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
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
(Signature) 9/29/99
Date

Approved for the Department of Chemistry and Biochemistry

Paul A. Grieco
(Signature) 9/25/99
Date

Approved for the College of Graduate Studies

Bruce 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  

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

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Overview of Denitrification</td>
<td>1</td>
</tr>
<tr>
<td>Key Enzymes of Denitrification</td>
<td>4</td>
</tr>
<tr>
<td>Regulation of Denitrification</td>
<td>6</td>
</tr>
<tr>
<td>Structure of Nitrous Oxide Reductase</td>
<td>8</td>
</tr>
<tr>
<td>The nos (Nitrous Oxide Reductase) Gene Cluster</td>
<td>9</td>
</tr>
<tr>
<td>Metallochaperones</td>
<td>12</td>
</tr>
<tr>
<td>Research Goals</td>
<td>13</td>
</tr>
<tr>
<td>SEQUENCE ANALYSIS OF THE nos GENE CLUSTER</td>
<td>15</td>
</tr>
<tr>
<td>Introduction</td>
<td>15</td>
</tr>
<tr>
<td>Experimental Methods</td>
<td>19</td>
</tr>
<tr>
<td>Isolation of nos DNA</td>
<td>19</td>
</tr>
<tr>
<td>DNA Sequencing and Analysis</td>
<td>20</td>
</tr>
<tr>
<td>Protein Characterization and Expression</td>
<td>20</td>
</tr>
<tr>
<td>Molecular Modeling</td>
<td>20</td>
</tr>
<tr>
<td>Results</td>
<td>21</td>
</tr>
<tr>
<td>The nos Gene Cluster</td>
<td>21</td>
</tr>
<tr>
<td>NosR</td>
<td>23</td>
</tr>
<tr>
<td>N₂OR (NosZ)</td>
<td>26</td>
</tr>
<tr>
<td>Model of the Cu₄A Site in N₂OR</td>
<td>29</td>
</tr>
<tr>
<td>NosD, F, Y, and L</td>
<td>32</td>
</tr>
<tr>
<td>NosX</td>
<td>36</td>
</tr>
<tr>
<td>Discussion</td>
<td>37</td>
</tr>
<tr>
<td>Conclusions</td>
<td>45</td>
</tr>
<tr>
<td>CHARACTERIZATION OF NOSL, A NOVEL CU(I) PROTEIN</td>
<td>47</td>
</tr>
<tr>
<td>Introduction</td>
<td>47</td>
</tr>
<tr>
<td>Experimental Methods</td>
<td>51</td>
</tr>
<tr>
<td>Design of Recombinant Constructs</td>
<td>51</td>
</tr>
<tr>
<td>Expression of rNosD</td>
<td>52</td>
</tr>
<tr>
<td>Expression and Purification of rNosL</td>
<td>52</td>
</tr>
<tr>
<td>Analysis of rNosL</td>
<td>54</td>
</tr>
<tr>
<td>Results</td>
<td>55</td>
</tr>
<tr>
<td>Expression Trials of rNosD</td>
<td>55</td>
</tr>
<tr>
<td>Protein Characterization of rNosL</td>
<td>56</td>
</tr>
<tr>
<td>Metal Content of rNosL</td>
<td>57</td>
</tr>
</tbody>
</table>
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 nos Gene Cluster into a Broad Host Range Plasmid ........ 74
Construction of nosL Deletion Plasmids ................................................. 74
Construction of a nosZ Deletion Plasmid ................................................ 75
Construction of nosL Antisense Plasmids ............................................... 75
Introduction of pML122-Based Plasmids into A. cycloclastes .......... 76
Introduction of pML122-Based Plasmids into S. meliloti ....................... 77
GC Assay Of N2O Reductase Activity ..................................................... 77
Purification of N2OR .................................................................................... 78
Interactions of rNosL and apo-N2OR ..................................................... 79
Reduction of N2OR by rNosL ................................................................. 79
Oxidation of Cu(I)-rNosL During the Catalytic Turnover of N2OR .... 80
Results .................................................................................................................. 80
Purification and Preliminary Characterization of Apo- and Holo-N2OR.. 80
The Possible Function of rNosL as an Electron-transfer Protein .......... 80
In Vitro Study of the Possible Function of rNosL as a Copper Chaperone to N2OR ................................................................. 82
In Vivo Studies of the Possible Function of rNosL as a Copper Chaperone to N2OR ................................................................. 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 N2OR .......... 113
Appendix B - Genbank File of the Sequence of the nos Cluster from A. cycloclastes ................................................................. 128
Appendix C - Genbank File of the Sequence of nosR from S. meliloti ......... 135
Appendix D - XAS and EXAFS Data and Analysis for rNosL ............... 137
LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Copper Analysis of rNosL and N$_2$OR</td>
<td>83</td>
</tr>
<tr>
<td>2. Rates of N$_2$O Reduction in Different Strains of <em>S. meliloti</em></td>
<td>84</td>
</tr>
<tr>
<td>3. Rates of N$_2$O reduction at Different Concentrations of Copper in the Growth Media</td>
<td>86</td>
</tr>
</tbody>
</table>
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. The Global Nitrogen Cycle</td>
<td>1</td>
</tr>
<tr>
<td>2. The Denitrification Pathway</td>
<td>3</td>
</tr>
<tr>
<td>3. The Genes of the <em>nos</em> Gene Cluster</td>
<td>9</td>
</tr>
<tr>
<td>4. Possible Function of NosD, F, Y, and L</td>
<td>11</td>
</tr>
<tr>
<td>5. The Arrangement of Genes in the <em>nos</em> Gene Clusters of Four Denitrifiers</td>
<td>21</td>
</tr>
<tr>
<td>6. Comparison of the Deduced Amino Acid Sequences of NosR</td>
<td>24</td>
</tr>
<tr>
<td>7. Sequence Alignment of N$_2$OR</td>
<td>27</td>
</tr>
<tr>
<td>8. Ribbon Diagram of a Model of the Cu$_A$ Domain</td>
<td>30</td>
</tr>
<tr>
<td>9. Surface Map of the N$_2$OR Cu$_A$ Model</td>
<td>30</td>
</tr>
<tr>
<td>10. Close-up of the Modeled Cu$_A$ Site</td>
<td>31</td>
</tr>
<tr>
<td>11. NosD Sequence Alignment</td>
<td>33</td>
</tr>
<tr>
<td>12. Sequence Alignment of NosF</td>
<td>34</td>
</tr>
<tr>
<td>13. Hydropathy Plots of NosY</td>
<td>34</td>
</tr>
<tr>
<td>14. Sequence Alignment of NosL</td>
<td>35</td>
</tr>
<tr>
<td>15. Sequence Alignment of NosX</td>
<td>36</td>
</tr>
<tr>
<td>16. The pET20b+ Expression Vector</td>
<td>49</td>
</tr>
<tr>
<td>17. The Lipobox Consensus Sequence of SpaseII</td>
<td>50</td>
</tr>
<tr>
<td>18. Chromatogram of the Purification of rNosL</td>
<td>53</td>
</tr>
<tr>
<td>19. SDS/PAGE Showing rNosD Expression Over Time</td>
<td>55</td>
</tr>
</tbody>
</table>
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 A. cycloclastes 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 N$_2$OR, 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 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.

