycloclastes nos* gene cluster was accomplished in collaboration with Dr. John Bollinger and Dr. Marcela Alvarez. Plasmids pHH92 and pEE19 (59) containing overlapping fragments of the *nos* cluster were digested and reassembled to create pCombo, a pUC19-based plasmid carrying the entire *nos* gene cluster and flanking regions. The *nos* cluster was trimmed and excised from pCombo by digestion with *Sbf*I. This 9.5 kb fragment was inserted into the *Sbf*I site of the broad host range plasmid pMLΔ to create plasmid pMLnos. Plasmid pMLΔ was created by removing a *Sal*I fragment from pML122 (133). The 9.5 kb *Sbf*I fragment containing the *nos* cluster was also ligated into the *Pst*I site of pUC19 to produce plasmid pCluster.

Construction of *nosL* Deletion Plasmids. Plasmid pCluster was digested with *Bgl*II, which cuts the *nos* cluster once at position 7582 in *nosL* (Appendix B). The linearized DNA was then digested for varying lengths of time with *Bal*31 exonuclease, which digests double stranded DNA. The digests were then precipitated, re-ligated, and transformed into chemically competent Top10F' *E. coli* cells (Invitrogen) using standard molecular biology procedures. Colonies containing truncated *nosL* genes were selected by PCR screening, using primers D7883L (5'TCCGAAAGGTGACCCCATGAGG) and U527L (5'GAGATAGTCT-TCGTCGTCCGC). Primer D7883L contains the stop codon for *nosY*, which overlaps the *nosL* initiation codon. Primer U527L matches *nosL* sequence downstream of the *Bgl*II site.

Consequently, any plasmids that yield PCR products with these primers should contain the entire *nosY* gene and any upstream *nosX* regulatory elements. Plasmid pClusLΔ11 was chosen from three candidates. DNA sequencing (performed on an ABI Prism 310 using the BigDye terminator kit) revealed that 215 nucleotides (7453 – 7668, Appendix B) have been deleted from *nosL* in pClusLΔ11. The mutant sequence is frame-shifted with respect to the original sequence, such that codons for only the first 13 amino acids of NosL are in-frame. The mutated *nos* cluster was cut from pClusLΔ11 with *SbfI* and spliced into pMLΔ to produce plasmid pMLLΔ11.

Construction of a *nosZ* Deletion Plasmid. Plasmid pCluster was digested with *FseI* and a 793-bp internal fragment (nucleotides 2569 – 3362) removed from *nosZ*. The plasmid was then re-ligated to create plasmid pClusFseΔ. The mutated *nos* cluster was cut from pClusFseΔ with *SbfI* and spliced into pMLΔ to produce plasmid pMLFseΔ.

Construction of *nosL* Antisense Plasmids. A transcriptional terminator cassette (called oligoTT) was assembled by annealing oligos TTXF (5'TCGAGCTGATTATCGAT-TCCCCATGCTGCTGACCCGACGCCGCTTGATTGCAATTTCTGCAGTTCTCGCAA) and TTCR (5'CGTTGCGAGAACTGCAGAAATTGCAATCAAGCGGCGTCGG-GTCAGCAGCATGGGGAATCGATAATCAGC). This cassette contains a 5' *XhoI* overhang (*XhoI* site is retained after ligation) followed by a spacer region and a new *Clal*

site. Immediately downstream of the *ClaI* site is the *nosL* transcriptional terminator sequence (nucleotides 8001 – 8053, Appendix B) followed by a *ClaI*-compatible overhang (this destroys the original *ClaI* site after ligation). The oligoTT cassette was ligated into pML122 prepared by digestion with *XhoI* and *ClaI*. These restriction sites are found in the MCS directly downstream of the neomycin promoter in pML122. The resulting plasmid was called pMLTT.

An antisense *nosL* fragment was produced by PCR amplification, using primers to amplify nucleotides 8000 - 7463 and introduce a 5' *XhoI* site and a 3' *ClaI* site into the PCR product. After restriction digest with *XhoI* and *ClaI*, the PCR product was ligated into the pMLTT vector (similarly prepared by restriction digest). The resulting construct was verified by PCR and called pTTal. The *nos*-containing *SbfI* fragment of pCluster was isolated as previously described and ligated into the *SbfI* site of pTTal to produce plasmid pTTalnos.

Introduction of pML122-Based Plasmids into *A. cycloclastes*. Electrocompetent *A.*

cycloclastes cells were prepared using a modification of a procedure developed by Dr. John Bollinger. A 50 mL culture of *A. cycloclastes* was grown in LB media at 30°C to $OD_{600} = 0.5$. The culture was placed on ice for two hours, then centrifuged at 4°C for 10 minutes at 5,000 rpm (SS34 rotor) to collect the cells. The pellet was washed once with 40 mL of ice-cold water, resuspended in 40 mL of ice-cold water, then allowed to sit on ice for an additional three hours. The cell suspension was centrifuged, washed with 20 mL of ice-cold 10% glycerol, and resuspended in residual 10% glycerol. After sitting on ice for one hour, the cells

were used for electroporation. An Invitrogen Electroporator was used with 0.1-cm sterile cuvettes. Settings were 50 μ F/1800V and 150 ohms. Approximately 200 – 400 ng plasmid DNA in 1 – 2 μ L was mixed with 50 μ L prepared cells. After electroporation, the cells were diluted with 1.0 mL SOC media, transferred to a sterile culture tube, and incubated at 30°C with shaking for one hour. Approximately 200 μ L cell suspension was plated on LB-agarose containing 50 μ g/mL gentamycin (pML122 selection marker). The plates were incubated for 2 – 3 days at 30°C.

Introduction of pML122-Based Plasmids into *S. meliloti*. Plasmids were transformed into *S. meliloti* strains RMYC2160 and RMYC2164, which were provided by Dr. Y-K Chan (81). RMYC2160 is a spontaneous rifamycin-resistant Balsac strain of *S. meliloti* that does not contain denitrifying genes. RMYC2164 is a nalidixic acid-resistant strain derived from *S. meliloti* ATCC9930. This strain contains the gene clusters for functional NaR, NiR, and NOR, but not N₂OR. Transformation was performed by conjugation with *E. coli* using pRK2013 as the helper plasmid (81).

GC Assay of N₂O Reductase Activity. One mL cultures (2 – 3 replicates per sample) were placed in sterile 12-mL vials equipped with screw caps and rubber septa. After purging the vials for ten minutes with argon, 200 μ L N₂O was introduced into the head space by gas-tight syringe. The vials were slowly shaken at room temperature over the course of the experiments. The amount of N₂O remaining in the head space was monitored over time by

gas chromatography (GC). Aliquots of 50 μL were removed and injected into a Hewlett Packard Series II 5890 gas chromatograph equipped with a GS-Q capillary column (J&W Scientific) and an electron capture detector. N_2O standards were prepared by injecting 50, 100, 150, 200, or 250 μL N_2O into vials containing 1 mL of water. Standard curves were constructed at every time point to account for the small, non-enzymatic loss of N_2O .

For *A. cycloclastes* strains, cultures were grown at 30°C and 240 rpm to $\text{OD}_{600} = 1.0$ in LB media containing 50 $\mu\text{g/mL}$ gentamycin. For RNA interference experiments, 50 $\mu\text{g/mL}$ kanamycin was also added (*A. cycloclastes* is Km^{R} (147)). The induction of *nos* genes was accomplished by adding 10 mM KNO_3 , $\text{Cu}(\text{NO}_3)_2$ as required, and lowering the shaking rate to 120 rpm to attain micro-aerobic conditions. Cultures were grown for an additional 24 hours prior to GC analysis. For *S. meliloti* strains, cultures were grown at 30°C and 150 rpm to $\text{OD}_{600} = 0.5$ in TYC media (81) containing 50 $\mu\text{g/mL}$ gentamycin. After the addition of 10 mM KNO_3 and any required $\text{Cu}(\text{NO}_3)_2$, cultures were grown for 10 – 24 hours prior to GC analysis.

Purification of N_2OR . A 500-mL culture of *A. cycloclastes* carrying plasmid pMLnos (strain AC/pMLnos) was grown and induced as described in the previous section. For the isolation of active N_2OR , 50 μM $\text{Cu}(\text{NO}_3)_2$ was added at the time of induction and an argon atmosphere was maintained during the purification. Cells were collected by centrifugation, washed once with 10 mM MES pH 6.5 (Buffer A), and frozen at -70°C overnight. The thawed cells were lysed with B-Per Reagent (Pierce) and centrifuged. The soluble fraction

was gently mixed with 50 mL DEAE Sepharose FastFlow (Sigma) equilibrated in Buffer A. After 30 minutes, the slurry was packed into a 5 x 15 cm glass column. The gel media was washed with 10 volumes of Buffer A. N₂OR was then eluted with Buffer A + 0.5 M NaCl. The eluent was concentrated to ~20 mL and dialyzed vs. Buffer A. N₂OR was purified to homogeneity by gradient elution on a MonoQ HR10/10 column, followed by gel filtration on a 1.6 x 100 cm Sephacryl S200HR column in Buffer A + 0.2 M NaCl. The activity of N₂OR was measured spectrophotometrically using benzyl viologen as the electron donor (39). The copper content of protein samples was measured by AAS or ICP-AES. Protein concentrations were determined using an extinction coefficient at 280 nm of 19.2 cm⁻¹ for a 1% solution (152).

Interactions of rNosL and apo-N₂OR. A 50 μ L sample of 0.174 mM apo-N₂OR was anaerobically added to 400 μ L of 0.59 mM Cu(I)-rNosL in Buffer A. After a 2-hour incubation at room temperature, the sample was applied to a 1.6 x 100 cm S200 gel filtration column equilibrated in anaerobic Buffer A. Two-mL fractions were collected and those containing protein were analyzed by UV/Vis spectroscopy and ICP/AES (copper determination).

Reduction of N₂OR by rNosL. The ability of Cu(I)-rNosL to reduce N₂OR was tested by monitoring the visible absorption bands associated with the Cu_A center of N₂OR. One mL of 22.5 μ M N₂OR (45 μ M Cu_A electron transfer sites) in Buffer A was anaerobically mixed with an 8-fold excess of Cu(I)-rNosL. As a control, N₂OR was reduced with ascorbate.

Oxidation of Cu(I)-rNosL During the Catalytic Turnover of N₂OR. The EPR spectrum of 0.4 mM Cu(I)-rNosL was monitored before and after the addition of a catalytic amount of N₂OR (~2 μ M) in an N₂O atmosphere. Reduced benzyl viologen, which serves as an electron donor in the spectrophotometric assay for N₂OR, was used in place of Cu(I)-rNosL in the control experiment. EPR conditions are the same as in Figure 23A.

Results

Purification and Preliminary Characterization of Apo- and Holo-N₂OR. N₂OR was purified from strain AC/pMLnos to ~95% homogeneity as judged by SDS/PAGE. When 50 μ M Cu(NO₃)₂ was added to growth media, holo-N₂OR was obtained. The enzyme contained ~ 8 copper ions per dimer (152) and had a specific activity of 1.2 μ mol/min/mg. The visible absorption spectrum of N₂OR is presented in Figure 26. The absorbance bands at 480, 540, 620, and 760 nm are associated with the Cu_A center of N₂OR. In contrast, the omission of exogenous copper resulted in the purification of apo-N₂OR. This protein was inactive, contained ~ 2 copper ions per dimer, and was devoid of spectral features in the visible range.

The Possible Function of rNosL as an Electron-transfer Protein. To ascertain if electron transfer between Cu(I)-rNosL and N₂OR was possible, an 8-fold excess of rNosL (relative to Cu_A sites) was anaerobically added to N₂OR. The visible absorbance bands of N₂OR are bleached when the enzyme is reduced with ascorbate. Dithionite, a stronger reductant, causes the formation of a

blue form of N_2OR that displays a single broad absorption band at 650 nm. No changes in the visible spectrum of N_2OR were observed when rNosL was added.

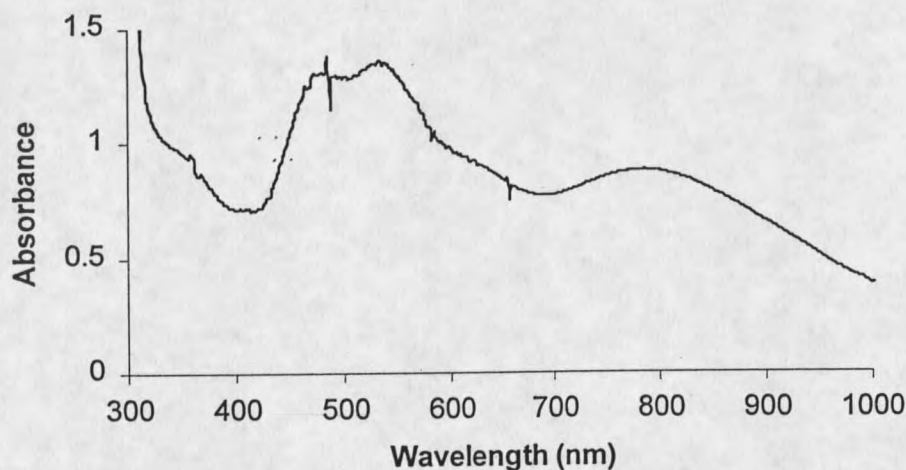


Figure 26. Visible absorption spectrum of *A. cycloclastes* N_2OR (30 mg/mL) in 10 mM MES pH 6.5 + 0.2 M NaCl. Path length, 1 cm.

To further test the ability of Cu(I)-rNosL to function as an electron transfer partner to N_2OR , the EPR spectrum of 0.4 mM Cu(I)-rNosL was monitored before and after the addition of a catalytic amount of N_2OR in an N_2O atmosphere (Figure 27). Reduced benzyl viologen, which serves as an electron donor in the spectrophotometric assay for N_2OR (39) was used in place of Cu(I)-rNosL in a control experiment. Whereas benzyl viologen lost its purple color within 1 minute of the addition of N_2O (indicating that it had been oxidized), the EPR spectrum

of rNosL remained unchanged after 20 minutes. This indicates that rNosL does not support catalysis by N_2OR and does not function as an electron donor to the enzyme.

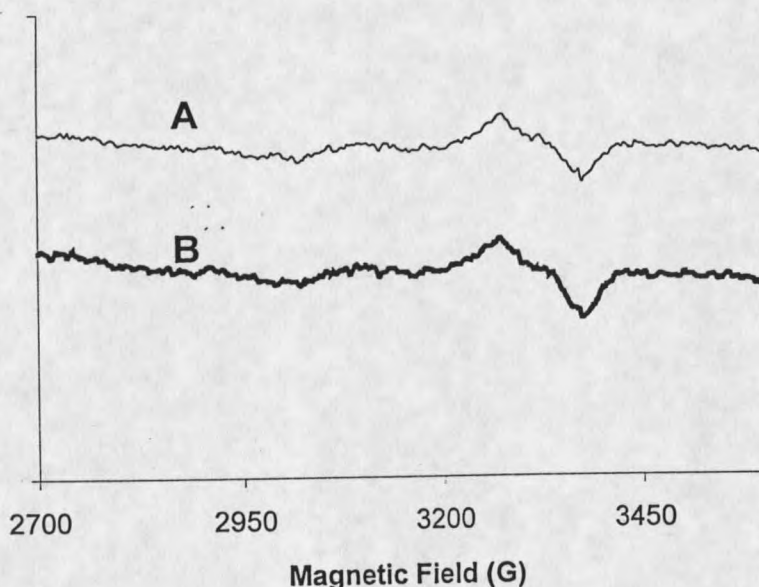


Figure 27. EPR spectra of 0.4 mM rNosL and 2 μM N_2OR before (A) and after (B) the addition of N_2O . Integrations vs. a CuTEA standard show the presence of < 5% EPR-detectable copper in the samples. EPR conditions are identical to those given in Figure 23A.

In vitro Study of the Possible Function of rNosL as a Copper Chaperone to N_2OR . The possibility that apo- N_2OR and rNosL would form a complex was investigated. A 450 μL mixture of 0.52 mM rNosL and 0.019 mM apo- N_2OR in Buffer A was anaerobically incubated for two hours. This represents a 3.4-fold excess of Cu(I)-rNosL over potential copper sites in N_2OR . After incubation, the sample was applied to a gel filtration column equilibrated in anaerobic Buffer A. Comparison of the chromatogram (Figure 28) with those obtained from

individual runs of rNosL and apo-N₂OR shows that the two proteins cleanly separate according to their respective sizes. Furthermore, copper analysis of fractions 17 and 23 shows that no metal transfer has occurred between the two proteins (Table 1). No evidence for protein-protein interactions between rNosL and N₂OR was detected.

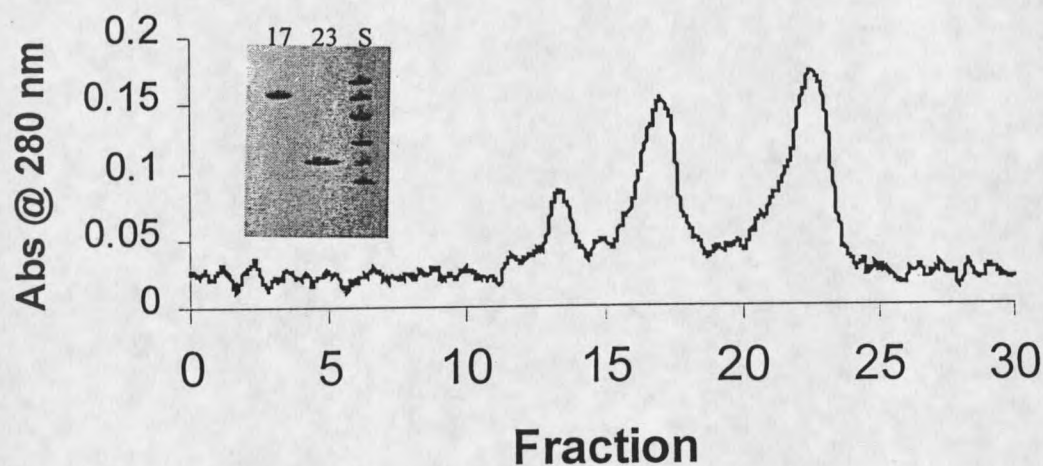


Figure 28. Chromatogram of a mixture of rNosL and apo-N₂OR. The first peak (fractions 13 and 14) represents a heme impurity found in apo-N₂OR. Inset: SDS/PAGE of fractions 17 and 23. S, standards as described in Figure 19.

Table 1. Copper Analysis of rNosL and N₂OR

Sample	Copper:Protein Ratio
rNosL as isolated	0.9
apo-N ₂ OR as isolated	2.3
fraction 17 (N ₂ OR)	2.1
fraction 23 (rNosL)	0.9

In vivo Studies of the Possible Function of rNosL as a Copper Chaperone to N₂OR. The lack of a *nos* minus strain of *A. cycloclastes* hinders phenotype studies designed to elucidate the function of NosL. Therefore, the phenotype studies were pursued using another host. Two strains of *S. meliloti* were obtained from Dr. Y-K Chan. RMYC2160 is a derivative of the *Balsac* species and contains genes associated with denitrification. RMYC2164 is a derivative of ATCC9930 and carries the *nar*, *nir*, and *nor* gene clusters, but not the *nos* cluster (81). The introduction of pMLnos into these strains confers N₂OR activity to the organisms (Table 2). However, the rate of N₂O reduction is greatly reduced compared with the *A. cycloclastes* native strain. It was also noted that the N₂OR-overproducing strain AC/pMLnos displays the same rate of N₂O reduction as the *A. cycloclastes* native strain under these assay conditions. The addition of copper to the growth media was necessary for maximal activity. SDS/PAGE analysis of the pMLnos and pMLLΔ11 variants of RMYC2160 and RMYC2164 indicated that little N₂OR is produced in these strains, as no bands of the expected MW were visible.

Table 2. Rates of N₂O reduction in different strains of *S. meliloti*

Strain	Rate ^{a,b}
<i>A. cycloclastes</i>	2.6 ± 0.2
AC/pMLnos	2.8 ± 0.3
RMYC2160	0
RMYC2160/pMLΔ	0
RMYC2160/pMLFseΔ	0
RMYC2160/pMLnos	0.3 ± 0.05
RMYC2160/pMLLΔ11	0.3 ± 0.05
RMYC2164	0
RMYC2164/pMLFseΔ	0
RMYC2164/pMLnos	0.7 ± 0.2
RMYC2164/pMLLΔ11	0.7 ± 0.2

^a μmol per hour per mL culture. ^b 10 μM Cu^{II} added to growth media.

Strains RMYC2160/pMLnos and RMYC2164/pMLnos displayed a slow loss of N₂OR activity when grown in continuous culture. Therefore, fresh cultures were prepared prior to each GC experiment when the phenotypes of strains containing pMLLΔ11 were evaluated. When 10 μM copper was included in the growth media, no differences in the rates of N₂O reduction were noted between strains RMYC2164/pMLnos and RMYC2164/pMLLΔ11 or between strains RMYC2160/pMLnos and RMYC2160/pMLLΔ11. Identical rates were obtained when the experiments were repeated with media supplemented with either 5 or 100 μM copper.

The phenotype experiments in *S. meliloti* indicate that NosL is not necessary for the bioassembly of active N₂OR when the bacteria are grown in copper-supplemented media. Therefore, the possibility that NosL might be required under copper-limited growth conditions was investigated. Experiments of this sort require the ability to manipulate the ratio of N₂OR produced to the available copper supply. Considering the low levels of N₂OR that are produced in the *S. meliloti* strains, it was decided to abandon heterologous expression in lieu of homologous RNA interference trials.

Comparative experiments measuring the rates of N₂O reduction were performed between strains AC/pMLnos and AC/pTTalnos and strains AC/pMLΔ and AC/pTTal. Copper concentrations in the growth media were varied from 0, 1, 2, 3, 4, 5, 10, 20, and 50 μM, depending on the experiment. The relative amounts of N₂OR in the samples were assessed by SDS/PAGE and Immunoblotting (59) and found to be the same within the error of the techniques. The results of a representative experiment are displayed in Table 3. No differences in the dependence of N₂OR activity on copper content were noted between

similar strains with or without the *antisenseNosL* coding sequence.

Table 3. Rates^{a,b} of N₂O reduction at different concentrations of copper in the growth media.

	0 $\mu\text{M Cu}^{\text{II}}$	2 $\mu\text{M Cu}^{\text{II}}$	3 $\mu\text{M Cu}^{\text{II}}$	4 $\mu\text{M Cu}^{\text{II}}$	5 $\mu\text{M Cu}^{\text{II}}$	20 $\mu\text{M Cu}^{\text{II}}$
AC/pTTalnos	0	0	0	0.2	0.3	2.2
AC/pMLnos	0	0	0	0.40	0.3	2.2

^a $\mu\text{mol per hour per mL culture}$. ^bError in rates was estimated to be 20% for rates < 1.0 and 10% for rates ≥ 1.0 .

Discussion

The purification of *A. cycloclastes* N₂OR is a significant step toward understanding the mechanism of the bioassembly of its active-site copper center. Yields of pure N₂OR from the over-expressing strain AC/pMLnos ranged from 15 – 40 mg per liter of cell culture. When isolated from copper-supplemented cultures, N₂OR was active and fully metallated (8 coppers per protein dimer). In contrast, the isolation of N₂OR from unsupplemented cultures produced a protein that only contained 25% copper. This enzyme is useful for studying potential interactions and copper transfer reactions between N₂OR and rNosL.

It is intriguing to speculate on the location of copper in this metal-depleted form (referred to as apo-N₂OR). One possibility is that the isolated protein is a mixture of 25% fully-loaded copper N₂OR and 75% fully-depleted N₂OR. However, the featureless visible absorbance spectrum of apo-N₂OR indicates that no Cu_A center is present in this enzyme form. It may also be inferred from this observation that the copper is not occupying 50% of the available Cu_A centers. One reasonable interpretation of the partial copper loading is that the

metal is adventitiously bound and does not occupy either Cu_A or Cu_{Cat} . Alternatively, it may be Cu_{Cat} that is partially occupied in this enzyme form. The Nos proteins required for the bioassembly of Cu_{Cat} are present in strain AC/pMLnos. Furthermore, if the Farrar/Thomson model for Cu_{Cat} is correct (130), only relatively weak visible absorption bands for a Cu_{Cat} -only form of N_2OR would be anticipated. The isolation of a Cu_A -free, Cu_{Cat} -loaded form of N_2OR would tremendously advance the elucidation of N_2OR structure and mechanism. Thus, the characterization of this form of the enzyme warrants further investigation.

The observation was made that wild type *A. cycloclastes* and its derivative strain AC/pMLnos show similar rates of N_2O reduction in the GC assay. This was surprising, as AC/pMLnos produces 10 – 100X more active N_2OR per mL culture than its parent strain. Together, the results indicate that biological N_2O reduction in AC/pMLnos may be limited by the efficiency of one or more components of the electron transport chain.

The possibility that rNosL functions as an electron donor to N_2OR was investigated. First, the ability of rNosL to reduce the Cu_A center of N_2OR was tested. No reduction was seen, as evidenced by the lack of change in the N_2OR visible spectrum. Next, the possibility that substrate binding might trigger electron transfer between rNosL and N_2OR was investigated. The large amount of rNosL that would be required to sustain measurable turnover in the N_2OR GC assay (41) is prohibitive. Therefore, an alternative strategy was employed in which the concentration of the electron donor was the limiting factor. During turnover, the electron donor becomes oxidized as it transfers electrons to N_2OR . For the artificial electron-donor benzyl viologen, oxidation is monitored by following the loss of the purple color of reduced benzyl viologen. This is the basis of the spectrophotometric assay (39). If rNosL were to donate

electrons to N_2OR during turnover, its newly oxidized copper center would become detectable by EPR. The EPR spectrum of rNosL did not change in the presence of a catalytic amount of N_2OR when large excesses of substrate were added. This indicates that rNosL is incapable of functioning as an electron-transfer partner to N_2OR .

The potential role of rNosL as a copper chaperone to N_2OR was investigated. Incubation of excess rNosL with apo- N_2OR did not result in the formation of a complex that was detectable by gel filtration. Moreover, the individual copper contents of the two proteins were identical before and after incubation, which indicates that no *in vitro* copper transfer had occurred.

The functional expression of *A. cycloclastes nos* genes carried on a pML-based plasmid in *S. meliloti* has been demonstrated, but only at low levels. The pML family of vectors is known to achieve high copy numbers in *S. meliloti*. Therefore, the low N_2OR expression levels observed for strains RMYC2160/pMLnos and RMYC2164/pMLnos suggest that the *A. cycloclastes* regulatory elements are not well-recognized by *S. meliloti* transcription factors. Furthermore, a gradual loss of N_2OR activity during continuous culture was noted. Antibiotic selection pressure is known to decrease as the concentration of bacteria rises (114). Hence, the data might indicate that the transcription of *A. cycloclastes nos* genes occurs merely as a result of transcriptional read-through at times when other genes on the plasmid are being heavily transcribed.

The rates of N_2O reduction in *S. meliloti* strains carrying pMLnos and strains carrying a *nosL* deletion (pML Δ 11) were compared (Table 2). No differences in N_2OR activity were noted. These data clearly indicate that NosL is not required for copper site assembly in N_2OR . The possibility remains that NosL might be conditionally required. For example, NosL might

provide a source of copper when the environmental levels of the metal become depleted. Experiments to test this hypothesis were not conducted in *S. meliloti* because of the low level of N₂OR expression in these strains. It was thought that accurately regulating such low levels of copper in the culture media would be difficult.

The technique of RNA interference was used to simulate an *A. cycloclastes nosL* knockout strain. The coding sequence for nucleotides 7463 – 8000 was inserted into the MCS of vector pMLTT such that its orientation was opposite that of the vector-derived neomycin promoter. Transcription of this sequence will produce mRNA that is complementary to the *nosL* transcript. The complementary mRNAs should form a duplex and prohibit translation of *nosL*. The first 44 N-terminal codons were omitted from the antisense construct to avoid potential problems with *nosY* translation. The comparative experiments between strains AC/pMLTT and AC/pTTal and strains AC/pMLnos and AC/pTTalnos showed no correlation of the cellular N₂OR activity with copper content. Significantly, this occurred under conditions where the activity was sensitive to the copper content (Table 3).

These results might indicate that NosL is not required under copper-limited growth conditions. However, several potential flaws in the experimental design exist. First, it is possible that NosL plays a catalytic role in the bioassembly of N₂OR and that the RNA interference technique was not 100% effective in preventing translation of *nosL*. In this case, even a small amount of NosL present in the cell could produce active N₂OR. Second, it is important to recall the insensitivity of the GC assay to the actual level of active N₂OR expression. Instead of reflecting a true difference in the amount of active N₂OR produced at the varying copper concentrations, the data may be measuring the response of another component. In this light, it is

important to recall that other *A. cycloclastes* denitrification enzymes are copper proteins. In particular, pseudoazurin has been identified as the electron-transfer partner to Cu-NiR (153). If pseudoazurin is also the physiological electron donor to N₂OR, then the copper sensitivity of the various *A. cycloclastes* strains may be reflecting the amount of metallated pseudoazurin present in the cells. Consequently, a definitive experiment of the conditional requirement of NosL for bioassembly of the copper centers in N₂OR would minimally require two things: (1) the development of a true *A. cycloclastes nosL* knockout strain to ensure a total *nosL* minus background; and (2) the purification of N₂OR from this strain. This would allow the direct comparison of the copper loading and specific activities of the purified enzymes produced under varying environmental copper levels.

Conclusions

While the results presented in this chapter do not suggest a definitive role for NosL, they do provide information about functions that NosL does not perform. Assuming that recombinant NosL behaves as the wild type protein would, NosL does not function as an electron-transfer partner to N₂OR. The expression of the *A. cycloclastes nos* gene cluster in *S. meliloti* has also permitted conclusive phenotype experiments to be performed. Deletion of *nosL* had no effect on N₂OR activity. No possibility of producing a truncated but functional NosL protein exists in the *S. meliloti* strains used for this research. This was a concern in the previously published study on *nosL* from *P. stutzeri*, which also showed no change in phenotype associated with interrupting *nosL* (98).

The possibility remains that NosL is conditionally required for the bioassembly of the catalytic copper center of N₂OR. The observation that *nosL* is conserved among the five known *nos* clusters argues that it plays some part in denitrification. Resolution of this point will require additional research and advances in the genetic manipulation of *A. cycloclastes*. Alternatively, another denitrifying organism for which knockout strains are feasible could be used in the phenotype studies. Homologous expression of the *S. meliloti nos* cluster in RMYC2164 or RMYC2160, using pML-based constructs, seems ideal for such investigations.

RESEARCH SIGNIFICANCE AND FUTURE DIRECTIONS

The primary goal of this research project was to investigate the mechanism of copper center assembly in N₂OR. At the onset of the project, the notion that copper enzymes might acquire their metallocofactors from other proteins was just beginning to be explored by scientists working in the field of bioinorganic chemistry. N₂OR was considered an attractive candidate for this type of study. Phenotype studies in *P. stutzeri* had indicated that additional proteins were required for copper insertion and/or processing in N₂OR. However, no auxiliary Nos protein had been previously isolated or characterized, and genetic information on the *nos* gene cluster was available for only one organism.

The research described here had made several contributions to the understanding of copper site assembly in N₂OR. The elucidation of the complete *nos* cluster sequence from *A. cycloclastes* helped to refine structure/function predictions for the Nos proteins. Techniques to explore *nos* function and over-express Nos proteins in the native host denitrifiers (*A. cycloclastes* and *S. meliloti*) have been developed. Heterologous expression in *E. coli* was also successful.

Perhaps the most significant contribution is the isolation and characterization of a recombinant form of NosL. rNosL has been shown to stoichiometrically bind Cu(I). Moreover, it appears to contain a novel type of biological copper site. The function of NosL is still not understood, but it does not act as the electron-transfer partner to N₂OR, nor is it strictly required for copper site assembly in the enzyme.

Further characterization of rNosL should be pursued. It would be worthwhile to measure the reduction potential of the Cu(I) center and to determine the three-dimensional structure of the protein. Both crystallization trials and 2-D NMR studies should be pursued. Site directed mutagenesis of potential copper ligands (Cys, His, and Met) may be informative on the nature of the metal center. At the same time, efforts to homologously express NosL should be pursued. If the wild type protein proves difficult to purify, its modification with a C-terminal fusion tag should be attempted. The confirmation that NosL is a lipoprotein would validate hypotheses that are based on studies of the recombinant protein.

The study of copper site assembly in N₂OR will require further investigation to understand the process at the molecular level. Both *in vitro* and *in vivo* studies should be pursued. If efforts to produce *nos* knockout strains in *A. cycloclastes* continue to fail, then another organism should be selected. *S. meliloti* is an attractive alternative. Both *nos* minus strains and non-denitrifying strains of *S. meliloti* are available, in addition to clones containing the *S. meliloti nos* gene cluster. The homologous expression of the *S. meliloti nos* gene cluster carried on the broad host range vector pML122 should provide the same level of over-expression as in *A. cycloclastes*.

A main focus of future research should be NosD. Efforts to isolate and characterize this protein should be made, using the strategies outlined in the third chapter. The HisXaaMet motif that is conserved in five NosD sequences may be part of a novel copper-binding site. Moreover, if NosL is not the copper chaperone to N₂OR, it is logical to suppose that NosD may play that role. It may be that NosL is the copper chaperone to NosD, being necessary

only when the levels of environmental copper are low. Hence, once NosD is purified and characterized, potential protein-protein interactions between NosD and N₂OR, and NosD and NosL should be investigated. Finally, it is important to remember that sequence analysis suggests that copper site assembly occurs by an energy-requiring process across a membrane. Therefore, *in vitro* reconstitution of the catalytic copper center in N₂OR may not be possible without the addition of NosF and NosY, and may require the correct localization of the various components in a membrane-like system.

REFERENCES CITED

1. Payne WJ. *Denitrification*. New York: Wiley, 1981.
2. Zumft WG. Cell biology and molecular basis of denitrification. *Microbiol Rev* 1997; 61:533-616.
3. Kobayashi M, Matsuo Y, Takimoto A, Suzuki S, Maruo F, Shoun H. Denitrification, a novel type of respiratory metabolism in fungal mitochondrion. *J Biol Chem* 1996;271:16263-16267.
4. Tsuruta T, Takaya N, Zhang L, Shoun H, Kimura K, Hamamoto M, Nakese T. Denitrification by yeasts and occurrence of cytochrome P450_{nor} in *Trichosporon cutaneum*. *FEMS Microbiol Lett* 1998;168:105-110.
5. Berks BC, Ferguson SJ, Moir JWB, Richardson DJ. Enzymes and associated electron transport systems that catalyse the respiratory reduction of nitrogen oxides and oxyanions. *Biochim Biophys Acta* 1995;1232:97-173.
6. Mosier AR, Schimel DS. Influence of agricultural nitrogen on atmospheric methane and nitrous oxide. *Chem Ind* 1991;874-877.
7. Smil V. Global population and the nitrogen cycle. *Sci Am* 1997;July:76-81.
8. Spector M. Production and decomposition of nitrous oxide during biological denitrification. *Water Environ Res* 1998;70:1096-1098.
9. Baird C. The chemistry of the ozone layer. In: *Environmental Chemistry*. New York: WH Freeman, 1999:15-84.
10. Krogulska B, Mycielski R. Removal of nitrogen from industrial wastewaters by three-step nitrification and denitrification. *Acta Microbiol Pol* 1979;28:301-309.

11. Wiesmann U. Biological nitrogen removal from wastewater. *Adv Biochem Eng Biotechnol* 1994;51:113-154.
12. Cizinska S, Mateju V, Krejci J, Janoch T, Kyslikova E. Water denitrification by an immobilized biocatalyst. *J Chem Tech Biotechnol* 1992;55:33-38.
13. Thiemens MH, Trogler WC. Nylon production: An unknown source of atmospheric nitrous oxide. *Science* 1991;251:932-934.
14. Cofer WR, III, Levine JS, Winstead EL, Stocks BJ. New estimates of nitrous oxide emissions from biomass burning. *Nature* 1991;349:689-691.
15. Christensen JP, Murray JW, Devol AH, Codispoti LA. Denitrification in continental shelf sediments has major impact on the oceanic nitrogen budget. *Global Biogeochem Cycles* 1987;1:97-116.
16. Seitzinger SP, Giblin AE. Estimating denitrification in North Atlantic continental shelf sediments. *Biogeochemistry* 1996;35:235-260.
17. Bedzyk L, Wang T, Ye RW. The periplasmic nitrate reductase in *Pseudomonas* sp strain G-179 catalyzes the first step of denitrification. *J Bacteriol* 1999;181:2802-2806.
18. Liu HP, Takio S, Satoh T, Yamamoto I. Involvement in denitrification of the *napKEFDABC* genes encoding the periplasmic nitrate reductase system in the denitrifying phototrophic bacterium *Rhodobacter sphaeroides* f. sp. *denitrificans*. *Biosci Biotechnol Biochem* 1999;63:530-536.
19. Bell LC, Richardson DJ, Ferguson SJ. Periplasmic and membrane-bound respiratory nitrate reductases in *Thiosphaera pantotropha*: The periplasmic enzyme catalyzes the first step in aerobic denitrification. *FEBS Lett* 1990;265:85-87.
20. Antipov AN, Lyalikova NN, Khiznjak TV, L'vov NP. Some properties of dissimilatory nitrate reductases lacking molybdenum and molybdenum cofactor. *Biochemistry-Moscow* 1999;64:483-487.

21. Adman ET, Godden JW, Turley S. The structure of copper-nitrite reductase from *Achromobacter cycloclastes* at five pH values, with NO₂⁻ bound and with type II copper depleted. *J Biol Chem* 1995;270:27458-27474.
22. Kukimoto M, Nishiyama M, Murphy MEP, Turley S, Adman ET, Horinouchi S, Beppu T. X-ray structure and site-directed mutagenesis of a nitrite reductase from *Alcaligenes faecalis* S-6: Roles of two copper atoms in nitrite reduction. *Biochemistry* 1994;33:5246-5252.
23. Strange RW, Murphy LM, Dodd FE, Abraham ZHL, Eady RR, Smith BE, Hasnain SS. Structural and kinetic evidence for an ordered mechanism of copper nitrite reductase. *J Mol Biol* 1999;287:1001-1009.
24. Dodd FE, Van Beeumen J, Eady RR, Hasnain SS. X-ray structure of a blue-copper nitrite reductase in two crystal forms. The nature of the copper sites, mode of substrate binding and recognition by redox partner. *J Mol Biol* 1998;282:369-382.
25. Fulop V, Moir JW, Ferguson SJ, Hajdu J. The anatomy of a bifunctional enzyme: Structural basis for reduction of oxygen to water and synthesis of nitric oxide by cytochrome *cd1*. *Cell* 1995;81:369-377.
26. Cheesman MR, Ferguson SJ, Moir JW, Richardson DJ, Zumft WG, Thomson AJ. Two enzymes with a common function but different heme ligands in the forms as isolated. Optical and magnetic properties of the heme groups in the oxidized forms of nitrite reductase, cytochrome *cd1*, from *Pseudomonas stutzeri* and *Thiosphaera pantotropha*. *Biochemistry* 1997;36:16267-16276.
27. Nurizzo D, Cutruzzola F, Arese M, Bourgeois D, Brunori M, Cambillau C, Tegoni M. Conformational changes occurring upon reduction and NO binding in nitrite reductase from *Pseudomonas aeruginosa*. *Biochemistry* 1998;37:13987-13996.
28. Nurizzo D, Cutruzzola F, Arese M, Bourgeois D, Brunori M, Cambillau C, Tegoni M. Does reduction of c heme trigger the conformational change of crystalline nitrite reductase? *J Biol Chem* 1999;274:14997-15004.

29. Wilson EK, Bellelli A, Liberti S, Arese M, Grasso S, Cutruzzola F, Brunori M, Brzezinski P. Internal electron transfer and structural dynamics of *cd1* nitrite reductase revealed by laser CO photodissociation. *Biochemistry* 1999;38:7556-7564.
30. Nurizzo D, Silvestrini MC, Mathieu M, Cutruzzola F, Bourgeois D, Fulop V, Hajdu J, Brunori M, Tegoni M, Cambillau C. N-terminal arm exchange is observed in the 2.15 Å crystal structure of oxidized nitrite reductase from *Pseudomonas aeruginosa*. *Structure* 1997;5:1157-11
31. Glockner AB, Jongst A, Zumft WG. Copper-containing nitrite reductase from *Pseudomonas aureofaciens* is functional in a mutationally cytochrome *cd1*-free background (NirS-) of *Pseudomonas stutzeri*. *Arch Microbiol* 1993;160:18-26.
32. Kastrau DHW, Heiss B, Kroneck PMH, Zumft WG. Nitric oxide reductase from *Pseudomonas stutzeri*, a novel cytochrome *bc* complex--Phospholipid requirement, electron paramagnetic resonance and redox properties. *Eur J Biochem* 1994;222:293-303.
33. Sakurai N, Sakurai T. Isolation and characterization of nitric oxide reductase from *Paracoccus halodenitrificans*. *Biochemistry* 1997;36:13809-13815.
34. Girsch P, De Vries S. Purification and initial kinetic and spectroscopic characterization of NO reductase from *Paracoccus denitrificans*. *Biochim Biophys Acta Bio-Energetics* 1997;1318:202-216.
35. Sakurai T, Sakurai N, Matsumoto H, Hirota S, Yamauchi O. Roles of four iron centers in *Paracoccus halodenitrificans* nitric oxide reductase. *Biochem Biophys Res Commun* 1998; 251:248-251.
36. Cheesman MR, Zumft WG, Thomson AJ. The MCD and EPR of the heme centers of nitric oxide reductase from *Pseudomonas stutzeri*: Evidence that the enzyme is structurally related to the heme-copper oxidases. *Biochemistry* 1998;37:3994-4000.
37. Hendriks J, Warne A, Gohlke U, Haltia T, Ludovici C, Luebben M, Saraste M. The active site of the bacterial nitric oxide reductase is a dinuclear iron center. *Biochemistry* 1998;37:13102-13109.

38. Hollocher TC, Hibbs JB, Jr. Enzymes of bacteria, plants and fungi that process free nitrogen oxides. In: Feelisch M, Stamler JS, eds. *Methods in nitric oxide research*. Chichester: John Wiley and Sons, Ltd. 1996:119-140.
39. Coyle CL, Zumft WG, Kroneck PM, Korner H, Jakob W. Nitrous oxide reductase from denitrifying *Pseudomonas perfectomarina*: purification and properties of a novel multicopper enzyme. *Eur J Biochem* 1985;153:459-467.
40. Berks BC, Baratta D, Richardson DJ, Ferguson SJ. Purification and characterization of a nitrous oxide reductase from *Thiosphaera pantotropha*--Implications for the mechanism of aerobic nitrous oxide reduction. *Eur J Biochem* 1993;212:467-476.
41. Ferretti S, Grossmann JG, Hasnain SS, Eady RR, Smith BE. Biochemical characterization and solution structure of nitrous oxide reductase from *Alcaligenes xylosoxidans* (NCIMB 11015). *Eur J Biochem* 1999;259:651-659.
42. Sato K, Okubo A, Yamazaki S. Anaerobic purification and characterization of nitrous oxide reductase from *Rhodobacter sphaeroides* f. sp. *denitrificans* IL106. *J Biochem (Tokyo)* 1999;125:864-868.
43. Soohoo CK, Hollocher TC. Purification and characterization of nitrous oxide reductase from *Pseudomonas aeruginosa* strain P2. Proceedings of the 78th Annual Meeting of the American Society of Biological Chemistry 1987;46:2241.
44. Snyder SW, Hollocher TC. Purification and some characteristics of nitrous oxide reductase from *Paracoccus denitrificans*. *J Biol Chem* 1987;262:6515-6525.
45. Hole UH, Vollack KU, Zumft WG, Eisenmann E, Siddiqui RA, Friedrich B, Kroneck PMH. Characterization of the membranous denitrification enzymes nitrite reductase (cytochrome *cd1*) and copper-containing nitrous oxide reductase from *Thiobacillus denitrificans*. *Arch Microbiol* 1996;165:55-61.
46. Antholine WE, Kastrau DHW, Steffens GCM, Buse G, Zumft WG, Kroneck PMH. A comparative EPR investigation of the multicopper proteins nitrous-oxide reductase and cytochrome c oxidase. *Eur J Biochem* 1992;209:875-881.

47. Kroneck PMH, Antholine WA, Riester J, Zumft WG. The nature of the cupric site in nitrous oxide reductase and of CuA in cytochrome *c* oxidase. *FEBS Lett* 1989;248:212-213.
48. Scott RA, Zumft WG, Coyle CL, Dooley DM. *Pseudomonas stutzeri* N₂O reductase contains CuA-type sites. *Proc Natl Acad Sci USA* 1989;86:4082-4086.
49. Kelly M, Lappalainen P, Talbo G, Haltia T, van der Oost J, Saraste M. Two cysteines, two histidines, and one methionine are ligands of a binuclear purple copper center. *J Biol Chem* 1993;268:16781-7.
50. Blackburn NJ, Barr ME, Woodruff WH, van der Oost J, de Vries S. Metal-metal bond in biology: EXAFS evidence for a 2.5 Å copper-copper bond in the Cu_A center of cytochrome oxidase. *Biochemistry* 1994;33:10401-10407.
51. Tsukihara T, Aoyama H, Yamashita E, Tomizaki T, Yamaguchi H, Shinzawa-Itoh K, Nakashima R, Yaono R, Yoshikawa S. Structures of metal sites of oxidized bovine heart cytochrome *c* oxidase at 2.8 Å. *Science* 1995;269:1069-1074.
52. Iwata S, Ostermeier C, Ludwig B, Michel H. Structure at 2.8 Å resolution of cytochrome *c* oxidase from *Paracoccus denitrificans*. *Nature* 1995;376:660-669.
53. Farrar JA, Thomson AJ, Cheesman MR, Dooley DM, Zumft WG. A model of the copper centres of nitrous oxide reductase (*Pseudomonas stutzeri*): Evidence from optical, EPR and MCD spectroscopy. *FEBS Lett* 1991;294:11-15.
54. Dooley DM, McGuirl M, Rosenzweig AC, Landin JA, Scott RA, Zumft WG, Devlin F, Stephens PJ. Spectroscopic studies of the copper sites in wild-type *Pseudomonas stutzeri* N₂O reductase and in an inactive protein isolated from a mutant deficient in copper-site biosynthesis. *Inorg Chem* 1991;30:3006-3011.
55. Dooley DM, Alvarez ML, Rosenzweig AC, Hollis RS, Zumft WG. Exogenous ligand binding to *Pseudomonas stutzeri* nitrous oxide reductase. *Inorg Chim Acta* 1996;242:239-244.

56. Sakurai N, Sakurai T. Genomic DNA cloning of the region encoding nitric oxide reductase in *Paracoccus halodenitrificans* and a structure model relevant to cytochrome oxidase. *Biochem Biophys Res Commun* 1998;243:400-406.
57. Van der Oost J, De Boer APN, De Gier JWL, Zumft WG, Stouthamer AH, Van Spanning RJM. The heme-copper oxidase family consists of three distinct types of terminal oxidases and is related to nitric oxide reductase. *FEMS Microbiol Lett* 1994;121:1-10.
58. Saraste M, Castresana J. Cytochrome oxidase evolved by tinkering with denitrification enzymes. *FEBS Lett* 1994;341:1-4.
59. McGuirl MA, Nelson LK, Bollinger JA, Chan YK, Dooley DM. The *nos* (nitrous oxide reductase) gene cluster from the soil bacterium *Achromobacter cycloclastes*: Cloning, sequence analysis, and expression. *J Inorg Biochem* 1998;70:155-169.
60. Tomlinson GA, Jahnke LL, Hochstein LI. *Halobacterium denitrificans* sp. nov., an extremely halophilic denitrifying bacterium. *Int J Syst Bacteriol* 1986;36:66-70.
61. Werber MM, Mevarech M. Induction of a dissimilatory reduction pathway of nitrate in *Halobacterium* of the Dead Sea. *Arch Biochem Biophys* 1978;186:60-65.
62. Volkl P, Huber R, Drobner E, Rachel R, Burggraf S, Trincone A, Stetter KO. *Pyrobaculum aerophilum* sp. nov., a novel nitrate-reducing hyperthermic archaeum. *Appl Environ Microbiol* 1993;59:2918-2926.
63. Hartig E, Zumft WG. Kinetics of *nirS* expression (cytochrome *cd1* nitrite reductase) in *Pseudomonas stutzeri* during the transition from aerobic respiration to denitrification: Evidence for a denitrification-specific nitrate- and nitrite-responsive regulatory system. *J Bacteriol* 1999;181:161-166.
64. Hartig E, Schiek U, Vollack KU, Zumft WG. Nitrate and nitrite control of respiratory nitrate reduction in denitrifying *Pseudomonas stutzeri* by a two-component regulatory system homologous to NarXL of *Escherichia coli*. *J Bacteriol* 1999;181:3658-3665.

65. Van Spanning RJM, De Boer APN, Reijnders WNM, Spiro S, Westerhoff HV, Stouthamer AH, Van der Oost J. Nitrite and nitric oxide reduction in *Paracoccus denitrificans* is under the control of NNR, a regulatory protein that belongs to the FNR family of transcriptional activators. *FEBS Lett* 1995;360:151-154.
66. Arai H, Kodama T, Igarashi Y. Effect of nitrogen oxides on expression of the *nir* and *nor* genes for denitrification in *Pseudomonas aeruginosa*. *FEMS Microbiol Lett* 1999;170:19-24.
67. Van Spanning RJ, De Boer AP, Reijnders WN, Westerhoff HV, Stouthamer AH, Van der Oost J. FnrP and NNR of *Paracoccus denitrificans* are both members of the FNR family of transcriptional activators but have distinct roles in respiratory adaptation in response to oxygen limitation. *Mol Microbiol* 1997;23:893-907.
68. Tosques IE, Kwiatkowski AV, Shi JR, Shapleigh JP. Characterization and regulation of the gene encoding nitrite reductase in *Rhodobacter sphaeroides* 2.4.3. *J Bacteriol* 1997;179:1090-1095.
69. Kwiatkowski AV, Shapleigh JP. Requirement of nitric oxide for induction of genes whose products are involved in nitric oxide metabolism in *Rhodobacter sphaeroides* 2.4.3. *J Biol Chem* 1996;271:24382-24388.
70. Arai H, Kodama T, Igarashi Y. Cascade regulation of the two CRP/FNR-related transcriptional regulators (ANR and DNR) and the denitrification enzymes in *Pseudomonas aeruginosa*. *Mol Microbiol* 1997;25:1141-1148.
71. Tosques IE, Shi J, Shapleigh JP. Cloning and characterization of *nnrR*, whose product is required for the expression of proteins involved in nitric oxide metabolism in *Rhodobacter sphaeroides* 2.4.3. *J Bacteriol* 1996;178:4958-4964.
72. Arai H, Igarashi Y, Kodama T. Expression of the *nir* and *nor* genes for denitrification of *Pseudomonas aeruginosa* requires a novel CRP/FNR-related transcriptional regulator, DNR, in addition to ANR. *FEBS Lett* 1995;371:73-76.