*Figure 1.* The global nitrogen cycle.
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 N₂ (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 O₂. 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:

\[
\begin{align*}
\text{NO}_3^- & \rightarrow \text{NO}_2^- \rightarrow \text{NO} \rightarrow \text{N}_2\text{O} \rightarrow \text{N}_2
\end{align*}
\]  

The electrons for the reaction are provided by small, soluble electron transfer proteins, which are in turn supplied by a membrane-bound bc₁ complex (5). In the bc₁ 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). N₂O has a remarkably stable half-life of about 150 years (7,8). Studies show that N₂O reaches the stratosphere, where it is oxidized to NO by reaction with singlet oxygen, which is produced
Figure 2. The Denitrification Pathway. ▸ represents electron transfers between proteins. ▶ 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 O₃ (9). NO continues reacting with O₃, 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 N₂O pollution. In addition to the N₂O produced from biological sources (1,8), significant amounts of N₂O are
released as a waste product during industrial nylon production (13) and biomass burning (14). Studying the biological mechanism of N₂O reduction to N₂ 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 (cdNiR) (25-30). While denitrifiers contain either CuNiR or cdNiR, but not both, it has been shown that CuNiR can be functionally produced in a cdNiR 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 nitroxyanion (NO$^-$), which is then non-enzymatically converted to N$_2$O (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 (N2OR) is the terminal enzyme of denitrification and reduces N2O by two electrons to N2. Although N2ORs are typically soluble, periplasmic enzymes (39-44), at least one may be membrane-associated (45). To date, all N2ORs have been found to be homodimers that contain two distinct dinuclear copper centers per subunit. The electron transfer copper site, called CuA, is analogous to the CuA site in cytochrome c oxidase (46-52), and the catalytic copper site, called CuCat, is where N2O 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 N2OR (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 O2, 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 on 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 N$_2$OR, 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 N$_2$OR, Glu is replace 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$_A$
The authors postulate that the CuCat 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 N\(_2\)OR by NosD and NosL via the integral membrane protein NosY (Figure 4). NosD has been suggested to be a chaperone protein that keeps apo-N\(_2\)OR 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 \textit{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 \textit{in vivo} or \textit{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 \textit{et al} were able to show that CCS is required for copper incorporation \textit{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 the the cdNiR 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 N$_2$OR 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 CLUSTER

Introduction

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 \(\alpha\)-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 ferridoxins, 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\(_2\)OR 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 N₂OR. 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, N₂OR was the only Nos protein that had been isolated and characterized. Biophysical investigations revealed that N₂OR is a homodimer containing two dinuclear copper centers (39-44). One copper site, Cuₐ, 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 CuA 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 N2ORs and COXs proved especially insightful. Data will be presented that identify probable ligands to CuA in N2OR. Moreover, the homology studies suggest that a His residue located in the CuA domain may be a ligand to the catalytic copper center. A model of the CuA domain in N2OR based on the COX structure will be presented. While the overall sequence identity between the CuA domains of COX and N2OR is low, two factors indicate that homology modeling is valid in this case. First, a comparison of the COX CuA 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 CuA 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 N2OR. 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 CuA from COX and N2OR are very similar, it is plausible that the protein environment of this dinuclear center in N2OR will closely resemble that of COX.
Experimental Methods

Isolation of nos DNA. A 3 kbp EcoRI - SphI 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 EcoRI, SphI, Aval, PstI, or HincII 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 EcoRI fragment containing the nos cluster from genomic DNA, using standard Southern and colony hybridization techniques. The EcoRI fragment was subsequently cloned into pUC19 to produce plasmid pEE19. E. coli strain INVαF' (Invitrogen) was used as the host for all pUC19-derived plasmids and cultures were grown in LB medium containing either 60 μg/mL ampicillin or 100 μg/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 ([γ^32-P]ATP) with either the fmol or the Silver Sequencing kit (Promega), as well as the label-incorporation method ([35S]dCTPαS) 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 N2OR 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 N2OR 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 Cuₐ 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.](image-url)
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 *nosRZDFYLYX* 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 (TTGATN$_4$ATCAA) 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 1/2-boxes (ATCAA) in the promoter region of *nosZ* in *A. cycloclastes*. In *S. meliloti* there is a
degenerate $\sigma^{54}$ binding motif (CTGNNAN$_{1,6}$TGCAA) upstream of the nosR initiation codon, although its associated transcriptional start site falls within an FNR box. Degenerate $\sigma^{54}$ 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.
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 CXjCP 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 \([Fe_4S_4]\)-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).

\(N_2OR\) (NosZ). The alignment of the deduced amino acid sequence of \(N_2OR\) from \(A.\) cycloclastes with six published sequences is shown in Figure 7. \(N_2ORs\) 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 \(N_2ORs\) range from 65,000 to 67,350 Da per monomer, with an average pI of 5.9.

\(N_2ORs\) 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\(_{Cal}\) 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
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*. CuA ligands are underlined; potential CuC ligands are in bold, with a possible semi-conserved Cys indicated by V.

An alignment of the consensus sequence (C) CuA 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$_{\text{cat}}$ 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$_{\text{A}}$ domain identified in other N$_2$ORs (79,80).

Model of the CuA Site in N$_2$OR. A model of the C-terminal domain of A. cycloclastes N$_2$OR, based on the crystal structure of the Cu$_{\text{A}}$ 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 N$_2$ORs (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$_{\text{A}}$ ligand in cytochrome oxidase, this amino acid is replaced by Trp (His in R. eutropha) in N$_2$ORs. 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$_{\text{A}}$ from solvent (51,52). A surface patch of negatively charged residues (Figure 8, white) is present in both cytochrome oxidase and N$_2$OR: 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 N$_2$ORs but not cytochrome oxidases and are potential Cu$_{\text{cat}}$ ligands. A surface map of the model is also shown in Figure 9.
Figure 8. Ribbon diagram showing a model of the Cu₄ domain in N₂OR (dark) based on the Cu₄ domain of COX Subunit II (light) as template.

Figure 9. A surface map of the N₂OR Cu₄ model, showing the spatial relationship of the purported Cyt c docking site (negatively charged patch) to the two Cu₄ ligands H621 and H664, and to the proposed Cu_{Cat} ligand H661.
A closeup of the modeled Cu\textsubscript{A} site in N\textsubscript{2}OR is shown in Figure 10. The Cu\textsubscript{A} ligands (black) are H621, C656, the peptide carbonyl of W658, C660, H664, and M667. Examination of the positions of the potential Cu\textsubscript{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 \textit{trans} to that of C660 and its ring nitrogens are within 9.5 Å of Cu\textsubscript{A}. In cytochrome oxidase, the site corresponding to the position of H661 is located at the interface of the Cu\textsubscript{A} domain of Subunit II with Subunit I, which contains the catalytic, Fe-Cu metallocenter.