73. Van Spanning RJ, Houben E, Reijnders WN, Spiro S, Westerhoff HV, Saunders N. Nitric oxide is a signal for NNR-mediated transcription activation in *Paracoccus denitrificans*. *J Bacteriol* 1999;181:4129-4132.
74. Vollack KU, Hartig E, Koerner H, Zumft WG. Multiple transcription factors of the FNR family in denitrifying *Pseudomonas stutzeri*: characterization of four *fnr*-like genes, regulatory responses and cognate metabolic processes. *Mol Microbiol* 1999;31:1681-1694.
75. Ye RW, Haas D, Ka J, Krishnapillai V, Zimmermann A, Baird C, Tiedje JM. Anaerobic activation of the entire denitrification pathway in *Pseudomonas aeruginosa* requires *anr*, an analog of *fnr*. *J Bacteriol* 1995;177:3606-3609.
76. Cuypers H, Zumft WG. Anaerobic control of denitrification in *Pseudomonas stutzeri* escapes mutagenesis of an *fnr*-like gene. *J Bacteriol* 1993;175:7236-7246.
77. Cuypers H, Berghoefer J, Zumft WG. Multiple *nosZ* promoters and anaerobic expression of *nos* genes necessary for *Pseudomonas stutzeri* nitrous oxide reductase and assembly of its copper centers. *Biochim Biophys Acta* 1995;1264:183-190.
78. Spiro S. The FNR family of transcription regulators. *Antonie Van Leeuwenhoek* 1994; 66:23-36.
79. Zumft WG, Dreusch A, Lochelt S, Cuypers H, Friedrich B, Schneider B. Derived amino acid sequences of the *nosZ* gene (respiratory N₂O reductase) from *Alcaligenes eutrophus*, *Pseudomonas aeruginosa* and *Pseudomonas stutzeri* reveal potential copper-binding residues - Implications for the CuA site of N₂O reductase and cytochrome *c* oxidase. *Eur J Biochem* 1992;208:31-40.
80. Hoeren FU, Berks BC, Ferguson SJ, McCarthy JE. Sequence and expression of the gene encoding the respiratory nitrous-oxide reductase from *Paracoccus denitrificans*: New and conserved structural and regulatory motifs. *Eur J Biochem* 1993;218:49-57.
81. Holloway P, McCormick W, Watson RJ, Chan YK. Identification and analysis of the dissimilatory nitrous oxide reduction genes, *nosRZDFY*, of *Rhizobium meliloti*. *J Bacteriol* 1996;178:1505-1514.

82. Viebrock A, Zumft WG. Molecular cloning, heterologous expression, and primary structure of the structural gene for the copper enzyme nitrous oxide reductase from denitrifying *Pseudomonas stutzeri*. *J Bacteriol* 1988;170:4658-4667.
83. Bedmar EJ. *Bradyrhizobium japonicum* nosRZDFYLX gene cluster for nitrous-oxide reduction. 1998; GenBank Accession # AJ002531.
84. Sabaty M, Schwinter C, Richaud P, Vermeglio A. *Rhodobacter sphaeroides* f. sp. *denitrificans* nitrous oxide reductase operon, partial sequence. 1999; Genbank Accession #AF125260.
85. Philoppot L, Cheneby D, Hartmann A, Germon JC. *Pseudomonas fluorescens* regulatory protein NosR (*nosR*) and nitrous oxide reductase catalytic subunit (*nosZ*) genes, partial cds. 1998; Genbank Accession #AF056319.
86. Scala DJ, Kerkhof LJ. Nitrous oxide reductase (*nosZ*) gene-specific PCR primers for detection of denitrifiers and three *nosZ* genes from marine sediments. *FEMS Microbiol Lett* 1998;162:61-68.
87. Scala DJ, Kerkhof LJ. Diversity of nitrous oxide reductase (*nosZ*) genes in continental shelf sediments. *Appl Environ Microbiol* 1999;65:1681-1687.
88. Dooley DM, Moog RS, Zumft WG. Characterization of the copper sites in *Pseudomonas perfectomarina* nitrous oxide reductase by resonance Raman spectroscopy. *J Am Chem Soc* 1987;109:6730-6735.
89. Riester J, Zumft WG, Kroneck PMH. Nitrous oxide reductase from *Pseudomonas stutzeri* - Redox properties and spectroscopic characterization of different forms of the multicopper enzyme. *Eur J Biochem* 1989;178:751-762.
90. Farrar JA, Zumft WG, Thomson AJ. CuA and CuZ are variants of the electron transfer center in nitrous oxide reductase. *Proc Natl Acad Sci USA* 1998;95:9891-9896.

91. Chan YK, McCormick WA, Watson RJ. A new *nos* gene downstream from *nosDFY* is essential for dissimilatory reduction of nitrous oxide by *Rhizobium* (*Sinorhizobium*) *meliloti*. *Microbiology* 1997;143:2817-2824.
92. Cuypers H, Viebrock-Sambale A, Zumft WG. NosR, a membrane-bound regulatory component necessary for expression of nitrous oxide reductase in *denitrifying Pseudomonas stutzeri*. *J Bacteriol* 1992;174:5332-5339.
93. Glockner AB, Zumft WG. Sequence analysis of an internal 9.72-kb segment from the 30-kb denitrification gene cluster of *Pseudomonas stutzeri*. *Biochim Biophys Acta Bio-Energetics* 1996;1277:6-12.
94. Reijnders WNM. *Paracoccus denitrificans nirI* and *nirX* genes, and flanking regions. 1998: Genbank Accession #AJ001308.
95. McGuirl MA. Unpublished TBlastN search results. 1999.
96. Zumft WG, Viebrock-Sambale A, Braun C. Nitrous oxide reductase from denitrifying *Pseudomonas stutzeri*-Genes for copper-processing and properties of the deduced products, including a new member of the family of ATP/GTP-binding proteins. *Eur J Biochem* 1990;192:591-599.
97. Hausinger RP, Eichhorn GL. Mechanisms of Metallocenter Assembly. New York: VCH Publishers, 1996.
98. Dreusch A, Riester J, Kroneck PMH, Zumft WG. Mutation of the conserved Cys165 outside of the CuA domain destabilizes nitrous oxide reductase but maintains its catalytic activity – Evidence for disulfide bridges and a putative protein disulfide isomerase gene. *Eur J Biochem* 1996;237:447-453.
99. Fath MJ, Kolter R. ABC transporters: Bacterial exporters. *Microbiol Rev* 1993;57:995-1017.
100. Auld DS. Methods for metal substitution. *Methods Enzymol* 1988;158:71-79.

101. Dooley DM, Brown DE, Clague AW, Kemsley JN, McCahon CD, McGuirl MA, Turowski PN, McIntire WS, Farrar JA, Thomson AJ. Structure and reactivity of copper-containing amine oxidases. In: Karlin KD, Tyeklar Z, eds. *Bioinorganic Chemistry of Copper*. New York: Chapman & Hall, 1993:459-470.
102. Rae TD, Schmidt PJ, Pufahl RA, Culotta VC, O'Halloran TV. Undetectable intracellular free copper: The requirement of a copper chaperone for superoxide dismutase. *Science* 1999;284:805-808.
103. Palmiter RD. The elusive function of metallothioneins. *Proc Natl Acad Sci USA* 1998;95:8428-8430.
104. Huckle JW, Morby AP, Turner JS, Robinson NJ. Isolation of a prokaryotic metallothionein locus and analysis of transcriptional control by trace metal ions. *Mol Microbiol* 1993;7:177-187.
105. Lamb AL, Wernimont AK, Pufahl RA, Culotta VC, O'Halloran TV, Rosenzweig AC. Crystal structure of the copper chaperone for superoxide dismutase. *Nature Struct Biol* 1999;6:724-729.
106. Rosenzweig AC, Huffman DL, Hou MY, Wernimont AK, Pufahl RA, O'Halloran TV. Crystal structure of the Atx1 metallochaperone protein at 1.02 Å resolution. *Struct Fold Des* 1999;7:605-617.
107. Pufahl RA, Singer CP, Peariso KL, Lin SJ, Schmidt PJ, Fahrni CJ, Culotta VC, Penner-Hahn JE, O'Halloran TV. Metal ion chaperone function of the soluble Cu(I) receptor Atx1. *Science* 1997;278:853-856.
108. Moncrief MBC, Hausinger RP. Nickel incorporation into urease. In: Hausinger RP, Eichhorn GL, eds. *Mechanisms of Metallocenter Assembly*. New York: VCH Publishers, 1996:151-171.
109. Brayman TG, Hausinger RP. Purification, characterization, and functional analysis of a truncated *Klebsiella aerogenes* UreE urease accessory protein lacking the histidine-rich carboxyl terminus. *J Bacteriol* 1996;178:5410-5416.

110. Sankaran K, Wu HC. Bacterial lipoprotein signal peptidase. *Methods Enzymol* 1995;248:169-180.
111. Braun V, Wu HC. Lipoproteins, structure, function, biosynthesis and model for protein export. Ghuysen J-M, Hakenbeck R., eds. In: *Bacterial Cell Wall*. Amsterdam: Elsevier Science B V, 1994:319-339.
112. Adman ET. Copper protein structures. *Adv Protein Chem* 1991;42:145-97.
113. Pierloot K, De Kerpel JOA, Ryde U, Olsson MHM, Roos BO. relation between the structure and spectroscopic properties of blue copper proteins. *J Am Chem Soc* 1998;120:13156-13166.
114. Sambrook J, Fritsch EF, Maniatis T. *Molecular Cloning*. New York: Cold Spring Harbor Laboratory Press, 1989.
115. Sanger F, Nicklen S, Coulson AR. DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci USA* 1977;74:5463-5467.
116. Roe BA, Lin SP, Song L, Yuan X, Clifton S, Dyer DW. *Gonococcal Genome Sequencing Project*, the University of Oklahoma. 1998.
117. Hofmann K, Stoffel W. TMbase - A database of membrane spanning proteins segments. *Biol. Chem. Hoppe-Seyler* 1993;347:166.
118. Dreusch A, Burgisser DM, Heizmann CW, Zumft WG. Lack of copper insertion into unprocessed cytoplasmic nitrous oxide reductase generated by an R20D substitution in the arginine consensus motif of the signal peptide. *Biochim Biophys Acta* 1997;1319:311-318.
119. von Heijne G. Computer analysis of DNA and protein sequences. *Eur J Biochem* 1991;199:253-256.

120. Chan Y-K, Wheatcroft R. Detection of a nitrous oxide reductase structural gene in *Rhizobium meliloti* strains and its location on the *nod* megaplasmid of JJ1c10 and SU47. *J Bacteriol* 1993;175:19-26
121. Teidje JM, *personal communication*.
122. Dooley DM, McDermott T. *unpublished observations*.
123. Meile L, Stettler R, Banholzer R, Kotik M, Leisinger T. Tryptophan gene cluster of *Methanobacterium thermoautotrophicum* Marburg: molecular cloning and nucleotide sequence of a putative *trpEGCFBAD* operon. *J Bacteriol* 1991;173:5017-5023.
124. Weiner JH, Bilous PT, Shaw GM, Lubitz SP, Frost L, Thomas GH, Cole JA, Turner RJ. A novel and ubiquitous system for membrane targeting and secretion of cofactor-containing proteins. *Cell* 1998;93:93-101.
125. Santini CL, Ize B, Chanal A, Muller M, Giordano G, Wu LF. A novel sec-independent periplasmic protein translocation pathway in *Escherichia coli*. *EMBO J* 1998;17:101-112.
126. Andrew CR, Han J, de Vries S, van der Oost J, Averill BA, Loehr TM, Sanders-Loehr J. CuA of cytochrome c oxidase and a site of N₂O reductase are tetrahedrally distorted type 1 Cu-cysteineates. *J Am Chem Soc* 1994;116:10805-10806.
127. Hulse CL, Averill BA. Isolation of a high specific activity pink, monomeric nitrous oxide reductase from *Achromobacter cycloclastes*. *Biochem Biophys Res Commun* 1990;166:729-735.
128. Moulis JM, Sieker LC, Wilson KS, Dauter Z. Crystal structure of the 2[4Fe-4S] ferredoxin from *Chromatium vinosum*: evolutionary and mechanistic inferences for [3/4Fe-4S] ferredoxins. *Protein Sci* 1996;5:1765-1775.
129. Shen B, Jollie DR, Diller TC, Stout CD, Stephens PJ, Burgess BK. Site-directed mutagenesis of *Azotobacter vinelandii* ferredoxin I: cysteine ligation of the [4Fe-4S] cluster with protein rearrangement is preferred over serine ligation. *Proc Natl Acad Sci USA* 1995;92:10064-10068.

130. Farrar JA, Zumft WG, Thomson AJ. CuA and CuZ are variants of the electron transfer center in nitrous oxide reductase. *Proc Natl Acad Sci USA* 1998;95:9891-9896.
131. Schmidt PJ, Rae TD, Pufahl RA, Hamma T, Strain J, O'Halloran TV, Culotta VC. Multiple protein domains contribute to the action of the copper chaperone for superoxide dismutase. *J Biol Chem* 1999;274:23719-23725.
132. Rees E, Siddiqui RA, Koster F, Schneider B, Friedrich B. Structural gene (*nirS*) for the cytochrome *cd1* nitrite reductase of *Alcaligenes eutrophus* H16. *Appl Environ Microbiol* 1997;63:800-802.
133. Labes M, Puhler A, Simon R. A new family of RSF1010-derived expression and lac-fusion broad-host-range vectors for gram-negative bacteria. *Gene* 1990;89:37-46.
134. Mayer MP, Prestwich GD, Dolence JM, Bond PD, Wu HY, Poulter CD. Protein farnesyltransferase: production in *Escherichia coli* and immunoaffinity purification of the heterodimer from *Saccharomyces cerevisiae*. *Gene* 1993;132:41-47.
135. Gennity J, Goldstein J, Inouye M. Signal peptide mutants of *Escherichia coli*. *J Bioenerg Biomembr* 1990;22:233-269.
136. Inukai M, Takeuchi M, Shinizu K, Arai M. Mechanism of action of globomycin. *J Antibiot (Tokyo)* 1978;31:1203-1205.
137. Metzger JW, Beck-Sickinger AG, Loleit M, Eckert M, Bessler WG, Jung GJ. Synthetic S-(2,3-dihydroxypropyl)-cysteinyl peptides derived from the N-terminus of the cytochrome subunit of the photoreaction centre of *Rhodospseudomonas viridis* enhance murine splenocyte proliferation. *J Pept Sci* 1995;1:184-190.
138. Hayashi S, Chang SY, Chang S, Giam CZ, Wu HC. Modification and processing of internalized signal sequences of prolipoprotein in *Escherichia coli* and in i. *J Biol Chem* 1985;260:5753-5759.

139. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. *Current Protocols in Molecular Biology* 1989.
140. Miroux B, Walker JE. Over-production of proteins in *Escherichia coli*: mutant hosts that allow synthesis of some membrane proteins and globular proteins at high levels. *J Mol Biol* 1996;260:289-298.
141. Holmquist B, Vallee BL. Tryptophan quantitation by magnetic circular dichroism in native and modified proteins. *Biochemistry* 1973;12:4409-4417.
142. Hanna PM, Tamilarasan R, McMillin DR. Cu(I) analysis of blue copper proteins. *Biochem J* 1988;256:1001-1004.
143. Morpurgo L, Savini I, Mondovi B, Avigliano L. Removal of type 2 Cu from ascorbate oxidase and laccase by reaction with N, N-diethyldithiocarbamate. *J Inorg Biochem* 1987;29:25-31.
144. Ramachandran S, Udgaonkar JB. Stabilization of Barstar by chemical modification of the buried cysteines. *Biochemistry* 1996;35:8776-8785.
145. Jocelyn PC. Spectrophotometric assay of thiols. *Meth Enzymol* 1987;143:44-67.
146. Husain M, Davidson VL, Smith AJ. Properties of *Paracoccus denitrificans* amicyanin. *Biochemistry* 1986;25:2431-2436.
147. Bollinger JM, *personal communication*.
148. Jenkinson HF. Adherence, coaggregation, and hydrophobicity of *Streptococcus gordonii* associated with expression of cell surface lipoproteins. *Infect Immun* 1992;60:1225-1228.
149. Cavard DJ. Synthesis and functioning of the colicin E1 lysis protein: comparison with the colicin A lysis protein. *J Bacteriol* 1991;173:191-196.

150. Hunter CP. Genetics: a touch of elegance with RNAi. *Curr Biol* 1999;9:R440-R442.
151. Kobayashi T, Ito K. Respiratory chain strongly oxidizes the CXXC motif of DsbB in the *Escherichia coli* disulfide bond formation pathway. *EMBO J* 1999;18:1192-1198.
152. Dooley DM, Henery SH. *unpublished observations*.
153. Liu MY, Liu MC, Payne WJ, LeGall J. Properties and electron transfer specificity of copper proteins from the denitrifier *Achromobacter cycloclastes*. *J Bacteriol* 1986;166:604-608.

APPENDICES

APPENDIX A

Comparison of 50 Amino Acid Sequences of N₂OR

Amino Acid Lineups of all available NosZ (N₂OR) sequences. * indicates a sequence that may have a frame shift in the italicized residues (D Scala, personal communication). The Consensus sequence has been calculated without including the italicized residues. Sequences are identified by their Genbank Accession Numbers (see list below); accession numbers in bold represent complete sequences.

Genbank #	Organism	Genbank #	Organism
AF119953	Uncultured marine bacterium	AF119925	Uncultured marine bacterium
*AF016059	Uncultured marine bacterium	AF119948	Uncultured marine bacterium
X65277	<i>Pseudomonas aeruginosa</i>	AJ002531	<i>Bradyrhizobium japonicum</i>
M22628	<i>Pseudomonas stutzeri</i>	U47133	<i>Sinorhizobium meliloti</i>
AF056319	<i>Pseudomonas fluorescens</i>	AF047429	<i>Achromobacter cycloclastes</i>
AF119942	Uncultured marine bacterium	X74792	<i>Paracoccus denitrificans</i>
AF119940	Uncultured marine bacterium	AF125260	<i>Rhodobacter sphaeroides</i>
AF119947	Uncultured marine bacterium	AF016055	Uncultured marine bacterium
AF119944	Uncultured marine bacterium	AF016056	Uncultured marine bacterium
AF119952	Uncultured marine bacterium	AF119927	Uncultured marine bacterium
AF119941	Uncultured marine bacterium	AF119926	Uncultured marine bacterium
AF119936	Uncultured marine bacterium	AF119922	Uncultured marine bacterium
AF119949	Uncultured marine bacterium	AF119921	Uncultured marine bacterium
AF016057	Uncultured marine bacterium	AF119920	Uncultured marine bacterium
AF119945	Uncultured marine bacterium	AF119933	Uncultured marine bacterium
AF119938	Uncultured marine bacterium	AF119930	Uncultured marine bacterium
AF119943	Uncultured marine bacterium	AF119928	Uncultured marine bacterium
AF119950	Uncultured marine bacterium	AF119932	Uncultured marine bacterium
AF119954	Uncultured marine bacterium	AF119931	Uncultured marine bacterium
AF119939	Uncultured marine bacterium	AF119934	Uncultured marine bacterium
AF119935	Uncultured marine bacterium	AF119929	Uncultured marine bacterium
AF119918	Uncultured marine bacterium	AF119919	Uncultured marine bacterium
AF119951	Uncultured marine bacterium	AF119924	Uncultured marine bacterium
AF119946	Uncultured marine bacterium	AF119923	Uncultured marine bacterium
AF119937	Uncultured marine bacterium	X65278	<i>Ralstonia eutropha</i>

1		50
AF119953		
AF016059		
X65277	MSDK QTDKD....E KTGLSRRGFL GASALTRSAV A.ASGLVGGV	
M22628	MSDK DSKNTPQVPE KLGLSRRGFL GASAVTGAAV A.ATALGGAV	
AF056319	MSDK KDQVPGAVEA PRGVSRRSFL GTGAVTGAVL AGATALGAGT	
AF119942		
AF119940		
AF119947		
AF119944		
AF119952		
AF119941		

AF119936					
AF119949					
AF016057					
AF119945					
AF119938					
AF119943					
AF119950					
AF119954					
AF119939					
AF119935					
AF119918					
AF119951					
AF119946					
AF119937					
AF119925					
AF119948					
AJ002531	MSDSDNIKGV	SRRTLLGTTA	..AAAGVGLA	GGAVVTKDGA	GFVSTADAQT
U47133	MSNEETKMRL	NRQMLGTTA	FMAAAGAVGA	GGALTLGGT	A..TPARAQE
AF047429	MESKE	HKGLSRRALF	SATAGSAILA	G.....TVG	PAALSLGAAG
X774792	MESKQ	EKGLSRRALL	GATAGGAAVA	GAFGGRLALG	PAALGLGTAG
AF125260					
AF016055					
AF016056					
AF119927					
AF119926					
AF119922					
AF119921					
AF119920					
AF119933					
AF119930					
AF119928					
AF119932					
AF119931					
AF119934					
AF119929					
AF119919					
AF119924					
AF119923					
X65278	..MSKEKASI	GNGPGGIGRR	QFLGTAALAG	LAGVVACTDK	GAAPAAAAG
Consensus	-----	-----	-----	-----	-----

51

100

AF119953					
AF016059					
X65277	MTRDSWAAAA	KEAQQKIHVA	PGELDEYYGF	WSGGHQGEVR	VLGVPSMREL
M22628	MTRESWAQAV	KESKQKIHVG	PGELDDYYGF	WSGGHQGEVR	VLGVPSMREL
AF056319	FTRESWAAAA	KEAKSKIHVG	PGELDEYYGF	WSGGHQGEVR	VLGVPSMREL
AF119942					
AF119940					
AF119947					
AF119944					
AF119952					
AF119941					
AF119936					

AF119949					
AF016057					
AF119945					
AF119938					
AF119943					
AF119950					
AF119954					
AF119939					
AF119935					
AF119918					
AF119951					
AF119946					
AF119937					
AF119925					
AF119948					
AJ002531	KSAAPKAPPA	RPAVQKTEVA	PGELDEYYVF	FSSGQSGEMR	IIGLPSMREL
U47133	TSGSS.....	YEVK	PGELDEYYVF	FSSGQSGEIR	IVGAPSMREM
AF047429	LATPARA...	..ATGADGSVA	PGKLDDYYGF	WSSGQTGEMR	ILGIPSMREL
X774792	VATVAGSGAA	LAASGDGSVA	PGQLDDYYGF	WSSGQSGEMR	ILGIPSMREL
AF125260			PGELDEYYGF	WSSGQTGELR	ILGIPSMREL
AF016055					
AF016056					
AF119927					
AF119926					
AF119922					
AF119921					
AF119920					
AF119933					
AF119930					
AF119928					
AF119932					
AF119931					
AF119934					
AF119929					
AF119919					
AF119924					
AF119923					
X65278	APASGAHGAA	HGAGASVHLK	PGELDTYYGL	WSGGHTGDMR	VLGLPSGREI
Consensus	-----	-----	PG-LD-YY--	--S-G--G--R	--G-PS-RE-

101

150

AF119953		..GWGLTNES	KRVLGDS...		..AHFQNGDC
AF016059		..GWGLTNES	KRVSADS...		..AHFQNGDC
X65277	MRIPVFNVD	ATGWGLTNES	RHILGDT...		..AKFLNGDC
M22628	MRIPVFNVD	ATGWGLTNES	RHIMGDS...		..AKFLNGDC
AF056319	MRIPVFNVD	ATGWGLTNES	KRILGDS...		..AKYQNGDC
AF119942		..GWGLTNES	VKIITENMLP	RETEFWK..N	KGGFPLNGDL
AF119940		..GWGLTNES	VKILTENMLP	RETEYWK..N	KGGFPLNGDL
AF119947		..GWGLTNES	VKILTENMLP	RETEYWK..N	KGGFPLNGDL
AF119944		..GWGLTNES	VKILTENMLP	RETEYWK..N	KGGFPLNGDL
AF119952		..GWGLTNES	LKILTENMLP	REKEFWK..N	KGGFPLNGDL
AF119941		..GWGLTNES	IKILTENMLP	REKEYWQ..D	KGGFPLNGDL
AF119936		..GWGLTNES	IKTLTENMLP	REKEFWK..N	KGGFPLNGDL