**Figure 10.** Closeup of the modeled Cu\textsubscript{A} site. White, surface Asp residues; black, proposed Cu\textsubscript{A} ligands; spheres, copper ions. The position of COX subunit I (surface dots) with Cu\textsubscript{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
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.
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.
Figure 14. Sequence alignment of NosL from *A. cycloclastes* (AC), *R. sphaeroides* (RS), *B. japonicum* (BJ), *S. meliloti* (SM), and *P. stutzeri* (PS). C represents the conserved amino acids. An arrow \( \downarrow \) marks the predicted start of mature *A. cycloclastes* NosL. RL represents the sequence of the mature *A. cycloclastes* recombinant NosL protein described in the third chapter. Potential copper ligands as predicted from EXAFS (His, Cys, and Met) are in bold.
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.
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^5^4 (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-CX_3-CP-helix-helix-CX_5-CP-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 N_2_ORs (NosZ) has been confirmed
for P. stutzeri (118) and now A. cycloclastes. This presequence contains the double Arg motif
that is found in all N_2_ORs 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 Cuₐ 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 Cuₐ site is highly conserved between N₂OR and cytochrome oxidase (126). Hence, a model of the N₂OR C-terminal Cuₐ domain was constructed, based on the crystallographically-determined Cuₐ site of bovine heart cytochrome oxidase (51). The model suggests that of the two potential Cuₐ ligands located in the Cuₐ domain, only H661 is available for metal binding; the other, M671, is buried within the protein. If H661 is a Cuₐ 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 Cuₐ 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 Cuₐ 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 N₃OR 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 Cu₄ site ligand.

An alternative explanation is possible, however, and is precedented by mutational studies on the Cys ligands of a Fe₂-S₂ ferridoxin (129). In these experiments, the site-directed mutagenesis (Cys→Ser) of a Cys ligand in *Azotobacter vinelandii* ferridoxin 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 N₃OR may explain the observation that mutagenesis did not appear to significantly perturb the Cu₄ site. If, in the absence of C190, another Cys became coordinated to Cu₄, 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* N₃OR. If the His residues of this sequence are Cu₄ 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\textsubscript{2}OR (130). The model is based on EPR and MCD quantitation of the N\textsubscript{2}OR 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\textsubscript{A} center (called Cu\textsubscript{2}). 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 form one site to another. Elucidating the nature of Cu\textsubscript{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\textsubscript{2}OR. 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_{Cax} biogenesis.

The role of NosL in formation of the Cu_{Cax} 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_{Cax} 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, \textit{nosX} resides at another location. Examples of this are found for several denitrification genes. For instance, in the denitrifier \textit{R. eutropha}, \textit{nosZ} is located on a megaplasmid but is not flanked by either \textit{nosR} or \textit{nosD} (79). Moreover, the nitrite reductase (\textit{nir}) genes of this denitrifier are evidently partitioned between megaplasmid and chromosomal DNA (132).

**Conclusions**

The comparison of \textit{nos} sequences from different organisms has confirmed several initial observations made for \textit{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$_4$S$_4$ 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 \textit{S. meliloti} and the recognition of \textit{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 N_2OR provides insight into the relationship of protein structure and function. For example, the conservation of an acidic surface patch in both N_2OR 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 N_2OR could be informative for other metalloproteins, given the emergence of metallochaperone chemistry as a general phenomenon. Since copper incorporation into N_2OR 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 N\textsubscript{2}OR 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 \textit{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 \textit{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 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.
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 KPO₄ 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 as hosts for pD21, but these showed no improvement in expression over the original strains. The addition of metals (10 μM 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 Cu(NO₃)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 ± 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 ($\Delta\varepsilon_m$) of a protein sample with UV absorbance $\Delta\varepsilon_{278} = 0.905$ was measured by MCD. Using the published value for Trp of $\Delta\varepsilon_m = 777$ deg cm$^2$ decimol$^{-1}$ 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, $\varepsilon_{278}$ was calculated to be 10.8 cm$^{-1}$ for a 1% solution, or 20.0 mM$^{-1}$ 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 mM⁻¹ 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 H₂O₂, 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-diethylthiodithiocarbamate (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 tryprophan) 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 μM rNosL with the Cys reagent PDS. A, apomonomer; 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 Histagged 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 CuS$_2$imidazole and CuS$_2$(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 resolublize 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 ~ 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 Cu(I)S2(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 H2O2 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 N\textsubscript{2}OR and NosD (59). Hence, NosL would not be strictly required for copper incorporation.

The finding that rNosL selectively binds Cu(I) (\textit{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 N\textsubscript{2}OR. On the other hand, the acidic nature of NosL (pI ~ 4) argues against this role. The electron-transfer docking site in N\textsubscript{2}OR 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 N\textsubscript{2}OR.

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 N\textsubscript{2}OR. Yet the insertional mutation study in \textit{P. stutzeri} (98) suggests that NosL is not necessary for the production of active N\textsubscript{2}OR. However, in that experiment the insertion occurred late in the \textit{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 \textit{nosL} in the heterologous \textit{S. melliloti} 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 *SbfI*. This 9.5 kb fragment was inserted into the *SbfI* site of the broad host range plasmid pMLA to create plasmid pMLnos. Plasmid pMLA was created by removing a *SalI* fragment from pML122 (133). The 9.5 kb *SbfI* fragment containing the *nos* cluster was also ligated into the *PstI* site of pUC19 to produce plasmid pCluster.

Construction of *nosL* Deletion Plasmids. Plasmid pCluster was digested with *BglII*, 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 *Bal1* 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'GAGATAGTCTTCGTCGTCCGC). Primer D7883L contains the stop codon for *nosY*, which overlaps the *nosL* initiation codon. Primer U527L matches *nosL* sequence downstream of the *BglII* site.
Consequently, any plasmids that yield PCR products with these primers should contain the entire *nosY* gene and any upstream *nosX* regulatory elements. Plasmid pClusLA11 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 pClusLA11. 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 pClusLA11 with *SbfI* and spliced into pMLA to produce plasmid pMLLA11.

**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 pClusFseA. The mutated *nos* cluster was cut from pClusFseA with *SbfI* and spliced into pMLA to produce plasmid pMLFseA.

**Construction of *nosL* Antisense Plasmids.** A transcriptional terminator cassette (called oligoTT) was assembled by annealing oligos TTXF (5’TCGAGCTGATTATCGAT-TCCCATGCTGCTGACCCGACCGCTTGAATTCTTGCAATTTCCTGCAGTTCTCGCAA) and TTCR (5’CGTTGCGAGAACTGCAGAAATTGCAATCAAGCGGCGTCGG-GTCAGCAGCATGGGGGAAATCGATAATCAGC). This cassette contains a 5’ *XhoI* overhang (*XhoI* site is retained after ligation) followed by a spacer region and a new *ClaI*
site. Immediately downstream of the Clal site is the nosL transcriptional terminator sequence (nucleotides 8001 – 8053, Appendix B) followed by a Clal-compatible overhang (this destroys the original Clal site after ligation). The oligoTT cassette was ligated into pML122 prepared by digestion with XhoI and Clal. 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’ Clal site into the PCR product. After restriction digest with XhoI and Clal, 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$_{500}$ = 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 μF/1800V and 150 ohms. Approximately 200 – 400 ng plasmid DNA in 1 – 2 μL was mixed with 50 μL 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 μL cell suspension was plated on LB-agarose containing 50 μg/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 μL 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. N₂O standards were prepared by injecting 50, 100, 150, 200, or 250 μL N₂O into vials containing 1 mL of water. Standard curves were constructed at every time point to account for the small, non-enzymatic loss of N₂O.