AF119949		..GWGLTNES	LKVLTDGLLP	ETKKYLE..S	RGGTWQNGDL
AF016057		..GWGLTNES	REVLTKGLLP	ETQKYLE..T	RGGTWQNGDL
AF119945		..GWGLTNES	RKILTEGLLP	DTKKFLE..S	RGGIWENGDL
AF119938		..GWGLTNES	RKILTEGLLP	ETRKFLE..S	RGGIYENGDL
AF119943		..GWGLTNES	RKVLRDGLLP	ETRKFLE..S	RGGIWENGDL
AF119950		..GWGLTNES	RKVLRDGLLP	ETRKYLE..S	RGGIWENGDL
AF119954		..GWGLTNES	RKVLSEGLLP	ETTKYLE..S	RGGIWENGDL
AF119939		..GWGLTNES	RKILSDGLLP	ETKQFLE..S	RGGIWENGDL
AF119935		..GWGLTNES	RKILSEGLLP	ETKKFLE..S	RGGVWTNGDL
AF119918		..GWGLTNES	RKILRDGLLP	ETKEFLDTGG	RGGVWTNGDL
AF119951		..GWGLTNES	LKILTEGLLP	ETKEFLTTRG	.GT.YMNGDA
AF119946		..GWGLTNES	LKILTDGLLP	ETKEFLSTRG	.GT.YMNGDA
AF119937		..GWGLTNES	RKILTEGLLP	ETREYLAAG	.GV.WHNGDA
AF119925		..GWGLTNES	LKILTENLLP	ETKAFLAASG	.QKNYENGDL
AF119948		..GWGLTNES	RKIMTDGMLP	ETKEYLK..D	KGGVMLNGDA
AJ002531	MRVPVFNRC	ATGWGQTNES	LKVLTEGMQP	ATREFLK..N	RGGTFMNGDL
U47133	MRIPVFNRC	ATGWGQTNES	RKVMTEGLLP	ETVEFLK..D	QGGLYLNGLD
AF047429	MRVPVFNRC	ATGWGQTNES	IRIHQRTMTE	KTKKQLAANG	KKI.HDNGDL
X774792	MRVPVFNRC	ATGWGQTNES	VRIHERTMSE	RTKKFLAANG	KRI.HDNGDL
AF125260	MRVPVFNRC	ATGWGQTNES	LRIMTENMSD	KTRAFLOANG	KKV.HDNGDL
AF016055		..GWGLTNES	LQIPTEGLLP	ETKELLANRG	MKT.YDNGDL
AF016056		..GWGLTNES	LQILTEGLLP	ETKEMLANRG	MKT.YDNGDL
AF119927		..GWGLTNES	LKVLTEGLLP	ETKAYLADQG	KVT.YDHGDL
AF119926		..GWGLTNES	LKVLTEGLLP	ETKAYLADQG	KVT.YDHGDL
AF119922		..GWGLTNES	LKILTEGLTE	KTKDFLAARN	MKT.CDNGDL
AF119921		..GWGLTNES	LKILTEGLTE	KTKNFLAARN	MKT.YDNGDL
AF119920		..GWGLTNES	LKILTEGLTE	KTKDFLAARN	MKT.YDNGDL
AF119933		..GWGLTNES	LKVLTEGLTE	ETKAFLAAG	QKT.YDNGDL
AF119930		..GWGLTNES	KKILTEGLLP	KTKEYLAANG	MET.YDNGDL
AF119928		..GWGLTNES	LKVMTEGLLP	ETKEFLAANG	KVT.YDNGDL
AF119932		..GWGLTNES	LKVLTEGLTD	ETKAYLASNG	KVT.YDNGDL
AF119931		..GWGLTNES	LKVLTEGLTD	ETKAFLAANG	KVT.YDNGDL
AF119934		..GWGLTNES	LKVLTEGLTD	ETKAFLAANG	KVT.YDNGDL
AF119929		..GWGLTNES	LKVLTEGLTQ	ETKDFLAANG	KVT.YDNGDL
AF119919		..GWGLTNES	LKVLTEGLTD	ETKKFLAANG	KVT.YDNGDL
AF119924		..GWGLTNES	PKVLTEGLTD	ETKKFLAANG	KVT.YDNGDL
AF119923		..GWGLTNES	LKILTEGLLP	ETKQFLAANG	KVT.YDNGDL
X65278	LRIPCFVPA	LVGWGITNES	KKVMG.....	ARPD	GTLRYTVADT
Consensus	-R-P-F----	--GWG-TN-S	-----	-----	-----D-

151

200

AF119953	HHPHISMTDG	KYDGKYLFD	ELANSRVARI	RLDIMKCDKI	TTIPNVQAIH
AF016059	HHPHISMTDG	KYDGKYLFIN	DKANSRVARI	RLDIMKCDKI	TTIPNVQAIH
X65277	HHPHISMTDG	KYDGKYLFIN	DKANSRVARI	RLDIMKCDKI	TTIPNVQAIH
M22628	HHPHISMTDG	KYDGKYLFIN	DKANSRVARI	RLDIMKCDKM	ITVPNVQAIH
AF056319	HHPHISMTDG	KYDGKYLFIN	DKANSRVARI	RLDIMKCDKI	LTVPNVQAIH
AF119942	HHPHMSFTDG	TYDGRYLFAN	DKANTRVARI	RCDIMKTDKI	ISIPNASDIH
AF119940	HHPHMSFTDG	TYDGRYLFAN	DKANTRVARI	RCDIMKTDKI	ISIPNASDIH
AF119947	HHPHMSFTDG	TYDGRYLFAN	DKANTRVARI	RCDIMKTDKI	ISIPNASDIH
AF119944	HHPHMSFTDG	TYDGRYLFAN	DKANTRVARI	RCDIMKTDKI	ISIPNASDIH
AF119952	HHPHMSFTDG	TYGGRYLFAN	DKANTRVARI	RCDIMKTDKI	ISIPNASDIH
AF119941	HHPHMSFTKG	TYDGRYIFAN	DKANTRVARV	RCDIMKTDKI	ISIPNASDIH
AF119936	HHPHMSFTDG	TYDGRYVFAN	DKANTRVARI	RCDIMKTDKI	ISIPNASDIH

AF119949 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RVDIMKCDKI IEIPNASDTH
 AF016057 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNASDIH
 AF119945 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119938 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119943 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119950 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119954 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119939 HHPHMSFTDG TCDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119935 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDIH
 AF119918 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDIMKCDKI IEIPNASDIH
 AF119951 HHPHMSFTDG TYDGRYLYIN DKANTRVARI RCDVMKVDKI IEIPNASDIH
 AF119946 HHPHMSFTDG TYDGRYLYIN DKANTRVARI RCDVMKVDKV IEIPNASDIH
 AF119937 HHPHMSFTDG TYDGRYLFIN DKANTRVARI RCDVMKCDKI VEIPNASDIH
 AF119925 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKV IEIPNASDIH
 AF119948 HHPHMSFTDG TYDGRYVYIN DKADTRVARI RCDVMKVDKM VEIPNAADII
 AJ002531 HHPHVSFTDG TYDGRYAFMN DKANSRVARV RLDVMKCDKI IELPNQHTVH
 U47133 HHPHPSFTDG TYDGRYLYAN DKSNSRVCRI RLDVMKCDKI IQLPNQHTVH
 AF047429 HHVHMSFTDG KYDGRYLFMN DKANTRVARV RCDVMKTDAL LEIPNAKGII
 X774792 HHVHMSFTEG KYDGRFLFMN DKANTRVARV RCDVMKCDAL LEIPNAKGII
 AF125260 HHVHMSFTEG KYDGRFLFMN DKANTRVARV RCDVMKTDAL LEIPNAKGII
 AF016055 HHPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKCDKI IEIPNAQDII
 AF016056 HHPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKCDKI IEIPNAQDII
 AF119927 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119926 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKRDKI IEVPNAHDIH
 AF119922 HHPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKCDKI IEIPNAHDIH
 AF119921 HHPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKCDKI IEIPNAHDIH
 AF119920 HHPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKCDKI IEIPNAHDIH
 AF119933 HHPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVTKCDKI IEIPNAHDIH
 AF119930 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDIMKCDKI IEIPNAHDIH
 AF119928 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119932 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119931 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119934 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119929 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119919 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119924 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 AF119923 HHPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
 X65278 HHTHASYKDG NYDGRYAWVN DKINSRIARI RLDYFICDKI TELPNVQGFH
 Consensus HH-H-S---G ---G----- -----R--R- R-D----D-- ---PN----H

201

250

AF119953 GLRLQ..... ..KVPHSKYV LCNAEFIIPH PNDGSSFDIT GDNAFTMYNA
 AF016059 GLRLQ..... ..KVPHTKYV LCNAEFIIPH PNDGSSFDIT GDNAFTMYNA
 X65277 GLRLQ..... ..KVPHTKYV FCNAEFIIPH PNDGKVFLE DENSFTMYNA
 M22628 GLRLQ..... ..KVPHTKYV FANAEFIIPH PNDGKVFLE DENSFTMYNA
 AF056319 GLRLQ..... ..KVPYTKYV FANAEFVIPH PNDGHTFDLQ DKNSFTMFNA
 AF119942 GLRPQ..... ..KFPRTGYI FANGEHRIPI PNDGSILD.D PSKYHAIFTA
 AF119940 GLRPQ..... ..KFPRTGYI FANGEHRIPI PNDGSILD.D PSKYHAIFTA
 AF119947 GLRPQ..... ..KFPRTGYI FANGEHCIPI PNDGSILD.D PSKYHAIFTA
 AF119944 GLRPQ..... ..KFPRTGYI FANGEHRIPI PNDGSILD.D PSKYHAIFTA
 AF119952 GLRPQ..... ..KFPRTGYV FANGEHRIPL PNDGMILD.E PDKYHAILTA
 AF119941 GLRPQ..... ..KFPRTGYV FANGEHRIPL PNDGTILD.D PSKYHAIFTA
 AF119936 GLRPQ..... ..KFPRTGYV FANGEHRIPL PNDGSVLD.D PDKYHAIFTA

AF119949	GLRPQ.....	..KYPRTGYV	FANGEHRVPI	PNDGSILD.E	PEKYHAIFTA
AF016057	GLRPQ.....	..KYPRTGYV	FANGEHRVPI	PNDGSILD.E	PEKYHAIFTA
AF119945	GLRPQ.....	..KYPRTGYV	FANGEHRVPL	PNDGSVLD.E	PEKYHAIFTA
AF119938	GLRPQ.....	..KYPRTGYI	FANGEHRVPL	PNDGSVLD.E	PEKYHAIFTA
AF119943	GPRPQ.....	..KYPRTGYV	FANGEHRVPV	PNDGSVLD.E	PEKYHAIFTA
AF119950	GLRPQ.....	..KYPRTGYI	FANGEHRVPV	PNDGSVLD.E	PEKYRAIFTA
AF119954	GLRPQ.....	..KYPRTGYV	FANGEHRVPV	PNDGSILD.E	PEKYHAISTA
AF119939	GLRPQ.....	..KYPRTGYI	FANGEHRVPL	PNDGRDLD.E	PTKYHAIFTA
AF119935	GLRPQ.....	..KYPRTGYV	FANGEHRVPV	PNDGSILD.E	PDKYHAIFTA
AF119918	GLRPQ.....	..KYPRTGYV	FANGEHRVPI	PNDGSILD.E	PEEYHAIFTA
AF119951	GLRPQ.....	..KFPKTGYI	FANGEHRVPL	PNDGMILD.E	PEKYHAVFTA
AF119946	GLRPQ.....	..KFPKTGYI	FANSEHRIPM	PNDGMILD.E	PEKYHAVFSA
AF119937	GMRPQ.....	..KFPRTGYI	FANGEHRIPI	PNDGMILD.E	PEKYHSVFTA
AF119925	GLRPQ.....	..KFPRTGYV	FANGEHRIPL	PNDGMILD.D	PEKYVAIFTA
AF119948	GLRPQ.....	..KYPKTGYI	FANGEHRIPI	PNDGKILD.E	PGKYNAIFSA
AJ002531	GLRLQ.....	..KYPRTGYV	FCNGEDGVPL	PNDGKVLD.N	PKEYHSIFTA
U47133	GLRVQ.....	..KYPKTGYV	FCNGEDAVPV	PNDGKTMG.D	KNSYQAIFTA
AF047429	GMRPQ.....	..KWPRSNYV	FCNGEDEAPL	VNDGSTMT.D	VATYVNIFTA
X774792	GLRPQ.....	..KWPRSNYV	FCNGEDÉTPL	VNDGTNME.D	VANYVNVFTA
AF125260	GLRPQ.....	..KWPRNTNYV	FANGEDEAPM	VNDGKVLD.D	PSQYVNFETA
AF016055	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGSVLD.K	PDEYVNIFTA
AF016056	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGSVLD.K	PDEYVNTFTA
AF119927	GLRPQ.....	..KFPRTGYV	FANSEHEAPL	INDGSILD.E	PDKYINVFTA
AF119926	GLRPQ.....	..KVPRTGYV	FANSEHEAPL	INDGTILD.D	PDKYINIFTA
AF119922	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGDILD.D	PSKYVNIFTA
AF119921	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGDILD.D	PSKYVNIFTA
AF119920	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGDILD.N	PSKYVNIFTA
AF119933	GMRPQ.....	..KFPRTGYV	FANGEHEAPL	VNDGQILD.E	PDKYVNVFTA
AF119930	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGSVLD.D	PSQYVNI FSA
AF119928	GLRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGKVLD.K	PEEYVNI FSA
AF119932	GMRPQ.....	..KFPRTGYV	FANGEHEAPL	VNDGQVLD.D	PSQYVNIFTA
AF119931	GMRPQ.....	..KFPRTGYV	FANGEHEAPL	VNDGQILD.D	PSQYVNIFTA
AF119934	GMRPQ.....	..KFPRTGYV	FANGEHEAPL	VNDGQILD.D	PSQYVNIFTA
AF119929	GMRPQ.....	..KFPRTGYV	FANGEHEAPL	VNDGQILD.D	PSQYVNIFTA
AF119919	GMRPQ.....	..KFPRTGYV	FANGEHEASL	VNNGEILD.D	PSQYVNIFTA
AF119924	GMRPQ.....	..KYPRTGYV	FANGEHEAPL	VNDGQFLD.D	PSQYVNI FTA
AF119923	GLRPQ.....	..KYPRTGYI	FANGEHEAPL	VNDGSILD.K	PEEYVNI FTA
X65278	GIFPDKRPV	DTKINYTTRV	FCGGEFGIPL	PSAPT...ED	AGKYRSLFTC
Consensus	G-----	--K-----	----E-----	-----	-----

251

300

AF119953	VDAETMEVAW	QVIVDGNLDN	SDMDYTGRFA	ASTCYNSEKA	TDLAGMMRNE
AF016059	VDAETMEVAW	QVIVDGNLDN	SDMDYTGRFA	ASTCYNSEKA	TDLTGMMRNE
X65277	VDAETMEVAF	QVIVDGNLDN	TDADYTGRFT	AATCYNSEKA	FDLGGMMRNE
M22628	IDAETMEMAF	QVIVDGNLDN	TDADYTGRFA	AATCYNSEKA	FDLGGMMRNE
AF056319	IDAETMEMAF	QVIVDGNLDN	SDADYTGRFA	ASTCYNSEKA	YDLGGMMRNE
AF119942	IDGDAMTVAW	QVIVDGNLDN	VDADYQGKYA	ISSCYNSEEG	INLEQMTSSE
AF119940	IDGDAMTVAW	QVIVDGNLDN	VDADYQGKYA	ISSCYNSEEG	INLEQMTSSE
AF119947	IDGDAMTVAW	QVIVDGNLDN	VDADYQGKYA	ISSCYNSEEG	INLEQMTSSE
AF119944	IDGDAMSVAW	QVIVDGNLDN	VDADYQGKYA	ISSCYNSEEG	INLEQMTSSE
AF119952	IDGDTMKVAW	QVVVDGNLDN	VDADYQGKYA	ISSCYNSEEG	INLEQMTSSE
AF119941	IDGDSMKVAW	QVIVDGNLDN	VDADYQGKYA	VSSCYNSEEG	VTLEQMTSSE
AF119936	IDADEMVAW	QVMVDRNLDN	VDADYQGKYA	VSSCYNSEEG	VTLAEMTSSE

AF119949	VDGDAMEVAW	QVMVDGNLDN	CDADYQGKYA	ISTCYNSEEA	VNLAGMMANE
AF016057	IDGDKMEIAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEA	ANLAGMMASE
AF119945	LDGDSMKVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VTLAEMTANE
AF119938	LDGDTMKVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VTLAEMTSNE
AF119943	LDGDTMKVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VTLAEMTSSE
AF119950	LDGDSMKVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VTLAEMASSE
AF119954	LDGDSMKVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VTLTEMTSSE
AF119939	LDGDSMKVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VTLAEMTANE
AF119935	LDGDSMEVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEG	VNLAEMTASE
AF119918	IDGDEMKEVAW	QVMVDGNLDN	CDADYQGLYA	FSTCYNSEEA	VNLSGMIANE
AF119951	IDGETMLVAW	QVIVDGNLDN	CDADYQGKYA	VSTSYNSEEG	VTLAEMTANE
AF119946	IDGETMTVAW	QVIVDGNLDN	CDADYQGKYA	VSTSYNSEEG	VTLAEMTANE
AF119937	LDGDTMKVAW	QVIVDGNLDN	CDADYQGLYA	VSTCYNSEEG	VTLAEMTASE
AF119925	IDGDTMKIAW	QVIVDGNLDN	CDADYQGKYA	VSTCYNSEEG	VTLAEMTSSE
AF119948	VDGEKMTVAW	QVIVSGNLDN	VDADYQGKYA	YSTCYNSEEG	VTLAEMTKAE
AJ002531	LDGDTMKVAW	QVMVSGNLDN	VDSYQGYC	FSTCYNSEEG	VTLAEMTANE
U47133	VDGETMEVAW	QVMVDGNLDN	VDADYQGKYC	FATCYNSEEG	FTLADMMASE
AF047429	VDADKWEVAW	QVKVSGNLDN	CDADYEGKWA	FSTSYNSEMG	MTLEEMTKSE
X774792	VDADKWEVAW	QVLVSGNLDN	CDADYEGKWA	FSTSYNSEMG	MTLPENTAAE
AF125260	IDADRMKVAW	QVLVSGNLDN	TDADYEGKWA	FSTSYNSEMG	TNLGEMMAAE
AF016055	IDGDEMTVAW	QVIVSGNLDN	CDADYQGKYA	FATSYNSEMG	MNLAEMTASE
AF016056	IDGDEMTVAW	QVIVSGNLDN	CDADYQGKYA	FATSYNSEMG	MSLAGMTASE
AF119927	IDGDEVKVAW	QVMVSGNLDN	TDCDYQGKYA	YSTSYNSEMG	MNLADMMAAE
AF119926	IDGDEMKEVAW	QVMVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLADMMAAE
AF119922	IDGDEMKEIAW	QVKVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTESE
AF119921	IDGDEMKEIAW	QVKVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTESE
AF119920	IDGDEMKEIAW	QVKVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTESE
AF119933	IDGDKMEVSW	QVTVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTESE
AF119930	IDGDEMKEVAW	QVMVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTSSE
AF119928	IDGDEMKEVAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTSSE
AF119932	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTANE
AF119931	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTANE
AF119934	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTANE
AF119929	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTANE
AF119919	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTANE
AF119924	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTANE
AF119923	IDGDTMEIAW	QVIVSGNLDN	TDCDYQGKYA	FSTSYNSEMG	MNLAEMTENE
X65278	VDAETMAVRW	QVLIDGNCDL	VATSYDGKLA	ATNQYNTENG	AHFEDMMSAE
Consensus	-D-----	QV-----	N-----	Y-G---	-----N-E--
				-----	M---E

301

350

AF119953	RDWVVVFEDIP	AIEKEIKAKR	FITLGDSKVP	VVDGRKKDGK	DSV..VTRYI
AF016059	RDWVVVFEDIP	AIEKEIKAKR	FITLGDSKVP	VVDGRKKDGK	DSV..VTRYI
X65277	RDWVVVFEDIP	AVEKEIKAGR	FITLGDSKVP	VVDGRKKDGK	DSV..VTRYI
M22628	RDWVVVFEDIH	AVEAAVKAGD	FITLGDSKTP	VLDGRKKDGK	DSK..FTRYV
AF056319	RDWVVVFENIL	RIEAAIKAGK	FITLDGSKVP	VVDGRKTDGK	DSE..FTRYI
AF119942	QDWAVVFEDIA	AIEAGVKAGD	FETYEGG.VP	VLDGRK..G.	.SK..YTRYI
AF119940	QDWAVVFEDIA	AIEAGVKAGD	FETYEGG.VP	VVDGRK..G.	.SK..YTRYI
AF119947	QDWAVVFEDIA	AIEAGVKAGD	FETYEGG.VP	VVDGRK..G.	.SK..YTRYI
AF119944	QDWAVVFEDIA	AIEAGVKAGD	FETYEGG.VP	VLDGRK..G.	.SK..YTRYI
AF119952	QDWAVVFEDIA	AIEAGVKAGD	FETYEGD.VP	VLDGRK..G.	.SK..YTRYI
AF119941	QDWAVVFEDIA	AIEAGVKAGD	FKTYEAD.VP	VLDGRK..G.	.SK..YTRYI
AF119936	QDWAVVFEDIE	AIEAGVKAGD	FKTYELD.VP	VLDGRK..G.	.SK..YTRYI

AF119949 QDHAVVFDIE AIEAGVKAGD FETI.NG.AK VLDGRK..G. .SK..YTRYI
 AF016057 QDWAVVFDIK AIEAGVKAGE FETI.GG.VL VLDGRK..G. .SK..YTRYI
 AF119945 QDWAVVFDIK RIEAAVKNGD YKKM.NG.VP VIDGRK..G. .SK..YTRYI
 AF119938 QDWAVVFDIK RIEAAVKSGD FKKM.NG.VP VIDGRK..G. .SK..YTRYI
 AF119943 QDWVVVFDIK RIEAGVKSGD FKKM.GG.VP VIDGRK..G. .SK..YTRYI
 AF119950 QDWVVVFDIK RIEAAVKNGD FKKM.GG.VP VIDGRK..G. .SK..YTRYI
 AF119954 QDWVVVFDIK RIEAAVKNGD FKKI.GG.VP VIDGRK..G. .SK..YTRYI
 AF119939 QDWVVVFDIK RIEAAVKNGD FKKM.GG.VP VIDGRK..G. .SK..YTRYI
 AF119935 QDWVVVFDIK RIEAAVKNGD FKKI.GG.VP VIDGRK..G. .SK..YTRYI
 AF119918 RDWAVVFDIK AIEAGVKAGD FKEI.HG.VP VLDGRK..G. .SK..YTRYI
 AF119951 QDWAVVFNK RIEDAVAAGE YEEM.SG.VP VLNGQK..G. .SA..LTRYI
 AF119946 QDWAVVFNK RIEEAVAAGD YEEM.NG.VP VLNGQK..G. .SD..YTRYI
 AF119937 QDWAVVFDIK AIEAAVAKGD FKEI.SG.VP VLDGQK..G. .SR..FTRYI
 AF119925 TDWAVIFNIK RIEEGVKAGD YTEH.SG.VP VLDGRK..G. .SK..YTAYV
 AF119948 QDWVVVFDIA RIEAAVKAGK AQKI.SG.VP VLDGRK..G. .SA..YTRYI
 AJ002531 QDWVVFNLK RIEEAVKKGD FKEM.NG.VP VIDGRK..G. .SP..YTRYV
 U47133 QDWVVFNLK RIEEAVAKGD YKEI.GG.VP VLDGRK..G. .SP..YTRYV
 AF047429 MDHVVFVNIA EIEKAIAAGD YEEI.NG.VK VVDGRK..EA KSL..FTRYI
 X774792 MDHIVVFVNIA EIEKAIAAGD YQEL.NG.VK VVDGRK..EA SSL..FTRYI
 AF125260 MDHVVFVNIP EIEKGIAAGD FQEL.NG.VK VIDGRK..GR NKN..YTRYI
 AF016055 TDHCVVFNK AIEDAVAAGE GQEL.NG.VR VLDGRK..QA SSQ..FTRYI
 AF016056 TDHCVVFNK AIEDAVAAGE GQEL.NG.VR VLDGRK..QA SSQ..FTRYI
 AF119927 MDHVVFIDLK AIEAEVAAGN YEEL.NG.VP VVDGRK..GS ..K..LTRYV
 AF119926 MDHVVFIDLK AIEAEVAAGN YEEL.NG.VP VVDGRK..GS ..K..LTRYV
 AF119922 LDHVVFNLK AIEEGVKSGD VEVL.NG.VP IINGMK..GE NKN..YTRYI
 AF119921 LDHVVFNLK AIEEGVKAGD VEVL.NG.VP IINGMK..GE NEN..YTRYI
 AF119920 LDHVVFNLK AIEEGVKAGD VEVL.NG.VP IINGLK..GE NKN..YTRYI
 AF119933 LDHVVLNLK AIEEGVKAGD VQIL.NG.VP VIDGRK..GS NKE..YTD.I
 AF119930 MDHVVFNLK EIEAGVASGD VQVL.NG.VP VIDGRK..GS NKK..YTRYI
 AF119928 MDHVVFNLK AIEEGVANGD VQML.KG.VP VIDGRK..GQ NKK..YTRYI
 AF119932 LDHVVFNLK EIEAGVAAGD FQEL.NG.VK VLDGRK..GQ NKN..YTRYV
 AF119931 LDHVVFNLK EIEAGVAAGD FQEL.NG.VK VLDGRK..GQ NKN..YTRYV
 AF119934 LDHVVFNLK EIEAGVAAGD FQEL.NG.VK VLDGRK..GQ NKN..YTRYV
 AF119929 LDHVVFNLK EIEAGVAAGD FQEL.NG.VK VLDGRK..GQ NKN..YTRYV
 AF119919 LDHVVFNLK EIEAGVAAGD YQEL.NG.IK VIDGRK..ET GKN..YTRYV
 AF119924 LDHVVFNLK EIEAGVAAGD FQEL.NG.VK VLDGRK..GQ NKN..YTRYV
 AF119923 QDHVVVFNLK EIEAGIAAGD FQEL.KG.VK VIDGRK..GQ NKN..YTRYI
 X65278 RDACVFFNIA RIEAAVQAGK FKTYGDSKVP VVDGTQAANK DPKTALTAYV
 Consensus -D--V----- --E----- -----G----- -----T---

351

400

AF119953 PVPKNPHGLN TSPDGKYFIA NGKLSPTCTM IAIDKLPDLF AGKCLKDPRDV
 AF016059 PVPKNPHGLN TSPDGKYFIA NGKLSPTCTM IAIDKLPDLF AGKCLKDPRDV
 X65277 PVPKNPHGLN TSTDGKYFIA NGKLSPTCSM IAIDLLPDLF AGKCLKDPRDV
 M22628 PVPKNPHGCN TSSDGKYFIA AGKLSPTCSM IAIDKLPDLF AGKLADPRDV
 AF056319 PVPKNPHGLN TSSDGKYFIA NGKLSPTVSM IAIDRLDDLF NDRYKDPREV
 AF119942 PISNSPHGVN TAPDGHHIVI NGKLSPTVSV IDVRKLPDVF ADKI.KPRDA
 AF119940 PISNSPHGVN TAPDGHHIVI NGKLSPTVSV IDVRKLPDVF ADKI.KPRDA
 AF119947 PISNSPHGVN TAPDGHHIVI NGKLSPTVSV IDVRKLPDVF ADKI.KPRDA
 AF119944 PISNSPHGVN SAPDGHHIVI NGKLSPTVSV IDVRKLPDVF ADKI.KPRDA
 AF119952 PISNSPHGVN TAPDGIHIVI NGKLSPTVSV IDVRKLPDVF ADKI.EPRAA
 AF119941 PISNSPHGVN TAPDGIHIVI NGKLSPTVSV LDVRKLPDVF ADKI.KPRDA
 AF119936 PIPNSPHGVN TAPDGHHIVI NGKLSPTVSV LDVRKFPDMF ADKI.KPRDA

AF119949	PVSNSPHGVN	TAPDGIHIVI	AGRLSPTMTV	LDVRKLPDLF	SDKI.GPRDV
AF016057	PVSNSPHGCN	TAPDGIHVVI	AGKLSPTVTV	IDVRRLPDLF	ADKI.GPRDV
AF119945	PISNSPHGCN	TAPDGHVVI	NGKLSPTVSI	LDVRRLPDLF	NDKI.KPRDV
AF119938	PISNSPHGCN	TAPDGHVVI	NGKLSPTVSV	LDVRRLPDLF	NDKI.KPRDV
AF119943	PISNSPHGCN	TAPDGHYVII	NGKLSPTVSV	LDVRRLPDLF	NDKI.KPRDV
AF119950	PISNSPHGCN	TAPDGHYVVI	NGKLSPTVSV	LDVRRLPDLF	NDKI.KPRDV
AF119954	PISNSPHGCN	TAPDGHYVVI	NGKLSPTVSI	LDVRRLPDLF	NDKI.KPRNV
AF119939	PISNSPHGCN	TAPDGHYVVI	NGKLSPTVSV	LDVRRLPDLF	NDKI.KPRDA
AF119935	PISNSPHGCN	TAPDGKYVVI	NGKLSPTVSV	LDVRRLPDLF	NDKI.KSRDV
AF119918	PVPNSPHGCN	TAPDGIHVVI	NGKLSPTVTV	IDVRKLPDLF	ADKI.KPRDT
AF119951	PISNSPHGCN	TAPDGKHIVI	NGKLSPTVSV	IDVTKLDALF	EEDA.DPRSA
AF119946	PISNSPHGCN	TAPDGKHIVI	NGKLSPTVSV	IDVEKIDALF	EEDA.DPRSC
AF119937	PISNSPHGCN	TAPDRRHIVI	NGKLSPTVSV	IDVTRLPDVF	ADRI.EPRGC
AF119925	PVANSPHGCN	TAPDGKHIVI	NGKLSPTCTV	LDVTKMDAVF	.DGA.DPRSC
AF119948	PVSKSPHGIN	TAPDGKHFV	AGKLSPIVTV	VDVTKLPDLF	NDKI.TPRDA
AJ002531	PVSNNPHGMN	TAPDGIHIVA	AGKLSPTVTV	MDVRLFNQLF	DDKI.KPRDV
U47133	PVPNSPHGIN	TAPDGIHVVA	NGKLSPTVTV	FDVRKFDDLF	DDKI.QARDT
AF047429	PIANNPHGCN	MAPDRKHLV	AGKLSPTVTV	LDVTKFDALF	YDNA.EPRSA
X774792	PIANNPHGCN	MAPDKKHLV	AGKLSPTATV	LDVTRFDVAV	YENA.DPRSA
AF125260	PVANNPHGCN	MAPDKIHLCI	NGKLSPTVTV	IDVRKLDALF	YEGA.EPRST
AF016055	PIANSPHGCN	AAPDKQHIVI	NGKLSPTVSV	IDVTKVDALF	SDNA.DPRSV
AF016056	PIANSPHGCN	AAPDKQHIVI	NGKLSPTVSV	IDVTKVDALF	SDNA.DPRSV
AF119927	PIPNSPHGIN	TAPDKKHVCI	AGKLSPTVSV	MDVTKIDALF	ADDA.DPRSC
AF119926	PIPNSPHGIN	TAPDKKHVCI	AGKLSPTVSV	MDVTKIDALF	ADDA.DPRSC
AF119922	PIPNSPHGVN	AAPDKKHIVI	NGKLSPTVSI	IDVEKVDALF	ADNA.DPRSC
AF119921	PIPNSPHGVN	AAPDKKHIVI	NGKLSPTVSI	IDVEKVDALF	ADNA.DPRSC
AF119920	PIPNSPHGVN	AAPDKKHIVI	NGKLSPTVSI	IDVEKVDALF	ASNA.DPRSC
AF119933	PIPNSPHGVN	TAPDKKHIVI	NGKLSPTVSI	IDVSKVDALF	ESDA.DPRSC
AF119930	PIPNSPHGIN	TAPDGKHVMV	AGKLSPTVSV	IDVTKLDALF	DSDA.EPRSC
AF119928	PIPNSPHGCT	AAPDQKHIMI	NGKLSPTVSV	IDVTKVDALF	ESDA.DPRSC
AF119932	PIPNSPHGIN	TAPDGKHVMI	NGKLSPTVSV	MDVEKLDALF	ASDV.DPRSC
AF119931	PIPNSPHGIN	TAPDGKHVMI	NGKLSPTVSV	MDVEKLDALF	ASDV.DPRSC
AF119934	PIPNSPHGIN	TAPDGKHVMI	NGKLSPTVSV	MDVEKLDALF	ASDV.DPRSC
AF119929	PIPNSPHGIN	TAPDGKHVMI	NGKLSPTVSV	MDVEKLDALF	ASDV.DPRSC
AF119919	PIPNSPHGIN	TAPDGKHVMI	NGKLSPTVSV	MDVEKLDALF	ADNA.DPRSC
AF119924	PIPNSPHGIN	TAPDGKHVMI	NGKLSPTVSV	MAVEKLDALF	ADNA.DPRSC
AF119923	PIPNSPHGVN	AAPDKKHIMI	NGKLSPTVSV	IDVEKVDGVF	AG.A.DERSC
X65278	SVPKNPHGVN	ASPDQKYFIC	AGKLSPTATV	IELSRVLGWF	DGKQEKLDDA
Consensus	-----PHG--	---D-----	-G-LSP----	-----F-----	-----

401

450

AF119953	VVGEPELGLG	PLHTTFDGRG	NAYTTLFIDS	QLVKWNLEDA	RRAYKGEK.V
AF016059	VVGEPELGLG	PLHTTFDGRG	NAYTTLFIDS	QLVKWNLEDA	RRAYKGEK.V
X65277	VVGEPELGLG	PLHTTFDGRG	NAYTTLFIDS	QVVKWNMEEA	RRAYKGEK.V
M22628	IVGEPELGLG	PLHTTFDGRG	NAYTTLFIDS	QVVKWNMEEA	VRAYKGEK.V
AF056319	IVAEPPELGLG	PLHTTFDGRG	NAYTTLFIDS	QVVKWNMDEA	IRAYKGEK.V
AF119942	IVAEPPELGLG	PLHTAFDNKG	NCYTTLFLDS	QVAKWSLDLA	RKKFAGED.V
AF119940	IVAEPPELGLG	PLHTAFDNKG	NCYTTLFLDS	QVAKWSLDLA	RKKFAGED.V
AF119947	IVAEPPELGLG	PLHTAFDNKG	NCYTTLFLDS	QVAKWSLDLA	RKKFAGED.V
AF119944	IVAEPPELGLG	PLHTAFDNKG	NCYTTLFLDS	QVAKWSLDLA	RKKFAGED.F
AF119952	IVAEPPELGLG	PLHAAYDNKG	NCYTTLFLDS	QVAKWSLDLA	RKKFAGED.V
AF119941	VVAEPPELGLG	PLHTAYDNKG	NCYTTLFLDS	QVAKWSIELA	KQQFAGKK.V
AF119936	VVAEPPELGLG	PLHTAFDNKG	NWYTTLFLDS	QVAKWSMDLA	RKEFAGED.V

AF119949	VVAEPELGLG	PLHTAYDGKG	IAFTTLFLDS	QVVKWDIQKA	VDAYAGKD.V
AF016057	VVAEPELGLG	PLHTAYDGTG	IAYTTLFFDS	QVVKWDIQKA	IDAYAGKD.V
AF119945	VVAEPELGLG	PLHTAFDGKG	NAFTTLFLDS	QVVKWNIQKA	IDAYSGKD.V
AF119938	VVAKPELGLG	PLHTAFDGKG	NAYTTLFLDS	QIVKWSMQKA	IDAYSGKD.A
AF119943	VAAEPELGLG	PLHTAFDGKG	NAFTTLFLDS	QAVKWNIQKA	VDAYSGKD.V
AF119950	VAAEPELGLG	PLHTAFDGKG	NAFTTLFLDS	QAVKWNIQKA	VDAYSGKD.V
AF119954	VAAEPELGLG	PLHTAFDGKG	NAFTTLFLDS	QAVKWNMQKA	VDAYNGKD.V
AF119939	VAAEPELGLG	PLHTAFDGKG	NAFTTLFLDS	QAVKWNIQKA	IDAYNGKD.V
AF119935	VAAEPELGLG	PLHTAFDGKG	NAFTTLFLDS	QAVKWNIQKA	VDAYAGKK.V
AF119918	VVAEPELGLG	PLHTAFDGKG	NCFTTLFLDS	QAVKWNMQKA	IDAYNGKN.V
AF119951	IVAEPPELGLG	PLHTAFDGQG	NCFTTLFLDS	QVCKWNLEAA	VKQYAGED.V
AF119946	IVAEPPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVAKWNMEAA	IKQYAGED.A
AF119937	IVAEPPELGLG	PLHTAYDDKG	NCYTTLFLDS	QVVKWSMEKA	IRAFKGEK.V
AF119925	VVAESQLGLG	PLHTAFDGKG	NCYTTLFLDS	QMVKWNMDLA	LRKYAGEE.V
AF119948	VVAEPKLGLG	PLHTAFDKKG	NAYTTLFLDS	QVAKWSIDKA	IRAYKGEK.V
AJ002531	VVAEPELGLG	PLHTAYDGKG	NAYTTLFLDS	QVVKWNIDLA	KRAFKEK.V
U47133	VVAEPELGLG	PLHTAYDGKG	NAYTTLFIDS	QVCKWNIEDA	KRAYAGEK.V
AF047429	VVAEPELGLG	PLHTAFDGRG	NAYTSLFLDS	QVVKWNIDEA	IRAYAGEK.I
X774792	VVAEPELGLG	PLHTAFDGRG	NAYTSLFLDS	QVVKWNIEDA	IRAYAGEK.V
AF125260	VVAEPELGLG	PLHTAFDGAG	NCYTSLFLDS	QVVKWNIDKA	IRAYNGEK.V
AF016055	IVAEPPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVAKWNMAKA	IEAYGGAD.V
AF016056	IVAEPPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVAKWNMAKA	IEAYGGAD.V
AF119927	IVAEPPELGLG	PLHTAYDGQG	NAFTTLFLDS	QIVKWNLDKA	IRAYAGEE.V
AF119926	IVAEPPELGLG	PLHTAYDGQG	NAFTTLFLDS	QIVKWNLDKA	IRAYTGEE.V
AF119922	IVGEPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVVKWNMEKA	IKAFAGEK.V
AF119921	IVGEPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVVKWNLEKA	IKAFAGEK.V
AF119920	IVGEPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVVKWNMERA	IKAFAGEK.V
AF119933	IVGEPELGLG	PLHTAFDGKG	NCFTTLFLDS	QVVKWNMEKA	IKAFAGEN.V
AF119930	IVAEPQLGLG	PLHTAFDGQG	NCFTTPFLDS	QVAKWNIDKA	IQAYGGAD.V
AF119928	IVAEPQLGLG	PLHTAFDGKG	NCFTTLFLDS	QVVKWNIDKA	IEAYGGAD.V
AF119932	IVAEPQLGLG	PLHTAFDGQG	NCFTTLFLDS	QVVKWNIDKA	LKQFAGED.V
AF119931	IVAEPQLGLG	PLHTAFDGQG	NCFTTLFLDS	QVVKWNIDKA	LKQFAGED.V
AF119934	IVAEPQLGLG	PLHTAFDGQG	NCSTTLFLDS	QVVKWNIDKA	LKQFAGED.V
AF119929	IVAEPQLGLG	PLHTAFDGQG	NCFTTLFLDS	QVVKWNIDKA	LKQFAGED.V
AF119919	IVAEPQLGLG	PLHTAFDGQG	NCFTTLFLDS	QVVKWNIDKA	LKQFAGDD.V
AF119924	IVAEPQLGLG	PLHTAFDGQG	NCFTTLFLDS	QVVEWNIDKA	LKQFAGED.V
AF119923	IVAEPPELGLG	PLHTAFDGKG	NAYTTLFLDS	QVVKWNIEDA	IASFGGAD.V
X65278	IVAEEVELGLG	PLHTAFDGRG	NAYTTLFLDS	QLVKWNLDAA	IKFHKGDKNA
Consensus	-----LGLG	PLH---D--G	---T--F-DS	Q---W----A	-----G-----

451

500

AF119953	DYIRQKLDVR	YQPGHNHASL	TETSEADGKW	VVLSKFSKD	RFLLTGPLHP
AF016059	DYIRQKLDVH	YQPGHNHASL	TETSEADGKW	VVLSKFSKD	RFLPTGPLHP
X65277	NYIKQKLDVH	YQPGHLHASL	CETSEADGKW	LVALSKFSKD	RFLPTGPLHP
M22628	NYIKQKLDVH	YQPGHLHASL	CETNEADGKW	LVALSKFSKD	RFLPVGPLHP
AF056319	NYIKQKLDVQ	YQPGHNHASL	TETSEADGKW	LVVLCKFSKD	RFLPTGPLHP
AF119942	DPILDKVDVH	YQPGHNHTSM	GETNKADGKW	LISLNKFSKD	RFINVGPLKP
AF119940	DPILDKVDVH	YQPGHNHTSM	GETNKADGKW	LISLNKFSKD	RFINVGPLKP
AF119947	DPILDKVDVH	YQPGHNHTSM	GETNKADGKW	LISLNKFSKD	RFINVGPLKP
AF119944	DPILDKVDVH	YQPGHNHTSM	GETNKADGKW	LISLNKFSKD	RFINVGPLKP
AF119952	DPILDKVDVH	YQPGHNHTSM	GETNKADGKW	LISLNKFSKD	RFINVGPLKP
AF119941	NPILDKVDVH	YQPGHNHTTM	GETNKADGKW	LISLNKFSKD	RFINVGPLKP
AF119936	EPILDKVDVQ	YQPGHNHTSM	GETNKADGKW	LVSLNKFSKD	RFINVGPLKP

AF119949	DPILDKVDVQ	YQPGHNHTSM	GETKEADGKW	LVSLNKFSKD	RFINVGPLKP
AF016057	SPILDKVDVQ	YQPGHNHSSM	GETKEADGKW	LVSLNKFSKD	RFINVGPLKP
AF119945	NPILEKIDVQ	YQPGHNHSSM	GETKEADGKW	LVSLNKFSKD	RFINVGPLKP
AF119938	NPILEKIDVQ	YQPGHNHSPM	GETREADGKW	LVSLNKFSKD	RFINVGPLKP
AF119943	NPILEKVDVH	YQPGHNHTSM	GETKDADGKW	LVSLNKFSKD	RFTNVGPLKP
AF119950	NPILEKIDVH	YQPGHNHTSM	GETKDADGKW	LVSLNKFSKD	RFINVGPLKP
AF119954	NPILEKIDVH	YQPGHNHTSM	GETKEADGKW	LVSLNKFSKD	RFINVGPLKP
AF119939	NPILEKIDVQ	YQPGHNHTSM	GETKEADGKW	LVSLNKFSKD	RFINVGPLKP
AF119935	DPIIEKIDVQ	YQPGHNHTSM	GETKEADGKW	LVSLNKFSKD	RFINVGPLKP
AF119918	DPIIDKLDVH	YQPGHNHTSM	GETKEADGKW	LVSLNKFSKD	RYINVGPLKP
AF119951	DPILOKIDVH	YQPGHNHTSM	GETKEADGNW	LVSLNKFSKD	RYINVGPLNP
AF119946	DPILOKIDVH	YQPGHNHTSM	GETKEADGKW	LVSLNKFSKD	RYINVGPLKP
AF119937	DPIILKVDVQ	YQPGHNHSSM	GETKEADGKW	LASLNKFSKD	RFLNVGPLKP
AF119925	DPILAKTDVQ	YQPGHNHSSM	GEAKDADGKW	LISLNKFSKD	RYINVGPLKP
AF119948	NPVLDKIDVH	YQPGHNHSSM	GETKDADGKW	LVSMNKFSKD	RFLNVGPLKP
AJ002531	DPIIQKLDVH	YQPGHNHSSM	GOTKEADGKW	LISLNKFSKD	RFLNVGPLKP
U47133	DPIRHKLDVH	YQPGHNHTSM	GOTKEADGKW	LISLNKFSKD	RYLNVGPLKP
AF047429	NPIKDKLDVQ	YQPGHLKTVM	GETLDAANDW	LVCLCKFSKD	RFLNVGPLKP
X774792	DPIKDKLDVH	YQPGHLKTVM	GETLDATNDW	LVCLSKFSKD	RFLNVGPLKP
AF125260	DPILOKLDVQ	YQPGHLKTSM	GETLEADGKW	LVALCKFSKD	RFLNVGPLKP
AF016055	DPIISKVDVH	YQPGHNSTSM	GETAEADGKW	LISMNKFSKD	RFLNTGPLKP
AF016056	DPIISKVDVH	YQPGHNSTSM	GETAEADGKW	LISMNKFSKD	RFLNTGPLKP
AF119927	DPIISKVDVA	YQPGHNSTTM	GETSEADGKW	LISMNKFSKD	RFLNVGPLKP
AF119926	DPIISKVDVA	YQPGHNSTTM	GETSEADGKW	LISMNKFSKD	RFLNVGPLKP
AF119922	DPIISKVDVH	YQPGHNSTSM	GETAEADGEW	LVSMNKFSKD	RFINVGPLKP
AF119921	DPIISKVDVH	YQPGHNSTSM	GETAEADGEW	LVSMNKFSKD	RFINVGPLKP
AF119920	DPIISKVDVH	YQPGHNSTSM	GETAEADGEW	LVSMNKFSKD	RFINVGPLKP
AF119933	DPIISKVDVH	YQPGHNSTSM	GETAEADGQW	LVSMNKFSKD	RFINVGPLKP
AF119930	DPILKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFLNTGPLKP
AF119928	DPILERVVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFLNVGPLKP
AF119932	DPIVEKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFINTGPLKP
AF119931	DPIVEKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFINTGPLKP
AF119934	DPIVEKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFINTGPLKP
AF119929	DPIVEKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFINTGPLKP
AF119919	DPIVEKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFINTGPLKP
AF119924	DPIVEKVDVH	YQPGHNSTSM	GETKEADGKW	LVSMNKFSKD	RFINTGPLKP
AF119923	DPIRDKVDVH	YQPGHNSTSM	GETAEADGKW	LISMNKFSKD	RFLNVGPLKP
X65278	KYVVDRDLQ	YQPGHVNASQ	SETVAADGKY	LAVGCKFSKD	RFLPVGPLHP
Consensus	-----D--	YQPGH-----	-----A----	-----KFSKD	R----GPL-P

501

550

AF119953	ENEQLID...				
AF016059	ENDQLIDISG	DEMMLVHDGP	TFAEPHDCIL	ARRDQIKTRK	I..WDRNDPF
X65277	ENDQLIDISG	DVMMLVHDGP	TFAEPHDCIM	ARRDQIKTRK	I..WDRNDPF
M22628	ENDQLIDISG	DEMMLVHDGP	TFAEPHDCIM	ARRDQIKTKK	I..WDRNDPF
AF056319	ENDQLIDISG	EEMMLVHDGP	AFAEPHDCIL	ARRDQIKTQK	I..WNRNDPF
AF119942	ENEQLID...				
AF119940	ENEQLID...				
AF119947	ENDQLID...				
AF119944	ENEQLID...				
AF119952	ENEQLID...				
AF119941	ENEQLID...				
AF119936	ENDQLID...				