For A. cycloclastes strains, cultures were grown at 30°C and 240 rpm to OD₆₀₀ = 1.0 in LB media containing 50 μg/mL gentamycin. For RNA interference experiments, 50 μg/mL kanamycin was also added (A. cycloclastes is KnR (147)). The induction of nos genes was accomplished by adding 10 mM KNO₃, Cu(NO₃)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 OD₆₀₀ = 0.5 in TYC media (81) containing 50 μg/mL gentamycin. After the addition of 10 mM KNO₃ and any required Cu(NO₃)₂, cultures were grown for 10 – 24 hours prior to GC analysis.

Purification of N₂OR. 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 N₂OR, 50 μM Cu(NO₃)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 µL sample of 0.174 mM apo-N2OR was anaerobically added to 400 µL 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 CuA center of N2OR. One mL of 22.5 µM N2OR (45 µM CuA 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 μM) 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 μM 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 μmol/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 Cuₐ 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 Cuₐ 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$_2$OR that displays a single broad absorption band at 650 nm. No changes in the visible spectrum of N$_2$OR were observed when rNosL was added.

![Visible absorption spectrum of A. cycloclastes N$_2$OR (30 mg/mL) in 10 mM MES pH 6.5 + 0.2 M NaCl. Path length, 1 cm.](image)

**Figure 26.** Visible absorption spectrum of *A. cycloclastes* N$_2$OR (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$_2$OR, the EPR spectrum of 0.4 mM Cu(I)-rNosL was monitored before and after the addition of a catalytic amount of N$_2$OR in an N$_2$O atmosphere (Figure 27). Reduced benzyl viologen, which serves as an electron donor in the spectrophotometric assay for N$_2$OR (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$_2$O (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 N₂OR and does not function as an electron donor to the enzyme.

**Figure 27.** EPR spectra of 0.4 mM rNosL and 2 μM N₂OR before (A) and after (B) the addition of N₂O. 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 N₂OR. The possibility that apo-N₂OR and rNosL would form a complex was investigated. A 450 μL mixture of 0.52 mM rNosL and 0.019 mM apo-N₂OR 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 N₂OR. 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**

<table>
<thead>
<tr>
<th>Sample</th>
<th>Copper:Protein Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>rNosL as isolated</td>
<td>0.9</td>
</tr>
<tr>
<td>apo-N2OR as isolated</td>
<td>2.3</td>
</tr>
<tr>
<td>fraction 17 (N2OR)</td>
<td>2.1</td>
</tr>
<tr>
<td>fraction 23 (rNosL)</td>
<td>0.9</td>
</tr>
</tbody>
</table>
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 pMLΔ11 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*

<table>
<thead>
<tr>
<th>Strain</th>
<th>Rate&lt;sup&gt;a,b&lt;/sup&gt;</th>
<th>Rate&lt;sup&gt;a,b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>A. cycloclastes</em></td>
<td>2.6 ± 0.2</td>
<td>2.8 ± 0.3</td>
</tr>
<tr>
<td>AC/pMLnos</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2160</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2160/pMLΔ</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2160/pMLFseΔ</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2160/pMLnos</td>
<td>0.3 ± 0.05</td>
<td>0.3 ± 0.05</td>
</tr>
<tr>
<td>RMYC2160/pMLΔ11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2164</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2164/pMLFseΔ</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMYC2164/pMLnos</td>
<td>0.7 ± 0.2</td>
<td>0.7 ± 0.2</td>
</tr>
<tr>
<td>RMYC2164/pMLΔ11</td>
<td>0.7 ± 0.2</td>
<td>0.7 ± 0.2</td>
</tr>
</tbody>
</table>

<sup>a</sup> µmol per hour per mL culture. <sup>b</sup> 10 µM Cu<sup>II</sup> added to growth media.
Strains RMYC2160/pMLnos and RMYC2164/pMLnos displayed a slow loss of N\textsubscript{2}OR activity when grown in continuous culture. Therefore, fresh cultures were prepared prior to each GC experiment when the phenotypes of strains containing pML\textalpha{}11 were evaluated. When 10 \textmu{}M copper was included in the growth media, no differences in the rates of N\textsubscript{2}O reduction were noted between strains RMYC2164/pMLnos and RMYC2164/pML\textalpha{}11 or between strains RMYC2160/pMLnos and RMYC2160/pML\textalpha{}11. Identical rates were obtained when the experiments were repeated with media supplemented with either 5 or 100 \textmu{}M copper.

The phenotype experiments in \textit{S. meliloti} indicate that NosL is not necessary for the bioassembly of active N\textsubscript{2}OR 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 N\textsubscript{2}OR produced to the available copper supply. Considering the low levels of N\textsubscript{2}OR that are produced in the \textit{S. meliloti} strains, it was decided to abandon heterologous expression in lieu of homologous RNA interference trials.

Comparative experiments measuring the rates of N\textsubscript{2}O reduction were performed between strains AC/pMLnos and AC/pTTalnos and strains AC/pML\textalpha{} and AC/pTTal. Copper concentrations in the growth media were varied from 0, 1, 2, 3, 4, 5, 10, 20, and 50 \textmu{}M, depending on the experiment. The relative amounts of N\textsubscript{2}OR 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 N\textsubscript{2}OR activity on copper content were noted between
similar strains with or without the *antisense*NosL coding sequence.

Table 3. Rates\(^{a,b}\) of N\(_2\)O reduction at different concentrations of copper in the growth media.

<table>
<thead>
<tr>
<th></th>
<th>0 (\mu)M Cu(^{II})</th>
<th>2 (\mu)M Cu(^{II})</th>
<th>3 (\mu)M Cu(^{II})</th>
<th>4 (\mu)M Cu(^{II})</th>
<th>5 (\mu)M Cu(^{II})</th>
<th>20 (\mu)M Cu(^{II})</th>
</tr>
</thead>
<tbody>
<tr>
<td>AC/pTTalnos</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.2</td>
<td>0.3</td>
<td>2.2</td>
</tr>
<tr>
<td>AC/pMLnos</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.40</td>
<td>0.3</td>
<td>2.2</td>
</tr>
</tbody>
</table>

\(^a\)\(\mu\)mol per hour per mL culture. \(^b\)Error in rates was estimated to be 20% for rates < 1.0 and 10% for rates ≥ 1.0.