AF119949	ENDQLID...				
AF016057	ENEQLIDISG	DKMAVLHDGP	TFAEPHDCII	VRADIVNPRS	V..WARDHAM
AF119945	ENEQLID...				
AF119938	ENEQLID...				
AF119943	ENEQLID...				
AF119950	ENEQLID...				
AF119954	ENEQLID...				
AF119939	ENEQLID...				
AF119935	ENEQLID...				
AF119918	ENEQLID...				
AF119951	ENDQLID...				
AF119946	ENDQLID...				
AF119937	ENEQLID...				
AF119925	ENEQLID...				
AF119948	ENEQLID...				
AJ002531	ENDQLIDISG	DQMKLVHDGP	SFAEPHDATI	VHRSKINPIS	I..WDRADPM
U47133	ENDQLIDISG	DEMVLVHDNP	TFAEPHDATI	VHASKINPVH	V..WNRDDPF
AF047429	ENDQLIDISG	DKMVLVHDGP	TFAEPHDAIA	VSPSILENIR	SVWDRKDPLW
X774792	ENDQLIDISG	DKMVLVHDGP	TFAEPHDAIA	VHPSILSDIK	SVWDRNDPMW
AF125260	ENDQLIDISG	DKMVLVHDGP	TFAEPHDAIA	VHASKM.NIR	SVWDRNDPMW
AF016055	ENEQLIDISS	DEMKVVDHGP	TFAEPHDSII	VHASKVNPV.	NIWSRDDPTW
AF016056	ENEQLIDISS	DEMKVVDHGP	TFAEPHDSII	VHASKVNPV.	NIWSRDDPTW
AF119927	ENEQLID...				
AF119926	ENEQLID...				
AF119922	ENEQLID...				
AF119921	ENEQLID...				
AF119920	ENEQLID...				
AF119933	ENEQLID...				
AF119930	ENEQLID...				
AF119928	ENDQLID...				
AF119932	ENEQLID...				
AF119931	ENEQLID...				
AF119934	ENEQLID...				
AF119929	ENEQLID...				
AF119919	ENEQLID...				
AF119924	ENDQLID...				
AF119923	ENEQLID...				
X65278	ENEQLIDISG	QKMVLMDHP	VRGEPHDFII	FKRELVRPKQ	VYALDDFPL.
Consensus	EN-QLIDIS-	--M-----P	---EPHD---	-----	-----

551

600

AF119953					
AF016059	FAETVAIAKR	DGIDLMQDNK	VIRDGKKVRV	YMTSSAPAYG	LTEFTVKQGD
X65277	FAPTVAIAKK	DGINLEEDNK	VIRGDNKVRV	YMTSMAPAYG	LTEFTVKQGN
M22628	FAPTVEIAKK	DGINLDTDNK	VIRDGKNVRV	YMTSMAPAFG	VQFTVKQGD
AF056319	FADTVALAKK	DGINLETDNK	VIRDGKNVRV	YMTSMAPAYG	LTEFTVKQGN
AF119942					
AF119940					
AF119947					
AF119944					
AF119952					
AF119941					
AF119936					

AF119949					
AF016057	WATTAAQAAK	DGVKLESGSE	VVREGNKVRV	YIWSVAPTFS	MEKFTVKQGD
AF119945					
AF119938					
AF119943					
AF119950					
AF119954					
AF119939					
AF119935					
AF119918					
AF119951					
AF119946					
AF119937					
AF119925					
AF119948					
AJ002531	FADAVKQAKA	DGINLEADSK	IIRDGNKVRV	YMTSTAPAFG	LEQFQVKQGD
U47133	FADAVAQAKA	DNIDLMVDSE	VIRDGNKVRV	YMTSAAPAFG	LDDFTVKQGD
AF047429	AETRKQAEAD	EVDIDEWTEA	VIRDGNKVRV	YMTSVAPSFS	QPSFTVKEGD
X774792	AETRAQAEAD	GVDIDNWTEE	VIRDGNKVRV	YMSSVAPSFS	IESFTVKEGD
AF125260	AETRKQAEAD	GVDIDDFTEG	VIRDGNKVRV	YLSGVAPSYS	LEKFTVTEGD
AF016055	EAARQQAAD	GVDLEDGHPE	PIRDGNKVRV	YMYSIAPEFS	LESFTVQQGD
AF016056	EAARQQAAD	GVDLEDGHPE	PIRDGNKVRV	YMYSIAPEFS	LESFTVQQGD
AF119927					
AF119926					
AF119922					
AF119921					
AF119920					
AF119933					
AF119930					
AF119928					
AF119932					
AF119931					
AF119934					
AF119929					
AF119919					
AF119924					
AF119923					
X65278		.AIKDPKESG	VFRNGRKVTV	KITSQAPAFS	LREFKLKKGD
Consensus	-----	-----	--R-G-KV-V	----AP--	---F-----

601

650

AF119953					
*AF016059	EVTVTITNLD	EIEDVSHGFV	MVNHGVSMEI	RPQQTASITF	TAGKP.GCSG
X65277	EVTVVITNMD	QIEDVSHGFV	MVNHGVSMEI	SPQQTSSITF	IADKPGLHWY
M22628	EVTVTITNID	QIEDVSHGFV	VVNHGVSMEI	SPQQTSSITF	VADKPGLHWY
AF056319	EVTVTITNID	QIEDVSHGFV	MTNHGASMEI	SPQQTSSITF	TADKAG....
AF119942					
AF119940					
AF119947					
AF119944					
AF119952					
AF119941					
AF119936					

AF119949					
AF016057	EVTLIVTNLD	KVDDLTHGIT	IANYGVAMEV	GPQDTSSVTF	NADRPGVHWF
AF119945					
AF119938					
AF119943					
AF119950					
AF119954					
AF119939					
AF119935					
AF119918					
AF119951					
AF119946					
AF119937					
AF119925					
AF119948					
AJ002531	QVTYIITNID	AVEDLTHGFC	IVNYGIQMEV	APMATASVSF	SADKAGVYWY
U47133	EVTYVVTNID	EVEDLTHGFC	IVNYGINMEV	APQATASVTF	KASRPGVYWY
AF047429	EVTVIVTNLD	EIDDLTHGFT	MGNHGVAMEV	GPQQTSSVTF	VAANPGVYWY
X774792	EVTVIVTNLD	EIDDLTHGFT	MGNYGVAMEI	GPQMTSSVTF	VAANPGVYWY
AF125260	EVTVIVTNLD	DVDDLTHGFT	MGNHGVAMEI	GPQATASVTF	VAAQPGVYWY
AF016055	EVTYVVTNLD	DIDDVTHGFC	LSNHGVAMEV	APQATASVTF	IADRPGVHWF
AF016056	EVTYVVTNLD	DIDDVTHGFC	LSNHGVAMEV	APQATASVTF	IADRPGVHWF
AF119927					
AF119926					
AF119922					
AF119921					
AF119920					
AF119933					
AF119930					
AF119928					
AF119932					
AF119931					
AF119934					
AF119929					
AF119919					
AF119924					
AF119923					
X65278	EVTLILTNLD	KIEDLTHGFA	IPKYNVNFIV	NPQETASVTF	VADKPGVFWC
Consensus	-VT---TN-D	---D--HG--	-----	-P---S-TF	-A---G--W-

651

675

AF119953		
*AF016059	ILHWFCHALH	ME.....
X65277	YCSWFCHALH	MEMVGRMMVE PA*..
M22628	YCSWFCHALH	MEMVGRMMVE PA*..
AF056319		
AF119942		
AF119940		
AF119947		
AF119944		
AF119952		
AF119941		
AF119936		

AF119949
 AF016057 YCQWFCHALH ME.....
 AF119945
 AF119938
 AF119943
 AF119950
 AF119954
 AF119939
 AF119935
 AF119918
 AF119951
 AF119946
 AF119937
 AF119925
 AF119948
 AJ002531 YCSWFCHAMH MEMKGRMFVE PKSV*
 U47133 YCTWFCHAMH MEMKGRMLVE AQGA*
 AF047429 YCQWFCHALH MEMRGRMFVE PKGA*
 X774792 YCQWFCHALH MEMRGRMLVE PKEA*
 AF125260 YCQWFCHALH MEMRGRMFVE PKAA*
 AF016055 YCQWFCHALH ME.....
 AF016056 YCQWFCHALH ME.....
 AF119927
 AF119926
 AF119922
 AF119921
 AF119920
 AF119933
 AF119930
 AF119928
 AF119932
 AF119931
 AF119934
 AF119929
 AF119919
 AF119924
 AF119923
 X65278 YCTHFCHALH LEMRTRMIVE A*...
 Consensus YC--FCHA-H -EM--R--VE -----

APPENDIX B

Genbank File of the Sequence of the nos Cluster from *A. cycloclastes*

LOCUS AF047429 9120 bp DNA BCT 28-MAY-1999
 DEFINITION *Achromobacter cycloclastes* regulatory protein NosR (nosR), nitrous oxide reductase precursor (nosZ), accessory protein NosD precursor (nosD), putative ATP/GTP binding protein NosF (nosF), accessory protein NosY (nosY), NosL precursor (nosL), and NosX (nosX) genes, complete cds.
 ACCESSION AF047429
 NID g4914312
 VERSION AF047429.2 GI:4914312
 KEYWORDS .
 SOURCE *Achromobacter cycloclastes*.
 ORGANISM *Achromobacter cycloclastes*
 Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae; *Achromobacter*.
 REFERENCE 1 (bases 1 to 9120)
 AUTHORS McGuirl, M.A., Nelson, L.K., Bollinger, J.A., Chan, Y.-K. and Dooley, D.M.
 TITLE The nos (nitrous oxide reductase) gene cluster from the soil bacterium *Achromobacter cycloclastes*: Cloning, sequence analysis, and expression
 JOURNAL J. Inorg. Biochem. (1998) In press
 REFERENCE 2 (bases 1 to 9120)
 AUTHORS McGuirl, M.A., Nelson, L.K., Bollinger, J.A., Chan, Y.-K. and Dooley, D.M.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-1998) Chemistry and Biochemistry, Montana State University/Bozeman, 108 Gaines Hall, Bozeman, MT 59717, USA
 REFERENCE 3 (bases 1 to 9120)
 AUTHORS McGuirl, M.A., Nelson, L.K., Bollinger, J.A., Chan, Y.-K. and Dooley, D.M.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-1999) Chemistry and Biochemistry, Montana State University/Bozeman, 108 Gaines Hall, Bozeman, MT 59717, USA
 REMARK Sequence update by submitter
 COMMENT On May 28, 1999 this sequence version replaced gi:2935349.
 FEATURES Location/Qualifiers
 source 1..9120
 /organism="*Achromobacter cycloclastes*"
 /strain="ATCC 21921; IAM 1013"
 /db_xref="taxon:223"
 /db_xref="ATCC:21921"
 gene 248..2443
 /gene="nosR"
 CDS 248..2443
 /gene="nosR"
 /note="two putative Fe4-S4 clusters at C-terminus"
 /codon_start=1
 /transl_table=11
 /product="regulatory protein NosR"
 /protein_id="AAD09156.1"

```

/db_xref="PID:g2935350"
/db_xref="GI:2935350"
/translation="MLRQRFAGKHPGMKLFRLVLLCLSLWGAVSASAAASVLQEFLPT
TTASDLVPGADGFGAVRDDVAAAPVLKGGQTVAVVFVTSDFVGTGYSKGPIHTLVAV
DKDAKIVGVRLVKHSEPIVLIGIPDAKVKALVAGYAGLDLVAEAKSGGTTHEVDIISG
ATVTVMVIDDSIVRSGLKVARALGLGGLTPEADNAGPRFEIDPAATAPADWVEAEGDG
TLRRLSLDVAQVNAAFAGHPDERARERAIQAPETTFFIEMQAALVSVPSIGKAFLGAA
EEKNLQAWLGEGEHAIAVVGRLYSFKGSGYVRGGIFDRIVLIQDDVSVRFRDSDHRR
LGAISLAGAPEYNEMDLFRIKAGLGFDPKPFRIQLLAQREVGPVEKFVHTFDLGYQL
PQKYLRAVAPPPVKDVAADVRDEETAQSDWLWKRIWDDSKLKIVVLGAMLVVLTGVFF
FQTYAVRNARAFYIFRIIFLTVTLVFLGWYANAQLSVNLMALFGSLVNGFSWQAFLL
DPLTFILWFVAAAALLFWGRGAYCGWLCFPGALQELTNQIARKLRIPQWTLFWGLHER
LWPIKYMIFLGLFGVSLMSVEQAHLAEIEPFKTAIILKFIRAWPFVAYATALLVAGL
FVERFYCRYLCLPGAALAI PARIRMFDWLKRYHECGNPCQCAHQCPVQAIHRTGEIN
PNECINCLHCQVLYQSETVCPVVIKKMSRAKLAASPGGVPTHQPAKV"
gene 2479..4407
/gene="nosZ"
sig_peptide 2479..2616
/gene="nosZ"
/note="periplasmic leader"
CDS 2479..4407
/gene="nosZ"
/codon_start=1
/transl_table=11
/product="nitrous oxide reductase precursor"
/protein_id="AAD09157.1"
/db_xref="PID:g2935351"
/db_xref="GI:2935351"
/translation="MESKEHKGLSRRALFSATAGSAILAGTVGPAALSLGAAGLATPA
RAATGADGSGVAPGKLDDYYGFWSSGQTGEMRILGIPSMRELMRVVFENRCSATGWGQT
NESIRIHQRTMTEKTKKQLAANGKKIHDNGDLHHVHMSFTDGKYDGRYLFMNDKANTR
VARVRCDVMKTDAILIIPNAKGIHGMRPQKWPRSNYVFCNGEDEAPLVNDGSTMTDVA
TYVNIFTAVDADKWEVAWQVKVSGNLDNCADADYEGKWAFFSTSYNSEMGMTLEEMTKSE
MDHVVFVFNIAEIEKAIAKAGQYEEINGVKVVDGRKEAKSLFTRYIPIANNPHGCNMAPD
RKHLVCVAGKLSPTVTVDVTKFDALFYDNAEPRSAVVAEPELGLGPLHTAFDGRGNAY
TSLFLDSQVVKWNIDEAIRAYAGEKINPIKDKLDVQYQPGHLKTVMGETLDAANDWLIV
CLCKFSKDRFLNVGPLKPENDQLIDISGDKMVLVHDGPTFAEPHDAIAVSPSILPNIR
SVWDRKDPWAETRKQAEADEVDIDEWTEAVIRDGNKVRVYMTSVAPSFQSPSFTVKE
GDEVTIVITNLDEIDDLTHGFTMGNHGVAMEVGPQQTSSVTFVAANPGVYWYYCQWFC
HALHMEMRGRMFVEPKGA"
mat_peptide 2617..4404
/gene="nosZ"
/EC_number="1.7.99.6"
/note="NosZ; multi-copper protein; CuA electron transfer
site; CuZ active site"
/function="denitrification enzyme; nitrous oxide reduction
to di-nitrogen"
/product="nitrous oxide reductase"
sig_peptide 4411..4464
/gene="nosD"
/note="putative periplasmic leader"
gene 4411..5739
/gene="nosD"
CDS 4411..5739
/gene="nosD"
/function="accessory protein for CuZ metallocenter
assembly"
/codon_start=1
/transl_table=11

```

```

/product="accessory protein NosD precursor"
/protein_id="AAD09158.1"
/db_xref="PID:g2935352"
/db_xref="GI:2935352"

/translation="MRLSVLLIGFLLASPLLAERAVMPGAGSLAAAIAGAQP GDVLV
LTDGAYDGPVTIDRPLTITGPRGAVVNGLKQGSVVTIAASDVTLTGFTVTGSGRINQD
LDAGVKIVQGADRARVEKLLVTDNMHGIDVHGGRDTIVSGNEIIGTRSARMNERGNGI
YVWNSPGTLLQDNIIRYGRDGIFSNASADSIYRRNIMRDLRFVAFHMYTRNTEVSDNI
SIGNHLGFAIMFSNRAKILNLSLGDREHGLMLNYANNADVSGNLVRGGTKKCLFIYN
AHKNLVWGNRFESCGIGIHFTAGSEKNVLTGNAFIANREQVKYVGTNRNMEWSHEGRGN
FWSHDPAFDLNGDGVADSFYRPNDLMDQILWSQPAASLLTGSPAVQIVRWSQRDFPAT
LPGGVDRDSAPLMRPLTIPVPLEILAYEAEAAGRWTENYDDTDADNLQAH"
mat_peptide 4465..5736
/gene="nosD"
/product="accessory protein NosD"
gene 5705..6604
/gene="nosF"
CDS 5705..6604
/gene="nosF"
/function="accessory protein for CuZ metallocenter
assembly"
/note="ABC-transporter like region; ATP/GTP binding motif"
/codon_start=1
/transl_table=11
/product="putative ATP/GTP binding protein NosF"
/protein_id="AAD09159.1"
/db_xref="PID:g2935353"
/db_xref="GI:2935353"
/translation="MTPTLTISRLTKRFQSVEALKSVSLALEPGRRAALLGHNGAGKS
TMMKIVLGLIPFDSGEVSVCGSPAGSPAARTQVAYLPENVAHFPAALTGEEQLRHYLAL
RGENPKQATELLARVGLGHAARRRIGTYSKGMRQRVGLAQLTIGRPRLVLVLEPTSGI
DPVSRRDYDILLDGLAAECTAILLSSHVLTEVEARTDSILILSGGELVAEGTLADLRT
RAALPVAFSVRPAPGHAPALAATFPEGVEGEDGLLRRLRGSQSEKPLLLARIAALGSAI
ADLDVIPPSLEDIYSHFSRRDQ"
gene 6601..7422
/gene="nosY"
CDS 6601..7422
/gene="nosY"
/function="accessory protein for CuZ metallocenter
assembly"
/note="putative integral membrane protein of an
ABC-transporter like system"
/codon_start=1
/transl_table=11
/product="accessory protein NosY"
/protein_id="AAD09160.1"
/db_xref="PID:g2935354"
/db_xref="GI:2935354"
/translation="MSRILATAVSEFRIALRNRWVSIA TGMMVLFALVLAAGSAPTG
DVGVDRLSVTVASLTSLAVYLVP LLALLMSFDAVAGEVERGTLPLLLTYPVSR LQILL
GKLLAHLAILGLAVTLGYGAAALAAVWFDPGATAGLGALWRLIWSSVLLGATFLGTGY
ALSALARPSGAAGLAVALWLVAVVLYDLALLALIVTDGGGAFTTHALPVALLANPAD
AFRVFNLSAAQAVSAAGGLGGAAGTIPLWQSAASLLAWPLAAIALAAA AFRKVTP"
sig_peptide 7419..7475
/gene="nosL"
/note="putative Type II signal peptidase leader"
gene 7419..8000

```

CDS /gene="nosL"
 7419..8000
 /gene="nosL"
 /note="putative N-terminal Cys anchor into membrane"
 /codon_start=1
 /transl_table=11
 /product="NosL precursor"
 /protein_id="AAD09161.1"
 /db_xref="PID:g2935355"
 /db_xref="GI:2935355"
 /translation="MRTRLRFVLVAAALALLSACKEDVAQSIVPQDMTPETLGHYCOM
 NLEHHPGPKAQIFLEGSPAPLFFSQVRDAIAYARGPEQIAPILVIYVNDMGAAGATWD
 QPGDGNWIAADKAFYVVGSAARRGGGAPEAVPFSSRDEAAAFVLAEGGQVLALADITD
 AMVLTPVETGSEPRADDEDYLGRRLRALPHPAGG"
 mat_peptide 7476..7997
 /gene="nosL"
 /product="NosL"
 gene 8005..8916
 /gene="nosX"
 CDS 8005..8916
 /gene="nosX"
 /note="possible regulatory partner of NosR"
 /codon_start=1
 /transl_table=11
 /product="NosX"
 /protein_id="AAD09162.2"
 /db_xref="PID:g4914313"
 /db_xref="GI:4914313"
 /translation="MLLTRRLIAISAVLATAPAIHANAAATRFWTGQALGARASIRI
 AHPEAGTIVSRVMAEIDRLLENILSLYRANSALCRLNRDGHLDQPPFELLECLSLADAV
 HLASGGRFDPTVQPLWLWAQSAANGTRPGPADVDRVRGNVGWGVKVLGAAAITLHPG
 MALTNLNGIGQGYVADRVAAMLEVEGLTDILIDTGEFRALGGRPGGGGWVPRLETGERL
 PLRQRAVATSAPLGTTFDQAARDGHILDPLTGMPAPANWRSVSVSASSAAIADALATA
 ACLLPDAGSIVAMNARFDGVRLEAASV"

BASE COUNT 1644 a 2923 c 2801 g 1752 t
 ORIGIN

```

1 gtagttgtgg gtggcgggct gcctgccgtc ccctggcaga tatcgcgggg agggcaggcc
61 acgggcccgg cgatccgact tcaaaggat tggtcttgcg ccggtcggcc gccggcttga
121 cgcgaaagtc tttccccagc tccacctgag catcaatccg tgcacgacat gcgacaatca
181 aggattgcaa tcaacgaagc tgtctttggt tgagctagat caaagaatgt cccgcgccaa
241 agtttcaatg ctccggcaac gatttgcggg aaaacaccct ggaatgaaac tcttcgctct
301 tgtccttctt ctctgcctca gcctgtgggg cgctgtatcc gccagcgag catcggtgct
361 tcaggagttt ttgccgacca ccacagctc cgatcttggt ccaggagcgg acggttttgg
421 cgctgtgcgt gatgatgtgg cggcgccccc ggtgctgaag ggcggtcaga cgggtggcgtg
481 ggttttcgtg acctcggatt tcgtcggaac gaccggctat tccggcaaac ccatccatac
541 gctcgtggcg gtggacaagg atgcgaagat cgtcggcgct cggctagtga aacattcgga
601 acccatcgct ctcacgcgca ttccggacgc caaggtgaag gcgctttagt ccggctacgc
661 aggccttgat ctcgctcgcg aggcgaaatc cggcggtacg acccacgagg tagatatcat
721 ctcgggcgcg acggtgacgg tcatggtcat cgacgattcg atcgctccgt cggggctaaa
781 ggttgcgcg gcgctcgggc ttggcggatt gacaccgga gcgacaatg ccggcccgcg
841 cttcgaaatc gatcccgcg ctaccgcgcc ggccgactgg gtcgaggcag aaggtgacgg
901 cacgctgcgc cgctgtcgc tcgacgttgc gcaggtcaat gctgctttcg ccggccatcc
961 cgatgagcgc gcccgcgagc gcgcgattgc gcaagcgccg gaaacgacat tcatcgagat
1021 gcaggcgcca ctggtctcgc tcccttccat cggcaaggcc tttctcgggg cagccgaaga
1081 gaaaaacctt caggcttggc ttggcgaagg cgaacacgcg atcgccgtgg tcggacgcgg
1141 gctctattcc ttcaagggat caggctatgt gcgcggcgcc atctttgatc gcatogtgct
1201 gatccaggac gatgtttccg ttcgctttcg tgatagcgac caccgtcgcc tggggcgaat
1261 ctcgcttgcc ggtgcgcgg aatacaacga gatggacctg ttccgcatta aagcaggcct
1321 cggcttcgac ccgacaaaac ctttcgcat acagctcctg gccaacgctg aggttggacc
  
```

1381 ggtcgaaaaa gtctttcata ccttcgatct gggttaccag cttccgcaga agtatctgag
 1441 tgctgtcgcg ccaccaccgg tgaaggacgt tgcagccgtt gctgatctgt acgaggagac
 1501 tgcacagtcc gacctgtgga aacgcatctg ggacgactcg aagctgaaga ttgtcgtgct
 1561 cggcgccatg ttggtcgtgc tgacgggctg gttcttcttc cagacctatg cggtcaggaa
 1621 cgcgcgagcg ttctacatct ttccgcatcat cttcctgaca gtaacgctgg tctttcttgg
 1681 ctggtatgcc aacgcgcagc tttctgtcgt caatctgatg gcgctcttcg gcagctcgt
 1741 caacgggttc agctggcagg ctttctctgt cgatccgctc accttcatcc tgtggtttgc
 1801 cgttgcccgcg gcaactcctgt tctggggagc aggagcctat tgcggctggc tctgcccctt
 1861 cggcgactcg caggagctga ccaaccagat cgcacgcaag ctgcgtatcc cgcaatggac
 1921 attgccatgg ggcctgcacg agcggctctg gccgatcaaa tacatgatct tcctcgccct
 1981 tttcggggtc tctctgatga gcgtcgagca ggcagagcat ctggccgaga tcgagccatt
 2041 caagacagcg atcatcctga agttcatccg cgccctggccg ttcgtggcct atgccacagc
 2101 acttctcgtc gccgggctct tcgtcgagag gttctattgc cgctacctct gcccgctcgg
 2161 tgcggcgctc gcgatccctg cgcgcatacg catgttcgac tggctcaagc gctaccatga
 2221 atgcggcaac ccctgccaga cctgcgcca ccaatgtccg gtgcaggcga tccaccggac
 2281 cggcgagatc aatccgaacg agtgcataca ctgcctgcac tgccagggtg tctaccagtc
 2341 cgaaccgctc tgcccgtcgc tgatcaagaa gatgaagtcg cgggcgaagc ttgcggcaag
 2401 cccggggcggc gtgcccagccc atcaaccgcg agccaaagtt taatcgaagc caaggttcaa
 2461 gcaaaaagga accagatcat ggaatcaaag gaacacaagg gactaagccg gcgagcactt
 2521 ttcagcgcgca ctgcaggcag cgccattctg gcgggactg tagggccggc cgcactcagc
 2581 ctccggcgctg caggattggc gacaccggcc cgtgcggcca cgggtcgcca cggctcggtc
 2641 gcaccgggca agctcgacga ttactacggc ttctgggtcct ccggtcagac cggcgagatg
 2701 gcatttctcg gcattccctc gatgcgcgaa ctgatgcgcg tgccgggtgt caaccgttgc
 2761 tcggccaccg gttgggggca gaccaatgaa tcgatccgga tccatcagcg cacgatgacg
 2821 gaaaagacga agaagcagct tgccgccaac ggcaagaaaa tccacgacaa cggcgacctg
 2881 catcacgtcc atatgtcctt caccgacggt aaatatgacg gtgcgtacct gttcatgaac
 2941 gacaaggcca atacgcgcgt ggcgcgctg cgtgcgacg tgatgaagac cgacgccatc
 3001 ctggagatcc ccaacgcaa gggcatccac ggcgtgcgtc cgcagaaatg gccgcgttca
 3061 aactatgtgt tctgcaacgg cgaggacgag gcccgcctcg tcaacgacgg ctcgacgatg
 3121 acggacgtgg cgacctacgt gaacatcttc accgcgctcg atgccgacaa gtgggaagtg
 3181 gcgtggcagg tgaaggtctc gggcaacctc gacaattgcg atgccgacta cgagggcaag
 3241 tgggcgttct ccaccagcta caactccgaa atgggcatga cgctggagga gatgaccaag
 3301 tccgagatgg atcatgtcgt cgtcttcaac atcgccgaaa tcgagaaggc cattaaggcc
 3361 ggccaatatg aggagatcaa cggcgctcaag gtgggtggac ggcgcaagga ggcaagtcg
 3421 ctcttcacgc gctacatccc gatcgccaac aacccccacg gctgcaacat ggcgccggac
 3481 aggaagcatc tgtgcgttgc cggcaagctt tcgccaaccg tcaccgtgct ggacgtgacg
 3541 aagttcgtat ccctgttcta cgacaatgcc gagccgcgca gcgcggtggt tgccgaaccg
 3601 gaaactgggccc ttggccatt gcacaccgcc ttcgacgggc gcggcaacgc ctatacctcg
 3661 ctgttcctcg acagccaggt ggtaaagtgg aacatcgatg aggccatcc cgcctacgcc
 3721 ggtgagaaga tcaaccgat caaggacaa gctcgatgcc gccaacgact ggctcgtttg cctgtgcaaa
 3781 aagacgggtga tgggcgaaac gctcgatgcc ccgctgaagc cggaacacga tcagttgatc
 3841 ttctccaagg accggttctt gaatgtcggc catgacggcc cgaccttcgc cgagccgcat
 3901 gacatttccg gtgacaagat ggtgctggtc catgacggcc ccgaacatcc gctcggtctg ggatcgcaag
 3961 gacgccatcg ccgtctcccc ctcgatcctg ccgaacatcc gccgaggccg atgaggtcga catcgacgaa
 4021 gatccgctgt gggccgaaac ccgcaagcag gccgaggccg tctacatgac ctcggtcgca
 4081 tggaccgagg ccgtgatccg cgacggcaac aaggttcgcg acgaagtcac ggtcatcgtc
 4141 cccagcttca gccagccgag ctttacggtg aaggaggcg acgaagtcac ggtcatcgtc
 4201 accaatctcg atgaaatcga tgaccttacc catggcttca ccatgggcaa tcacggcggtg
 4261 cggatggagg tcggcccgcga acagacgagc tccgttacct tcgttcgggc caatcctggc
 4321 gtctactggt actattgcca atggttctgc catgcccctg acatggaaat gcgcggccgc
 4381 atgttcgtgg aaccgaagg cgccctgatc atgcggctgt ccgttctcct gatcggtttt
 4441 cttctcgcct ctccgctctt agcggcgagg cgcgagtcga tggcggtgct gggcagctct
 4501 gccgcagcca tcgccggggc tcaaccggcg gatgtgctgg tgttgaccga cggggcctat
 4561 gacggccccg tcaccatcga tcggccactg accatcaccc gtccgcgtgg agcgggtggtg
 4621 aatgggctta agcagggtag cgtggtgacg attgccgctg cggacgtcac cttgacgggg
 4681 ttcaccgtca ccggttcggg tcgcatcaac caggatctcg acgcccgggt caaaatcggt
 4741 caaggcgccc atcgcgccc tgctgaaaag ctgctcgta cgcacaacat gcacggcatc
 4801 gacgtccatg gcggccgcga cacgattgtc tccggaaacg agatcatcgg cacgocgacg
 4861 gcccgcatga acgagcgcg taacggcatc tatgtctgga acagccccgg tacgctgctt

4921 caagacaata ttatccgcta cggctcgcgac ggcatctttt cgaacgccag cgccgacagc
 4981 atctatcgcc gcaacatcat gcgcgacctg cgctttgccc tgcacttcat gtacaccgcg
 5041 aacaccgagg tttccgacaa catctcgatc ggaaccatc tgggctttgc gatcatgttc
 5101 tccaaccgcg cgaagatcct gaacaatctc agcctcggcg accgcgagca tgggctgatg
 5161 ctgaactatg ccaacaatgc cgacgtgtca ggcaatctgg tgcgcggggg cacgaagaaa
 5221 tgcctgttca tctacaacgc ccacaagaat ctggtctggg gaaatcgctt cgagagctgc
 5281 ggcattggca tacatttcac cgccggttcg gaaaagaatg tgctgacagg caacgccttc
 5341 atcgccaacc gcgagcaggt caaatacgtc ggcaaccgca atatggaatg gagccatgaa
 5401 ggccgggggca atttctggtc cgatcatccg gccttcgacc tcaacggcga cggcggtgac
 5461 gacagtttct accgccccaa cgacctgatg gaccagatcc tctggtcgca gcccgcgga
 5521 agcctgctca cgggtcgcgc ggccgtgcag atcgttcgct ggagccagcg ggattttccc
 5581 gccacgctgc ctggaggcgt gcgcgacagc gcgcgctga tgcgccccct gacctccct
 5641 gtcccgcctg aaatccttgc ctatgaggcc gaggcgcgag ggcgctggac tgaaggaaat
 5701 tacgatgaca ccgacgctga caatctccag gctactaag cgtttccaat cggtggaagc
 5761 gctcaagtcc gtgtcccttg cgctggagcc gggacggcgc gcggcactgc tcggccataa
 5821 cgggtcggggc aagtcgacga tgatgaagat cgtgcttggc ctcatcccct tcgacagcgg
 5881 cgaggtttcg gtttgccggt cggcgcccg ttcgccccgc gcgcggaccc aggtggcgta
 5941 tctgcccga aatgtcgctt ttcacccggc tttgacggga gaagaacagc ttcgccacta
 6001 tctcgactgc cgtggcgaaa accggaagca ggccacagag cttcttgcca gggtcggact
 6061 tggccacgcg gcgcggcgtc ggatcggtag ctattccaag ggcagtcgcc agcgggtggg
 6121 tctggcccag acgctgatcg gccgcccccg tctcctggtg ctggacgaac cgacctcggg
 6181 cctcgaccgg gtgtcgcgac gggacttcta tgacctgctc gacgggttgg ccgcagaagg
 6241 aaccgccatc ctgctctcct cccacgtgtt gaccgaggtg gaggcccgca ccgacagcat
 6301 cctcatcctt tccggtgggg aactggtggc agaaggtacg cttgccgac ttcgaccccg
 6361 cgcggcactt cgggtcgcct tttcgttcg ccccgacccg ggccatgcgc cggcgctggc
 6421 agcgaccttt cccgaggggg tgggaaggga ggacggtctc ttgccccttc gcgggtcgca
 6481 gtccgaaaaa ctgccgctcc ttgcccgcag tgcggcgctt gggagcgcca ttgccgatct
 6541 cgatgtcatc ccgcccagcc tcgaggacat ctacagccac ttcagccgga gggacgggca
 6601 atgagccgca tccttgccac cgccgtcagc gaattccgta tcgctttgcg caatcgttgg
 6661 gtctctatcg ccaccggcat gatggtgctg cgtcgaccgg ctctccgtaa cggtcgcctc gctcacttcg
 6721 gcccgcagcg gcgatgtcgg atcgaccgg ctctccgtaa cggtcgcctc gctcacttcg
 6781 cttgcggtct atctggttcc gcttttgcca ttgctgatga gtttcgatgc ggtggctggc
 6841 gaagtgaac gcggcacgct gccactgctt ctacactatc ccgtctcag cctgcagatt
 6901 ctgctcggca aactgctggc gcactctcga atcctcggtc ttgccgtgac gcttggttac
 6961 ggcgcggcgg cgctcgcagc cgtctggttc gatcctggag caaccgcagg gcttggcgc
 7021 ctatggcggc tgatctggtc ctctgtgctg ctgggtgcga cctttcttgg caccggctat
 7081 gcgctttcgg ctttggcgcg acggccgctg ggggcggcgg gcttggcgtg cgcgctctg
 7141 ctggtggcgg tgggtctcta cgacctcgcc ctggtggcgt taatcgtgac ggatggaggc
 7201 ggtgccttca ccacacacgc attgcccgtt gcgctgcttg ccaaccctgc cgacgccttt
 7261 cgcgtcttta acctctcggc gcacacaggcc gtttctgccc caggcgagct gggtggcgc
 7321 gccggcacga tcccgtctg gcaatcggca gcctcccttc tcgcttgccc actcgccgc
 7381 atagcccttg ccgcccagc cttccgaaa gtagcccat gaggaccgga cttcgcttcg
 7441 tgctcgctgc ggccgtctg gcgctgctct ccgctgcaa ggagacgct gcccaaagca
 7501 tcgtgccgca ggacatgac cccgagacc tcggtcatta ctgccagatg aaccttctg
 7561 agcaccgccg ccgaaggcc cagatctttc tgggaaggaa ccccgcccg ctcttcttca
 7621 gtcaggtgcg cgatcgatc gcctatgcgc gtgggcccga gcagattgag ccgatcctc
 7681 tcatctatgt caatgacatg ggggtcgcg gggcaacgtg ggaccagccc ggtgacggca
 7741 actggatcgc cgccgacaag gcgttctacg tggtcggctc tgcccgcgga ggcgcatg
 7801 gtgcgcccga agccgtgccc ttttcaagcc gcgacgaggc tgccgcttc gttcttgccg
 7861 agggcgccca ggtgcttgcg ctgcgcgata ttaccgatgc catggtttt acgcccgtg
 7921 agaccggctc ggaaccgcga gcggacgagc aagactatct cggccgtctg cgcgacttgc
 7981 cccatccagc cggaggttga ccccatgctg ctgacccgac gccgcttgat tgcaatttct
 8041 gcagttctcg caacggcgcc cgccatgcgc cacgccaacg cggcaaccgc cttctggaca
 8101 ggtcaggctc ttggcgcgcg tgccctgatc cgcattgcgc accccgaggc cgggacgatc
 8161 gtctccaggg tcatggcgga gatcgaccgt ctggaaaaca tccttagcct ctaccgtgcc
 8221 aacagcgcgc tttgcccgtt gaaccgcgac ggtcatcttc aagaccgcc attcgagctt
 8281 ttggaatgtc tgtcactggc cgacgcagtc cacctggcta gcggcgggcg cttcgatccg
 8341 acggtacagc cgctctggac gctctgggct caatcggtg caaacggcgc gcgaccggc
 8401 ccggcgagc ttgatcgct gcgtggcaac gtcggctggg gtaagggtcaa gctcggtgca

8461 gcagcgatca cgttgcattc gggcatggcg ctgacgttga acggcatcgg ccaggggtat
8521 gtcgccgacc gcgtagcagc aatgcttgaa gtcgagggac tgacggacat tctgatcgat
8581 accggcgagt tccgtgcgct cgggggacga ccgggcggcg gcggatggcc cgtccggctg
8641 gaaacaggcg agcgctgcg gctgcgccag cgggcggtcg ccacctcgc ccctctcggc
8701 acgaccttcg atcaagcagc cagagacgga catattctcg acccctgac cggaatgcc
8761 gcgccagcaa actggcgctc tgtctcggcg tcggcatcat ccgctgccat cgccgacgcg
8821 ctggcgacgg ctgcatgcct gctgcctgat gccggatcga tcgttgcat gaacgcgcgt
8881 tttgacggcg tccgcctgga agccgcctcc gtttaaggcg ggcgaccaag gaacgtcatc
8941 gggggcttga gaagggtcgg atccgtctcg gcccgatcg gcgaaaggta ggcggctgcg
9001 accgcagcga cgatgcctg tgatggtgtt ggtgttcttc ctgagcgcgc cgaaggggca
9061 gtgcttgcg actacctgc ctccactggt cctgttccgc gcccatcgag aagccaggca

APPENDIX C

Genbank File of the Sequence of *nosR* from *S. meliloti*

LOCUS AF047430 2426 bp DNA BCT 06-MAR-1998
 DEFINITION Sinorhizobium meliloti regulatory protein NosR (nosR) gene, complete cds.
 ACCESSION AF047430
 NID g2935357
 VERSION AF047430.1 GI:2935357
 KEYWORDS .
 SOURCE Sinorhizobium meliloti.
 ORGANISM Sinorhizobium meliloti
 Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Sinorhizobium.
 REFERENCE 1 (bases 2109 to 2426)
 AUTHORS Holloway, P., McCormick, W., Watson, R.J. and Chan, Y.K.
 TITLE Identification and analysis of the dissimilatory nitrous oxide reduction genes, nosRZDFY, of Rhizobium meliloti
 JOURNAL J. Bacteriol. 178 (6), 1505-1514 (1996)
 MEDLINE 96198150
 REFERENCE 2 (bases 1 to 2426)
 AUTHORS McGuirl, M.A., Nelson, L.K., Bollinger, J.A., Chan, Y.-K. and Dooley, D.M.
 TITLE The nos (nitrous oxide reductase) gene cluster from the soil bacterium Achromobacter cycloclastes: Cloning, sequence analysis, and expression
 JOURNAL J. Inorg. Biochem. (1998) In press
 REFERENCE 3 (bases 1 to 2426)
 AUTHORS McGuirl, M.A., Nelson, L.K., Bollinger, J.A., Chan, Y.-K. and Dooley, D.M.
 TITLE Direct Submission
 JOURNAL Submitted (06-FEB-1998) Chemistry and Biochemistry, Montana State University/Bozeman, 108 Gaines Hall, Bozeman, MT 59717, USA
 FEATURES Location/Qualifiers
 source 1..2426
 /organism="Sinorhizobium meliloti"
 /db_xref="taxon:382"
 gene 159..2426
 /gene="nosR"
 CDS 159..2426
 /gene="nosR"
 /function="regulation of NosZ (nitrous oxide reductase)"
 /note="two putative Fe4-S4 clusters"
 /codon_start=1
 /transl_table=11
 /product="regulatory protein NosR"
 /protein_id="AAC05094.1"
 /db_xref="PID:g2935358"
 /db_xref="GI:2935358"
 /translation="MYQAAIHKNLFRLLRGLVALCIAMLLAGPSMAAGQLSNYLQKV
 QPQQIFPGATRFGEVTDGPPVFRGESLLGYAYLNSDVTSSVGYSGKPIHIVVGID
 PKGVVRGLKLVDHKEPIVLIGIPEAKVVASVNALIGKDLGRVSAGVEGPPQVDIVSGA

TVTTLVLMGDSIVRSALKLIRGNRFADAVAPEQSSEIRKVDLLRTGTSDWETLVGDGS
 VRSLRLTVGEVTEAFRQAGQPAADRPEHSPGDRFIDLIAIPVSVFVIGRSLLGDSR
 YEQMRALKLPGEAAIVVAGDGAYSFKGSGYVRGGIFDRIELIQDQGGLRFRDRYHTRL
 PTLAASGAERLREIALFVVPADFGFDIAAPWELQLLVQRTAVGRDKAVLPYNLSYVLP
 DAHVTVAAAAAPSPVEPSIAEQATPAEPIPEGEALWVKMWEMNRVSVAVTMAAVLVLT
 LIAFFQDWLVKRPALFAWVRRVYLLFTLVWLGWYANAQLSVVNVLTFFNSLVTGFHWE
 FFLSAPLVFLLWGSVAAALLFWGRGPFCEWLCPFGALQELTNNVARWLKVPQVRVPWG
 LHERLWPIKYIIFLGLFGLSLYLALAEAFVEPFKTAIILKFAREWPFVVFALTVL
 AAGLFIERFYCRYLCPLGAALAIIPGRIRMFEWLKRWPECSPCQRCAKECPVQSIHPE
 GAINVNECTIYCMHCQELYHDDQRCPHMIQVRLKREKFMALSTPASRGEAPAKTVVTHK
 GAPIRKADAAPENPV"

BASE COUNT 414 a 766 c 765 g 481 t
 ORIGIN

```

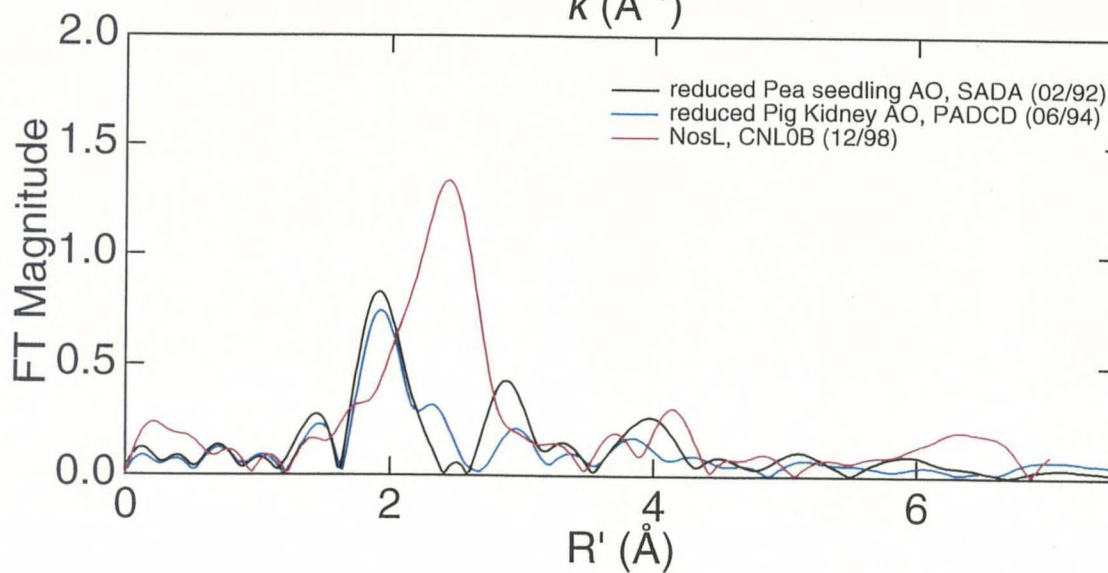
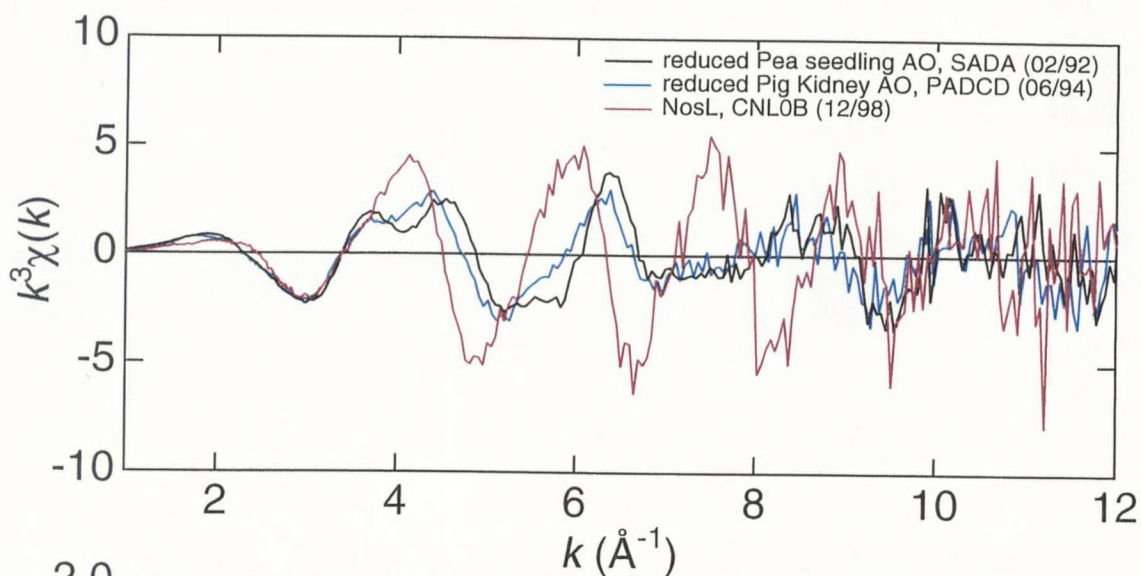
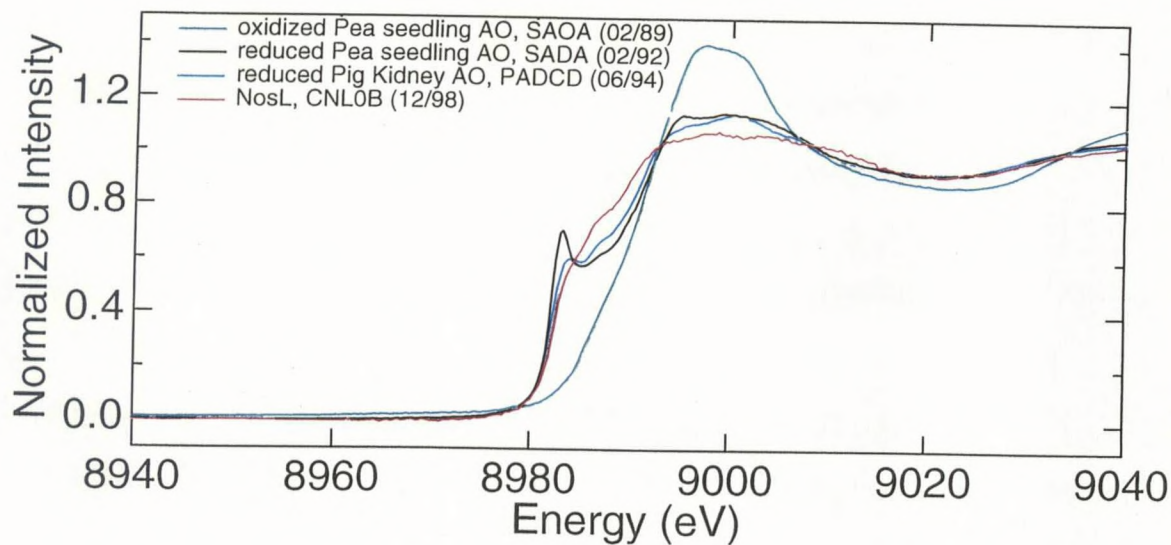
1  gtggcacgcg caaggtgcgc aatttgagct cgatcttctg gacctgagac acattggacc
61 gagttgaatt taatcaagtt gatctggcgc aaagaaattc atctcgggcc gtgcgaaatg
121 gccgcgatgc agtggctccg tgtggccacg gaaacccgat gtaccaagca gcgatacata
181 agaacctatt caagcgctt ctgcgccgct tcgtcgctct ctgcatcgct atgttggttg
241 ccggcccaag catggccgcg ggccaactgt ccaactatct tcaaaagggt cagccgcagc
301 agatttttcc aggtgcaacc cgcttcggcg aggtcaccgg cgatcccccg atcgcaccgg
361 tcttcgcgcg cgaaagcctg ctgggttacg cgtatctcaa ttccgacgtc acaagctccg
421 tcggctattc cggcaagccc atccatatcg tggtggaat cgatccgaaa ggcgtcgta
481 gcgggctgaa gctcgtcgac cataaggaac cgatcgttct aataggcatt cccgaagcga
541 aggtcgttgc ctctgtcaat cgcctcatcg gcaaggacct cgggcggggtc tcggcggggg
601 tcgaggggcc gccgcaggtc gacatcgta gccgcgccac cgtgaccgta ctggtgatgg
661 gcgacagcat tgtccgttcc gccctaaagc taatccgcgg gaacaggttc ggcgcgcagc
721 ccgtcgcgcc cgagcaatcg tccgaaatcc gcaaggtcga cctcttaagg acggggacga
781 gcgactggga gacgtgggtt ggcgacggat cgggtgcgaag ccttcgcctt acggtgggcg
841 aggtcacgga agcgttccgc caggccggcc agccggcagc cgccgaccgg cccgagacgc
901 acagtcccg cgatcggttc atcgatctct acatcgcgcc ggtgagcgtg ccggtgatcg
961 gaaggtcgct gctcggcgat tcccgttacg agcagatcg cgccaaggct aagccgggag
1021 aggaggcgat cgtcgttgcg ggagacgggg cctattcctt caagggctcc ggctacgtcc
1081 gtggcgcat ttcgaccgc attgagctta tccaggatgg ccagggcctg cggttccgcg
1141 accgtacca cacacggctg ccaacgctag ctgcgtccgg cgcgccgct cttcgcgaaa
1201 tcgccctggt tgtggttccg gcggacttcg gtttcgacat tgccgcaccg tgggaactgc
1261 agctactggt tcagcgacg gccgtcggcg gcgacaaggc ggtgctacct tacaactga
1321 gctatgtcct cccgagcgcc catgtgaccg tcgaggccgc agcagcgctg tcgccgtcg
1381 agccaagcat cgcggaacag gcaacacccg ccgagccgat cccggaaggc gaggccctct
1441 gggtagaat gtgggagatg aaccgggtgt cggtgccggt cacgatggct gccgtgctcg
1501 tgctgacct gatcttctt ttccaggact ggctggtgaa gcggccggcg ctgtttgcgt
1561 gggtcagacg cgtctacctg ctcttcacgc tgggtgtggt cggctggtat gccaacgccc
1621 agctctctgt tgtcaacgct ctgacattct ttaactccct ggtgacggcg ttccattggg
1681 agttctttct ctccgcccc ttggtgttcc tectgtgggg gagcgtcgcc gcgcgctgc
1741 tgttctgggg gcgcggcccc ttctgcggct ggctatgcc gttcggcgcg ctgcaggaac
1801 tgacgaacaa tgtcgcgcga tggtgaagg tgcctcaggt gagagtgcg tggggcctgc
1861 acgagcggct ctggccgatc aaatacatca tctttctcgg cctgttcggg ctgtcgctct
1921 actccctcgc gcttgccgag atgtttgctg aagtcgagcc gttcaagac gcgatcatoc
1981 tgaagttcgc ccgcgaatg cccttcgtcg tctttgcgct caccgtgctc gcggccggcc
2041 ttttcatoga acggttctac tgccgctatc tttgcccgt gggcgctgca ctgccattc
2101 ctggccggat ccgcatgttc gagtggctga agcggtgccc ggaatgcggc tcgccctgcc
2161 agcgtgcgc caaggagtgt ccggtccagt ccatccacc tgagggcgcc atcaatgtca
2221 acgagtgcac ctactgcacg cactgccagg agctttacca cgacgaccag cgctgtccgc
2281 acatgatcca ggtgcggctg aaacgcgaaa agttcatggc gctttccacc ccggcatcgc
2341 ggggggaagc gccggcgaag accgtcgtaa ccacaaagg cgccccgatc agaaaggctg
2401 acgcggcgcc ggaacacccc gtctga

```

APPENDIX D

XAS and EXAFS Data and Analysis of rNosL

The data presented on the following pages were collected, analyzed, and interpreted by Dr. Robert Scott and Mr. Nathaniel Cosper of the University of Georgia. Data for two samples were collected and are presented. The December 1998 data were collected on a 0.6 mM copper-rNosL sample and was considered to be of low quality. The April 1999 data were collected using a 1.1 mM copper-rNosL sample and gave good signal-to-noise throughout the spectral range. Thus, conclusions about the copper ligands have been based on the April 1999 dataset.