**Discussion**

The purification of *A. cycloclastes* N\(_2\)OR is a significant step toward understanding the mechanism of the bioassembly of its active-site copper center. Yields of pure N\(_2\)OR from the over-expressing strain AC/pMLnos ranged from 15 – 40 mg per liter of cell culture. When isolated from copper-supplemented cultures, N\(_2\)OR was active and fully metallated (8 coppers per protein dimer). In contrast, the isolation of N\(_2\)OR from unsupplemented cultures produced a protein that only contained 25% copper. This enzyme is useful for studying potential interactions and copper transfer reactions between N\(_2\)OR and rNosL.

It is intriguing to speculate on the location of copper in this metal-depleted form (referred to as apo-N\(_2\)OR). One possibility is that the isolated protein is a mixture of 25% fully-loaded copper N\(_2\)OR and 75% fully-depleted N\(_2\)OR. However, the featureless visible absorbance spectrum of apo-N\(_2\)OR 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 N2OR would be anticipated. The isolation of a Cu_A-free, Cu_Cat-loaded form of N2OR would tremendously advance the elucidation of N2OR 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 N2O reduction in the GC assay. This was surprising, as AC/pMLnos produces 10 – 100X more active N2OR per mL culture than its parent strain. Together, the results indicate that biological N2O 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 N2OR was investigated. First, the ability of rNosL to reduce the Cu_A center of N2OR was tested. No reduction was seen, as evidenced by the lack of change in the N2OR visible spectrum. Next, the possibility that substrate binding might trigger electron transfer between rNosL and N2OR was investigated. The large amount of rNosL that would be required to sustain measurable turnover in the N2OR 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 N2OR. 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\textsubscript{2}OR 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\textsubscript{2}OR when large excesses of substrate were added. This indicates that rNosL is incapable of functioning as an electron-transfer partner to N\textsubscript{2}OR.

The potential role of rNosL as a copper chaperone to N\textsubscript{2}OR was investigated. Incubation of excess rNosL with apo-N\textsubscript{2}OR 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 \textit{in vitro} copper transfer had occurred.

The functional expression of \textit{A. cyclodastes} nos genes carried on a pML-based plasmid in \textit{S. meliloti} has been demonstrated, but only at low levels. The pML family of vectors is known to achieve high copy numbers in \textit{S. meliloti}. Therefore, the low N\textsubscript{2}OR expression levels observed for strains RMYC2160/pMLnos and RMYC2164/pMLnos suggest that the \textit{A. cyclodastes} regulatory elements are not well-recognized by \textit{S. meliloti} transcription factors. Furthermore, a gradual loss of N\textsubscript{2}OR 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 \textit{A. cyclodastes} 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\textsubscript{2}O reduction in \textit{S. meliloti} strains carrying pMLnos and strains carrying a \textit{nosL} deletion (pML\textsubscript{L}A11) were compared (Table 2). No differences in N\textsubscript{2}OR activity were noted. These data clearly indicate that NosL is not required for copper site assembly in N\textsubscript{2}OR. 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 N$_2$OR, 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 N$_2$OR 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 N$_2$OR 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 N$_2$OR. 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 N$_2$OR 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 N$_2$OR. 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 N$_2$OR. 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. N$_2$OR 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 N$_2$OR. 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 N$_2$OR. 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 N$_2$OR, 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 N_{2}OR, 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 N_{2}OR 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


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.


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.


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.


94. Reijnders WNM. *Paracoccus denitrificans nirI* and *nirX* genes, and flanking regions. 1998: Genbank Accession #AJ001308.


116. Roe BA, Lin SP, Song L, Yuan X, Clifton S, Dyer DW. Gonococal Genome Sequencing Project, the University of Oklahoma. 1998.


121. Teidje JM, *personal communication*.

122. Dooley DM, McDermott T. *unpublished observations*.


147. Bollinger JM, *personal communication*.


152. Dooley DM, Henery SH. *unpublished observations*.

APPENDIX A

Comparison of 50 Amino Acid Sequences of N$_2$OR

Amino Acid Lineups of all available NosZ (N$_2$OR) 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.

<table>
<thead>
<tr>
<th>Genbank #</th>
<th>Organism</th>
<th>Genbank #</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF119953</td>
<td>Uncultured marine bacterium</td>
<td>AF119925</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>*AF016059</td>
<td>Uncultured marine bacterium</td>
<td>AF119948</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>Pseudomonas aeruginosa</td>
<td>AF002531</td>
<td>Bradyrhizobium japonicum</td>
</tr>
<tr>
<td>M22628</td>
<td>Pseudomonas stutzeri</td>
<td>Y47133</td>
<td>Sinorhizobium meliloti</td>
</tr>
<tr>
<td>AF056319</td>
<td>Pseudomonas fluorescens</td>
<td>AF047429</td>
<td>Achromobacter cycloclastes</td>
</tr>
<tr>
<td>AF119942</td>
<td>Uncultured marine bacterium</td>
<td>X74792</td>
<td>Rhodobacter sphaeroides</td>
</tr>
<tr>
<td>AF119940</td>
<td>Uncultured marine bacterium</td>
<td>AF125260</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119947</td>
<td>Uncultured marine bacterium</td>
<td>AF016055</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119944</td>
<td>Uncultured marine bacterium</td>
<td>AF016056</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119952</td>
<td>Uncultured marine bacterium</td>
<td>AF119927</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119941</td>
<td>Uncultured marine bacterium</td>
<td>AF119926</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119936</td>
<td>Uncultured marine bacterium</td>
<td>AF119922</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119949</td>
<td>Uncultured marine bacterium</td>
<td>AF119921</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF016057</td>
<td>Uncultured marine bacterium</td>
<td>AF119920</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119945</td>
<td>Uncultured marine bacterium</td>
<td>AF119933</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119938</td>
<td>Uncultured marine bacterium</td>
<td>AF119930</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119943</td>
<td>Uncultured marine bacterium</td>
<td>AF119928</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119950</td>
<td>Uncultured marine bacterium</td>
<td>AF119932</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119954</td>
<td>Uncultured marine bacterium</td>
<td>AF119931</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119939</td>
<td>Uncultured marine bacterium</td>
<td>AF119934</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119935</td>
<td>Uncultured marine bacterium</td>
<td>AF119929</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119918</td>
<td>Uncultured marine bacterium</td>
<td>AF119919</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119951</td>
<td>Uncultured marine bacterium</td>
<td>AF119924</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119946</td>
<td>Uncultured marine bacterium</td>
<td>AF119923</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF119937</td>
<td>Uncultured marine bacterium</td>
<td>X65278</td>
<td>Ralstonia eutropha</td>
</tr>
</tbody>
</table>

1 50

<table>
<thead>
<tr>
<th>Amino Acid Lineups</th>
<th>Accession Numbers</th>
<th>Organism</th>
</tr>
</thead>
<tbody>
<tr>
<td>X65277</td>
<td>AF119953</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>M22628</td>
<td>AF016059</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119942</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119940</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119947</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119944</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119952</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119941</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119936</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119949</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119945</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119938</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119943</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119950</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119954</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119939</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119935</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119918</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119951</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119946</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>X65277</td>
<td>AF119937</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119953</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF016059</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119942</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119940</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119947</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119944</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119952</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>AF056319</td>
<td>AF119941</td>
<td>Uncultured marine bacterium</td>
</tr>
<tr>
<td>Accession</td>
<td>Sequence</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
<td></td>
</tr>
<tr>
<td>AF119936</td>
<td>MSDSDNIKGV SRRTLLGTTA . AAAGVGLA GGAIVTKDGA GFVSTADAQT</td>
<td></td>
</tr>
<tr>
<td>AF119949</td>
<td>MSNEETKMLT RRQMLGTTA FMAAGAVGA GGALTLSGTA A TPARAFQ</td>
<td></td>
</tr>
<tr>
<td>AFO16057</td>
<td>MESKE HKGLSRRALF SATAGSAILA G TVG PAALGLGAAG</td>
<td></td>
</tr>
<tr>
<td>AF119945</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119938</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119943</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119950</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119954</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119939</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119935</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119918</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119946</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119937</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119925</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119948</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AJ002531</td>
<td>MSNEETKMLT RRQMLGTTA FMAAGAVGA GGALTLSGTA A TPARAFQ</td>
<td></td>
</tr>
<tr>
<td>U47133</td>
<td>MSNEETKMLT RRQMLGTTA FMAAGAVGA GGALTLSGTA A TPARAFQ</td>
<td></td>
</tr>
<tr>
<td>AFO47429</td>
<td>MESKE HKGLSRRALF SATAGSAILA G TVG PAALGLGAAG</td>
<td></td>
</tr>
<tr>
<td>X774792</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF125260</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AFO16055</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF016056</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119926</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119922</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119921</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119920</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119933</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119930</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119928</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119932</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119931</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119934</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119929</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119919</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119924</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>AF119923</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>X65278</td>
<td>MESKEKASI GNGPGGIGRR QFLEGTAAALG LAVVACTDK GAAPAAAAG</td>
<td></td>
</tr>
<tr>
<td>Consensus</td>
<td>MSDSDNIKGV SRRTLLGTTA . AAAGVGLA GGAIVTKDGA GFVSTADAQT</td>
<td></td>
</tr>
</tbody>
</table>

Consensus:

<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF119953</td>
<td>MTRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF016059</td>
<td>MTRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>X65277</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>M22628</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF056319</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119942</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119940</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119947</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119944</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119952</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119941</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>AF119936</td>
<td>MRRDSWAAAA KEAQQKIHVA PGEDEYGYGF WSGGHQGVR VLGVPSMREL</td>
</tr>
<tr>
<td>Accession</td>
<td>Sequence</td>
</tr>
<tr>
<td>-----------</td>
<td>----------</td>
</tr>
<tr>
<td>AF119949</td>
<td></td>
</tr>
<tr>
<td>AF016057</td>
<td></td>
</tr>
<tr>
<td>AF119945</td>
<td></td>
</tr>
<tr>
<td>AF119938</td>
<td></td>
</tr>
<tr>
<td>AF119943</td>
<td></td>
</tr>
<tr>
<td>AF119950</td>
<td></td>
</tr>
<tr>
<td>AF119954</td>
<td></td>
</tr>
<tr>
<td>AF119939</td>
<td></td>
</tr>
<tr>
<td>AF119935</td>
<td></td>
</tr>
<tr>
<td>AF119918</td>
<td></td>
</tr>
<tr>
<td>AF119951</td>
<td></td>
</tr>
<tr>
<td>AF119946</td>
<td></td>
</tr>
<tr>
<td>AF119925</td>
<td></td>
</tr>
<tr>
<td>AF119922</td>
<td></td>
</tr>
<tr>
<td>AF119921</td>
<td></td>
</tr>
<tr>
<td>AF119920</td>
<td></td>
</tr>
<tr>
<td>AF119913</td>
<td></td>
</tr>
<tr>
<td>AF119930</td>
<td></td>
</tr>
<tr>
<td>AF119928</td>
<td></td>
</tr>
<tr>
<td>AF119932</td>
<td></td>
</tr>
<tr>
<td>AF119931</td>
<td></td>
</tr>
<tr>
<td>AF119934</td>
<td></td>
</tr>
<tr>
<td>AF119929</td>
<td></td>
</tr>
<tr>
<td>AF119919</td>
<td></td>
</tr>
<tr>
<td>AF119924</td>
<td></td>
</tr>
<tr>
<td>AF119923</td>
<td></td>
</tr>
<tr>
<td>X65278</td>
<td>APASGAHGA</td>
</tr>
<tr>
<td></td>
<td>AAGASVHLK</td>
</tr>
<tr>
<td></td>
<td>PGELDTYGL</td>
</tr>
<tr>
<td></td>
<td>VLGLPSGREI</td>
</tr>
<tr>
<td>Consensus</td>
<td>-----------</td>
</tr>
<tr>
<td></td>
<td>PG-LD-YY--</td>
</tr>
<tr>
<td></td>
<td>G-P S-RE-</td>
</tr>
<tr>
<td>AF119953</td>
<td></td>
</tr>
<tr>
<td>AF016059</td>
<td></td>
</tr>
<tr>
<td>X65277</td>
<td>MRIPVFNVD</td>
</tr>
<tr>
<td></td>
<td>S RHILGDT</td>
</tr>
<tr>
<td>M22628</td>
<td>MRIPVFNVD</td>
</tr>
<tr>
<td></td>
<td>S RHIMGD</td>
</tr>
<tr>
<td>AF056319</td>
<td>MRIPVFNVD</td>
</tr>
<tr>
<td></td>
<td>S KRILGD</td>
</tr>
<tr>
<td>AF119942</td>
<td></td>
</tr>
<tr>
<td>AF119940</td>
<td></td>
</tr>
<tr>
<td>AF119947</td>
<td></td>
</tr>
<tr>
<td>AF119944</td>
<td></td>
</tr>
<tr>
<td>AF119952</td>
<td></td>
</tr>
<tr>
<td>AF119941</td>
<td></td>
</tr>
<tr>
<td>AF119936</td>
<td></td>
</tr>
</tbody>
</table>
Consensus HH-H-S----G -----R---- R-D----D---- PN-----H