CNLOB Fit 11

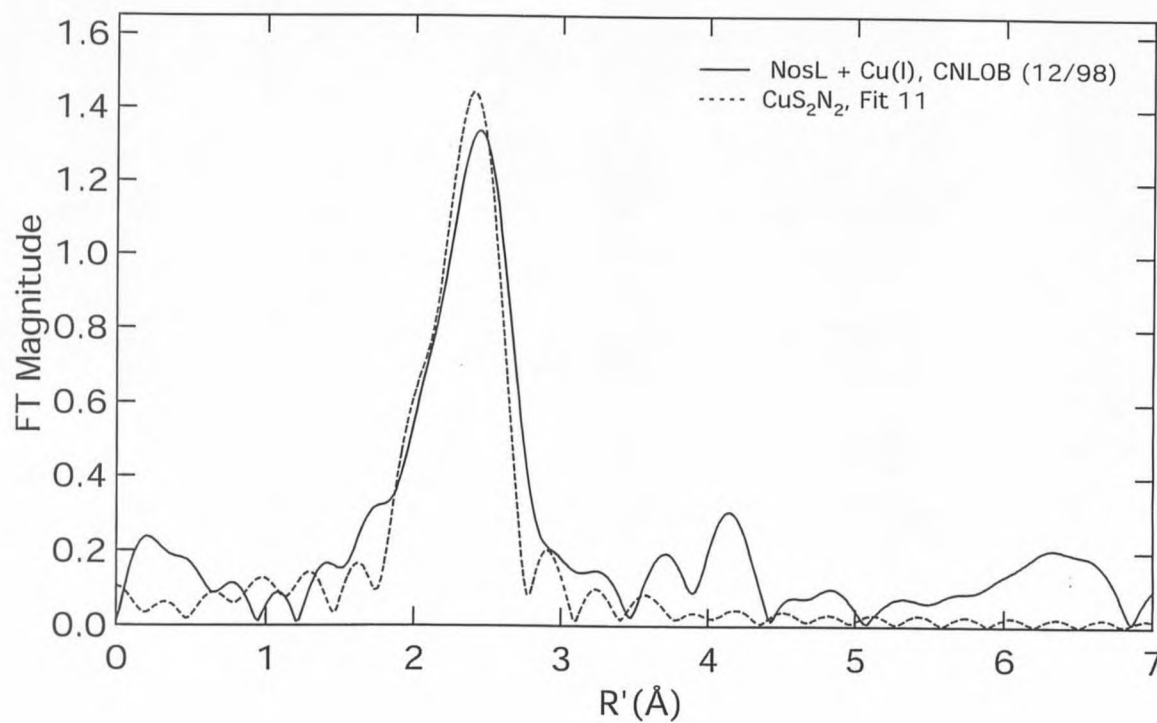
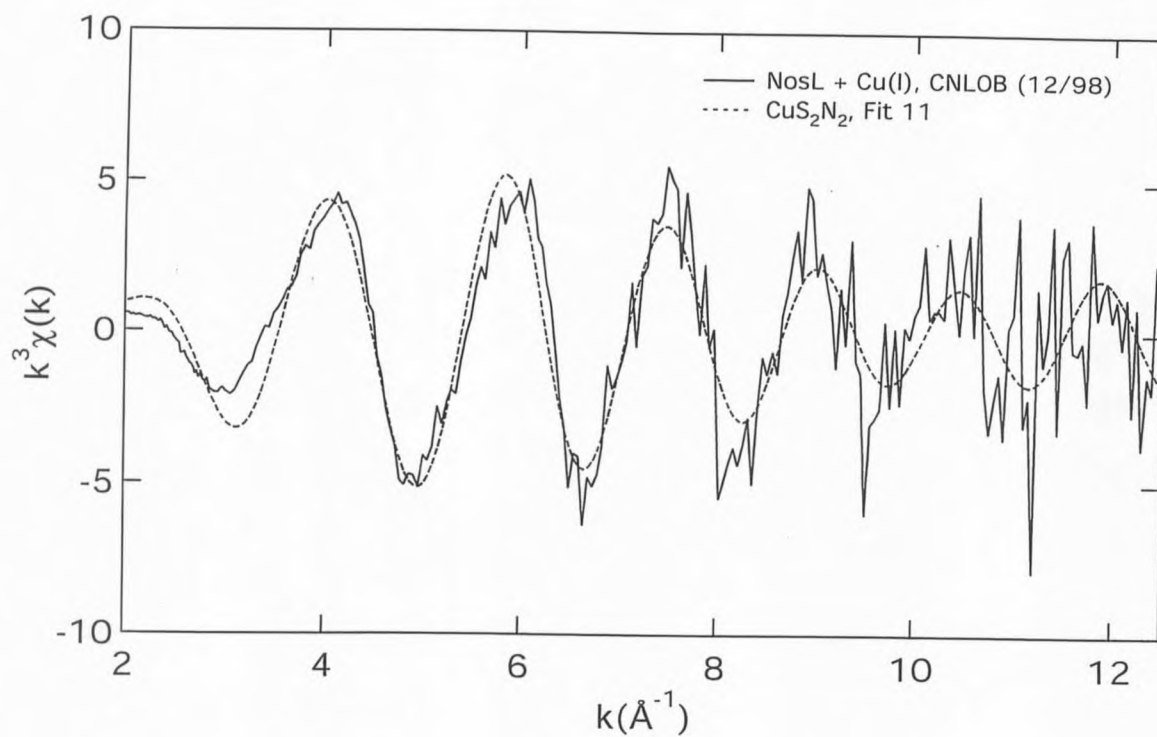


Table 1. Curve fitting results for Cu EXAFS modeled as CuN_xS_y^a

Sample filename (<i>k</i> range) $\Delta k^3 \chi$	Fit	x,y	Shell	N _s	R _{as} (Å)	σ_{as}^2 (Å ²)	ΔE_0 (eV)	f ^b	BVS ^c
NosL + Cu (I)	1	2,0	Cu-N	2	2.18	0.0027	4.51	0.174	
CNLOB (2-12.5 Å ⁻¹)	2	3,0	Cu-N	3	2.19	0.0041	6.05	0.177	
$\Delta k^3 \chi = 12.25$	3	4,0	Cu-N	4	2.18	0.0062	4.95	0.164	
	4	5,0	Cu-N	5	2.16	0.0085	3.54	0.167	
	5	6,0	Cu-N	6	2.15	0.0108	2.59	0.172	
	6	0,2	Cu-S	2	2.34	0.0046	-2.73	0.144	
	7	0,3	Cu-S	3	2.34	0.0070	-3.23	0.135	
	8	0,4	Cu-S	4	2.32	0.0093	-4.66	0.133	
	9	0,5	Cu-S	5	2.32	0.0115	-5.50	0.134	
	10	0,6	Cu-S	6	2.31	0.0136	-6.21	0.139	
	11	2,2	Cu-N	2	2.08	0.0032	-0.65	0.121	
			Cu-S	2	2.37	0.0032	[-0.65] ^d		
	12	2,3	Cu-N	2	2.06	0.0069	-2.19	0.127	
			Cu-S	3	2.35	0.0066	[-2.19]		
	13	3,2	Cu-N	3	2.09	0.0090	-0.34	0.125	
			Cu-S	2	2.37	0.0042	[-0.34]		
	14	3,1	Cu-N	3	2.12	0.0044	1.15	0.123	
			Cu-S	1	2.39	-0.0007	[1.15]		
	15	1,3	Cu-N	1	2.06	-0.0001	-2.76	0.123	
			Cu-S	3	2.35	0.0060	[2.76]		
	16	2,1	Cu-N	2	2.12	0.0010	1.32	0.122	
			Cu-S	1	2.39	-0.0009	[1.32]		
	17	1,2	Cu-N	1	2.08	-0.0014	-1.13	0.121	
			Cu-S	2	2.36	0.0027	[-1.13]		

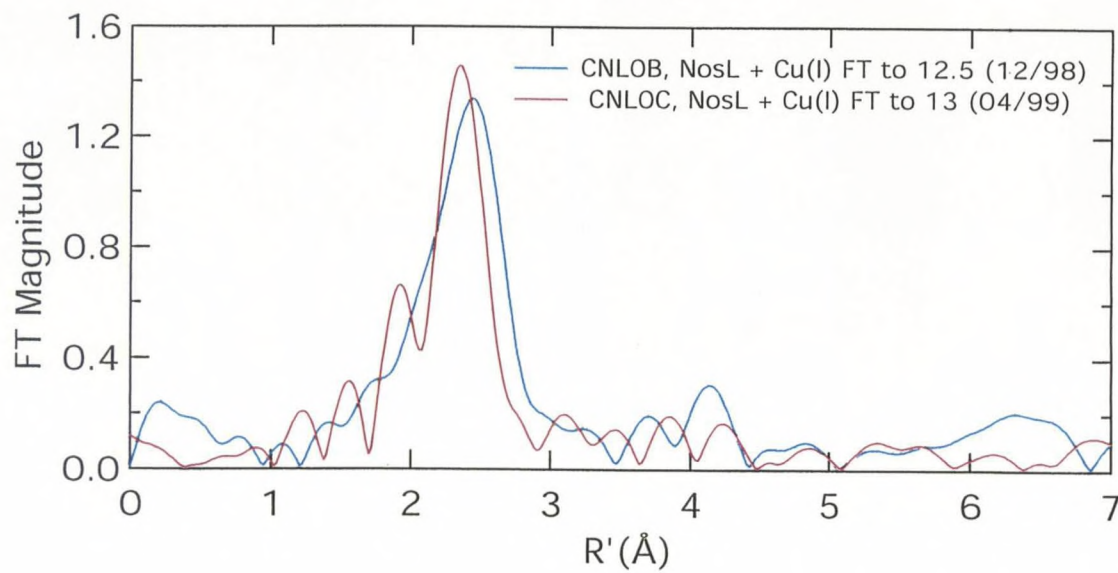
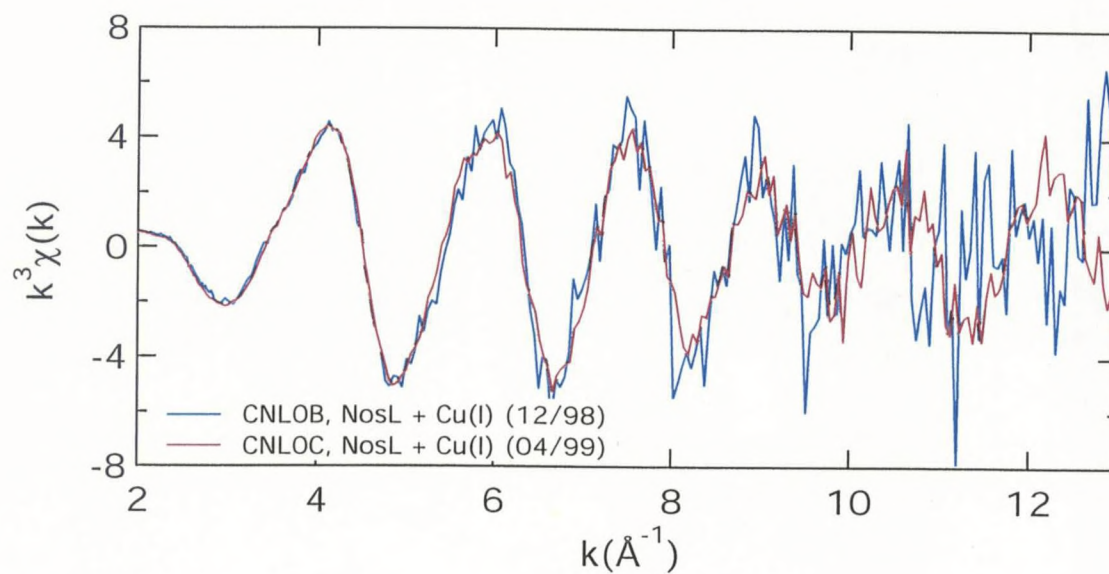
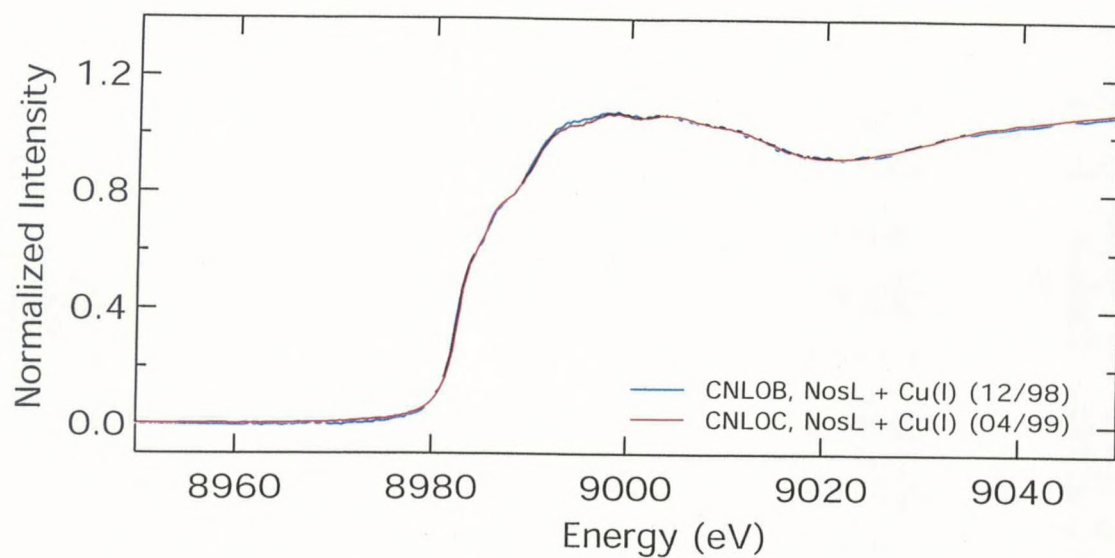
^a Group is the chemical unit defined for the multiple scattering calculation. N_s is the number of scatterers (or groups) per metal. R_{as} is the metal-scatterer distance. σ_{as}^2 is a mean square deviation in R_{as}. ΔE_0 is the shift in E₀ for the theoretical scattering functions.

^b f is a normalized error (chi-squared):

$$f = \frac{\left\{ \sum_i \left[k^3 (\chi_i^{obs} - \chi_i^{calc}) \right]^2 / N \right\}^{1/2}}{\left[(k^3 \chi^{obs})_{max} - (k^3 \chi^{obs})_{min} \right]}$$

^c BVS = $\sum \exp[(r_0 - r)/B]$, B = 0.37, $r_0(\text{Cu(I)-N}) = 1.595$, $r_0(\text{Cu(I)-S}) = 1.898$. (Liu & Thorp, 1993)

^d Numbers in square brackets were constrained to be either a multiple of the above value (σ_{as}^2) or to maintain a constant difference from the above value (R_{as}, ΔE_0).



CNLOC vs Fits17 & 18

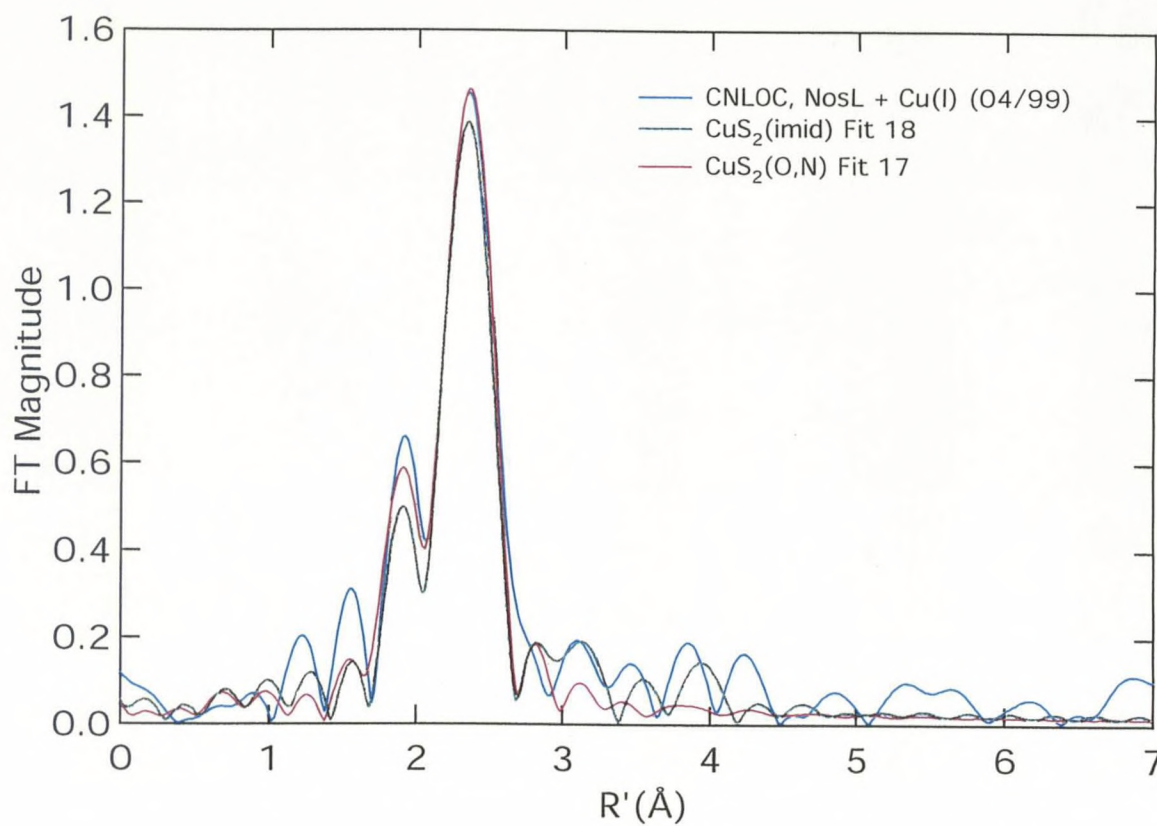
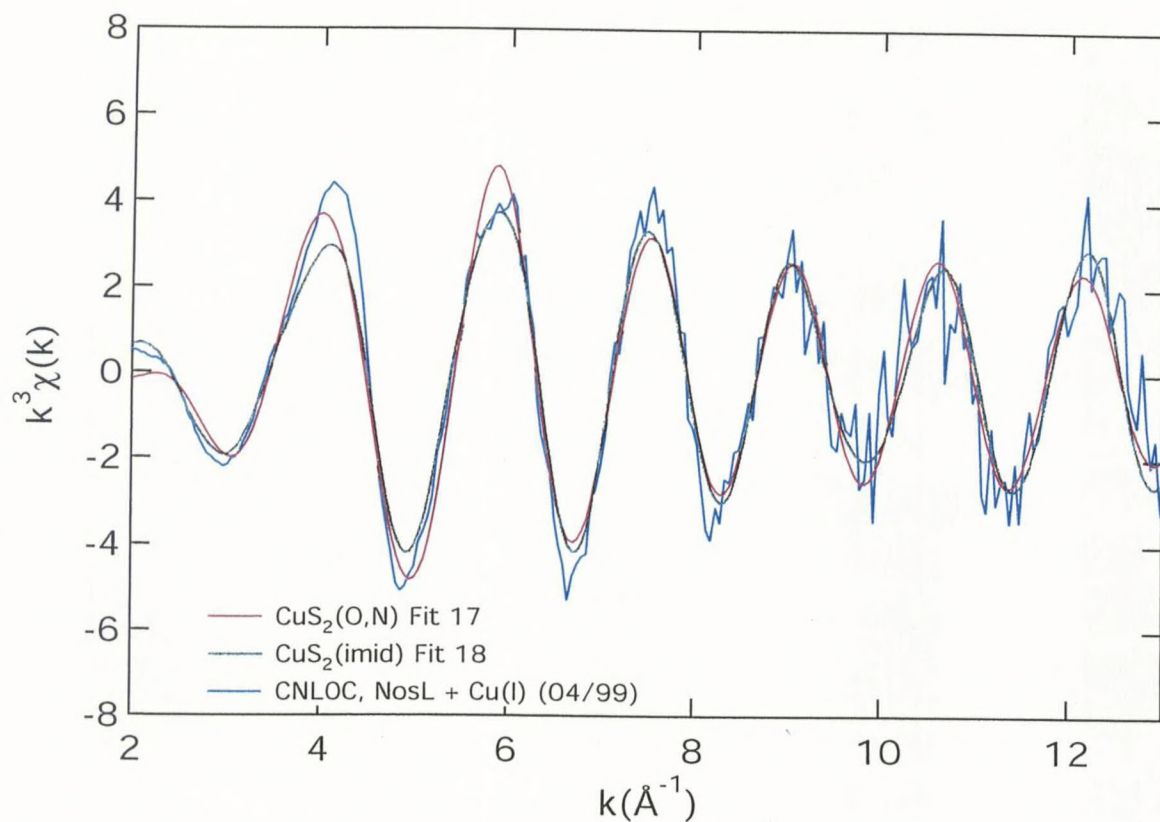


Table 1. Curve fitting results for Cu EXAFS modeled as CuN_xS_y ^a

Sample filename (<i>k</i> range) $\Delta k^3 \chi$	Fit	<i>x,y</i>	Shell	<i>N_s</i>	<i>R_{as}</i> (Å)	σ_{as}^2 (Å ²)	ΔE_0 (eV)	<i>r_b</i>
NosL + Cu (I)	1	2,0	Cu-N	2	2.19	0.0002	9.60	0.152
CNL0C (2-13 Å ⁻¹)	2	3,0	Cu-N	3	2.19	0.0027	19.29	0.153
$\Delta k^3 \chi = 9.55$	3	4,0	Cu-N	4	2.17	0.0063	7.78	0.167
	4	5,0	Cu-N	5	2.13	0.0109	5.38	0.180
	5	6,0	Cu-N	6	2.11	0.0146	3.68	0.187
	6	0,2	Cu-S	2	2.32	0.0033	-4.70	0.109
	7	0,3	Cu-S	3	2.32	0.0059	-5.01	0.101
	8	0,4	Cu-S	4	2.32	0.0083	-5.58	0.107
	9	0,5	Cu-S	5	2.31	0.0107	-6.19	0.117
	10	0,6	Cu-S	6	2.31	0.0131	-6.82	0.131
	11	2,2	Cu-N	2	2.00	0.0095	-3.77	0.079
			Cu-S	2	2.33	0.0032		
	12	2,3	Cu-N	2	2.08	0.0295	-2.74	0.093
			Cu-S	3	2.33	0.0062		
	13	3,2	Cu-N	3	2.07	0.0217	-0.85	0.084
			Cu-S	2	2.34	0.0037		
	14	3,1	Cu-N	3	2.07	0.012	0.41	0.086
			Cu-S	1	2.35	0.0002		
	15	1,3	Cu-N	1	1.95	0.0016	-6.43	0.079
			Cu-S	3	2.322	0.0056		
	16	2,1	Cu-N	2	2.04	0.0074	-1.70	0.090
			Cu-S	1	2.34	-0.0001		
	17	1,2	Cu-N	1	1.97	0.0025	-5.61	0.077
			Cu-S	2	2.32	0.0030		
	18	1,2	Cu-S	2	2.32	0.0031	-5.49	0.077
			Cu-N	1	1.99	0.0025		
			Cu-C	1	2.96	[0.0037]		
			Cu-C	1	2.97	[0.0037]		
			Cu-N	1	4.10	[0.0051]		
			Cu-C	1	4.11	[0.0051]		
	19	2,2	Cu-S	2	2.32	0.0031	-4.58	0.072
			Cu-N	2	2.00	0.0064		

Cu-C	2	2.98	[0.0096]
Cu-C	2	3.01	[0.0097]
Cu-N	2	4.13	[0.0133]
Cu-C	2	4.14	[0.0133]

^a Group is the chemical unit defined for the multiple scattering calculation. N_S is the number of scatterers (or groups) per metal. R_{as} is the metal-scatterer distance. σ_{as}^2 is a mean square deviation in R_{as} . ΔE_0 is the shift in E_0 for the theoretical scattering functions.

^b f is a normalized error (chi-squared):

$$f = \frac{\left[\sum_i \left(k^3 \chi_i^{obs} - \chi_i^{calc} \right)^2 \right]^{1/2}}{\left(k^3 \chi^{obs} \right)_{max} - \left(k^3 \chi^{obs} \right)_{min}}$$

^c $BVS = \sum \exp[(r_0 - r)/B]$, $B = 0.37$, $r_0(\text{Cu(I)-N}) = 1.595$, $r_0(\text{Cu(I)-S}) = 1.898$. (Liu & Thorp, 1993)

^d Numbers in square brackets were constrained to be either a multiple of the above value (σ_{as}^2) or to maintain a constant difference from the above value (R_{as} , ΔE_0).

MONTANA STATE UNIVERSITY - BOZEMAN



3 1762 10332018 8