AFl19949 HPHMSFTDG TYDGRYLFMN DKANTRVARV RVDIMKCDKI IEIPNASDTH
AF016057 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNASDTH
AF119945 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119938 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119943 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119950 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119954 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119939 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119935 HPHMSFTDG TCYDGRYLFMN DKANTRVARV RCDVMKCDKI TEIPNASDTH
AF119918 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNASDTH
AF119951 HPHMSFTDG TYDGRYLFIN DKANTRVARV RCDVMKVDKI IEIPNASDTH
AF119946 HPHMSFTDG TYDGRYLFIN DKANTRVARV RCDVMKVDKI IEIPNASDTH
AF119937 HPHMSFTDG TYDGRYLFIN DKANTRVARV RCDVMKVDKI VEIPNASDTH
AF119925 HPHMSFTDG TYDGRYLYIN DKANTRVARV RCDVMKVDKI IEIPNASDTH
AJ002531 HPHVSTFDG TYDGRYAFMN DKANSRVARV RCLDKMCDKI ILEPNQHTVH
U047133 HPHNTFDG TYDGRYLAN DKNSRNVCR RLDVMKCDKI IQLPQEVTWH
AF047429 HPHVSTFDG KYDGRYLMN DKANTRVARV RCDVMKTDKI LEIPNAKRGIH
X747492 HPHVSTFDG KYDGRYLMN DKANTRVARV RCDVMKTDKI LEIPNAKRGIH
AF125260 HPHVSTFDG KYDGRYLMN DKANTRVARV RCDVMKTDKI LEIPNAKRGIH
AF016055 HPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKVDKI IEIPNAQDIH
AF016056 HPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKVDKI IEIPNAQDIH
AF119927 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKVDKI IEIPNAHDIH
AF119926 HPHMSFTDG TYDGRYLMN DKANTRVARI RCDVMKVDKI IEIPNAHDIH
AF119921 HPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKVDKI IEIPNAHDIH
AF119920 HPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVMKVDKI IEIPNAHDIH
AF119933 HPHMSFTDG TYDGRYIFAN DKANTRVARI RCDVTMCDKI IEIPNAHDIH
AF119930 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDIMKCDKI IEIPNAHDIH
AF119928 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119932 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119931 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119934 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119929 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119919 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119924 HPHMSFTDG TYDGRYLMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
AF119923 HPHMSFTDG TYDGRYLFMN DKANTRVARV RCDVMKCDKI IEIPNAHDIH
X65278 HTHASYKDG TYDGRYAVMN DKNSRIARI RLDYFICDKI TELPNQGFH

Consensus HH-H-S----G -----R---- R-D----D---- PN-----H

201

AF119953 GLRLQ ....... KVPHTKYM LCNAEFIPFH PNDGSSFDIT GDNAFTMNYA
AF016059 GLRLQ ....... KVPHTKYM LCNAEFIPFH PNDGSSFDIT GDNAFTMNYA
X65277 GLRLQ ....... KVPHTKYM PCNAEFIPFH PNDGKVDLDE DENSFTMNYA
M22628 GLRLQ ....... KVPHTKYM VANAIFIPFH PNDGKVDLDE DENSFTMNYA
AF056319 GLRLQ ....... KVPHTKYM VANAIFIPFH PNDGKVDLDE DENSFTMNYA
AF119942 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119940 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119947 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119944 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119952 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119941 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119936 GLRPQ ....... KFPRTGYYI FANGEHTPI PNDGSI/L/D.D PSYHAI/FTA
AF119949  GLRPQ...  KYPTG YV  FANGEHRVPI  PNDGSILD.E  PEKYHAIFTA
AF016057  GLRPQ...  KYPTG YV  FANGEHRVPI  PNDGSILD.E  PEKYHAIFTA
AF119945  GLRPQ...  KYPTG YV  FANGEHRVPL  PNDGSLLD.E  PEKYHAIFTA
AF119938  GLRPQ...  KYPTG YV  FANGEHRVPL  PNDGSLLD.E  PEKYHAIFTA
AF119943  GPRPQ...  KYPTG YV  FANGEHRVPL  PNDGSLLD.E  PEKYHAIFTA
AF119950  GLRPQ...  KYPTG YV  FANGEHRVPL  PNDGSLLD.E  PEKYRAIFTA
AF119954  GLRPQ...  KYPTG YV  FANGEHRVPL  PNDGSLLD.E  PEKYHAIFTA
AF119939  GLRPQ...  KYPTG YI  FANGEHRVPL  PNDGRLDL.E  PTKYHAIFTA
AF119935  GLRPQ...  KFPKT GY  FANGEHRVPL  PNDGRLDL.E  PEKYHAIFTA
AF119946  GLRPQ...  KFPKT GY  FANGEHRVPL  PNDGRLDL.E  PEKYHAIFTA
AF119937  GMRPQ...  KFPKT GY  FANGEHRVPL  PNDGRLDL.E  PEKYHAIFTA
AF119925  GLRPQ...  KFPKT GY  FANGEHRVPL  PNDGRLDL.D  PEKYVAIFTA
AF119948  GLRPQ...  KFPKT GY  FANGEHRVPL  PNDGRLDL.E  PEKYHAIFTA
AJ002531  GLRLQ...  KWPBNSYVFCNGEDEAPL  VNQGMTMT.D  VAYYVNFITA
X677492  GLRPQ...  KWPBNSYVFCNGEDEAPL  VNQGTMNE.M  VANYYVNFITA
AF125260  GLRPQ...  KWPBNSYVFCNGEDEAPL  VNQGTVND.D  PSYQYNIFITA
AF016055  GMRPQ...  KYPR TGYVFANGEHEAPL  VNQGTMNE.K  PDEYVNFITA
AF016056  GMRPQ...  KYPR TGYVFANGEHEAPL  VNQGTMNE.K  PDEYVNFITA
AF119927  GLRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119926  GLRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119922  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119921  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119920  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.N  PDEYVNFITA
AF119933  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119930  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119928  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.K  PDEYVNFITA
AF119932  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119931  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119934  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119929  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119919  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119924  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
AF119923  GMRPQ...  KFPKT GY  FANGEHEAPL  VNQGTMNE.D  PDEYVNFITA
X65278  GIPDKRDVPD TKINIYTRVFCGGEFGIPL  PSAPT...EDAKYRSLFTC
Consensus  G---------  --K---------  ---E------  ---------------

251
AF119953  VDAETMEAVW  QIVVGDNLDN  SDMDYTGRFA  ASTCYNEKTA  TDLAGMMRNE
AF016059  VDAETMEAVW  QIVVGDNLDN  SDMDYTGRFA  ASTCYNEKTA  TDLAGMMRNE
X65277  VDAETMEAVF  QIVVGDNLDN  TDADYGRFT  AATCYNEKFA  FDLGMMRNE
M22628  IDAETMEMAF  QIVVGDNLDN  TDADYGRFT  AATCYNEKFA  FDLGMMRNE
AF056319  IDAEKMEMAF  QIVVGDNLDN  SDADYTGKYA  ASTCYNEKEK  YDLGMMRNE
AF119942  IDGDAMTVAW  QIVVGDNLDN  VDADYQGKYA  ISSCYNSEGK  INLEQMTSSE
AF119940  IDGDAMTAVW  QIVVGDNLDN  VDADYQGKYA  ISSCYNSEGK  INLEQMTSSE
AF119947  IDGDAMTVAW  QIVVGDNLDN  VDADYQGKYA  ISSCYNSEGK  INLEQMTSSE
AF119944  IDGDAMSVAV  QIVVGDNLDN  VDADYQGKYA  ISSCYNSEGK  INLEQMTSSE
AF119952  IDGDAMKVAW  QIVVGDNLDN  VDADYQGKYA  ISSCYNSEGK  VTLLEQMTSSE
AF119941  IDGDMSKVAV  QIVVGDNLDN  VDADYQGKYA  ISSCYNSEGK  VTLLEQMTSSE
AF119936  IDADEMKVAV  QVMVDRNLIN  VDADYQGKYA  VSSCYNSEGK  VTLAEMTSSS
AF119949 VGDAEMEVAW QVMVDGNLDN CDADYQGKYA ISTCYNSEEA VNLAGMMAE
AF016057 IDGDKMEIAW QVMVDGNLDN CDADYQGLYA FSTCNNSEAA ANLACMNASE
AF119945 LGDSMKVWAV QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTANE
AF119938 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSNE
AF119943 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119950 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119954 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119939 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119951 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119946 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119937 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119925 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AF119948 VDGEMKMTVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
AJ002531 LGDSTMKVAW QVMVDGNLDN CDADYQGLYA FSTCNSEEG VTTAEMTSE
U747133 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF047429 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
X774792 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF125260 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF016055 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF016056 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119927 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119928 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119921 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119922 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119930 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119938 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119939 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119919 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119924 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
AF119923 LDGDSMKVAW QVIVDGNLDN FDAGDQY Rica FSTCNSEEE VTTAEMTSE
X65278 VDAEMKNFVAVQVLGSDGLN CDADYEGKWA FSTSNSEEG MLLEMTKES
Consensus -D--------- QV----N---- -----Y-G---- ----N-E---- ----M---E

301
AF119953 RDWVVVFIDIP AIEKEIKAKR FITLGDSKVP VVDGRKKGDKG DSV...VTRYI
AF016059 RDWVVVFIDIP AIEKEIKAKR FITLGDSKVP VVDGRKKGDKG DSV...VTRYI
X65277 RDWVVVFIDIP AIEKEIKAKR FITLGDSKVP VVDGRKKGDKG DSV...VTRYI
M22626 RDWVVVFIDIP AIEKEIKAKR FITLGDSKVP VVDGRKKGDKG DSV...VTRYI
AF056319 RDWVVVFIDIP AIEKEIKAKR FITLGDSKVP VVDGRKKGDKG DSV...VTRYI
AF119942 QDWAVVFDIA AIEAGVKAGD FETYEGG.VP VLDGRK..G.. SK...YTRYI
AF119940 QDWAVVFDIA AIEAGVKAGD FETYEGG.VP VLDGRK..G.. SK...YTRYI
AF119947 QDWAVVFDIA AIEAGVKAGD FETYEGG.VP VLDGRK..G.. SK...YTRYI
AF119944 QDWAVVFDIA AIEAGVKAGD FETYEGG.VP VLDGRK..G.. SK...YTRYI
AF119952 QDWAVVFDIA AIEAGVKAGD FETYEGD.VP VLDGRK..G.. SK...YTRYI
AF119941 QDWAVVFDIA AIEAGVKAGD FETYED.VP VLDGRK..G.. SK...YTRYI
AF119936 QDWAVVFDIA AIEAGVKAGD FETYED.VP VLDGRK..G.. SK...YTRYI
AF119949 QDHVVFVDE AIEAVKAGD FETI.NG.AK VLDGRK.G.SK.YTRYI
AF016057 QDWAVFDIAR AIEAVKAGE FETI.GG.VL VLDGRK.G.SK.YTRYI
AF119945 QDWAVFDIAR AIEAVKNGD YKKM.NG.VP VIDGRK.G.SK.YTRYI
AF119938 QDWAVFDIAR AIEAVKSGD FKKM.NG.VP VIDGRK.G.SK.YTRYI
AF119943 QDWVVFDIAR AIEAVKSGD FKKM.GG.VP VIDGRK.G.SK.YTRYI
AF119950 QDWVVFDIAR AIEAVKNGD FKKM.GG.VP VIDGRK.G.SK.YTRYI
AF119954 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119939 QDWVVFDIAR AIEAVKNGD FKKM.GG.VP VIDGRK.G.SK.YTRYI
AF119935 QDWVVFDIAR AIEAVKNGD FKKI.GG.VP VIDGRR.G.SK.YTRYI
AF119918 QDWVVFDIAR AIEAVKAGD FKEI.HG.VP VLDGRK.G.SK.YTRYI
AF119951 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119946 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119937 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119925 QDWVVFDIAR AIEAVKAGD YTEH.SG.VP VLDGRK.G.SK.YTRYI
AF119948 QDWVVFDIAR AIEAVKAGD YTEH.SG.VP VLDGRK.G.SK.YTRYI
AF002531 QDWVFPLK AIEAVAAGD YEEM.SG.VP VLNQKG.G.SA.LTRYI
AF119946 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119947 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119922 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119921 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119920 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119933 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119930 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119928 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119932 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119931 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119934 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119929 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119921 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119924 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
AF119923 QDWVVFDIAR AIEAVVFDIK RIEAAVKSGD SK.YTRYI
X65278 RDACVFNNIA RIEAAVAGK FKTYGDSKVP VVDGTOAANK DPKTALTAYV
Consensus -D--V----- --E------ --------------- --G------ ------T-----

351
AF119953 PVKPNHGLN TSDPGKYFIA NGKLSPTCTM IAIKLPDLF AGKLKDPDRC
AF016053 PVKPNHGLN TSDPGKYFIA NGKLSPTCTM IAIKLPDLF AGKLKDPDRC
162277 PVKPNHCGC TSSDDGKYFIA NGKLSPTCSM IAIKLPDLF AGKLKDPDRC
M22628 PVKPNHCGN TSDPGKYFIA NGKLSPTCSM IAIKLPDLF AGKLKDPDRC
AF056319 PVKPNHGLN TSSDDGKYFIA NGKLSPTVSM IAIKLRLDDL NDRKYDREPV
AF119942 PISNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KPRDA
AF119940 PISNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KPRDA
AF119947 PISNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KPRDA
AF119944 PISNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KPRDA
AF119953 PISNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KEPPA
AF119941 PISNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KPRDA
AF119936 PIPNSHPGVN TADPGHHIVT NGKLSPTVSV IDVRKLDPDV ADK1.KPRDA

400
Consensus: PHG--D----G-LSP----F

401

AF119953 VVGEPGLG PLHTTFDGRG NAYTLFLDS QLVKWNLEDA RRAYKGEK.V
AF016059 VVGEPGLG PLHTTFDGRG NAYTLFLDS QLVKWNLEDA RRAYKGEK.V
X65277 VVGEPGLG PLHTTFDGRG NAYTLFLDS QVKKWNMLDA RRAYKGEK.V
M22628 VVGEPGLG PLHTTFDGRG NAYTLFLDS QVKKWNMLDA RRAYKGEK.V
AF056319 IVAEPELG PLHTTFDGRG NAYTLFLDS QVKKWNMLDA RRAYKGEK.V
AF119942 IVAEPELG PLHTAFDNKG NCTYTLFLDS QVAKWSDLDA RKKFAGED.V
AF119940 IVAEPELG PLHTAFDNKG NCTYTLFLDS QVAKWSDLDA RKKFAGED.V
AF119947 IVAEPELG PLHTAFDNKG NCTYTLFLDS QVAKWSDLDA RKKFAGED.V
AF119944 IVAEPELG PLHTAFDNKG NCTYTLFLDS QVAKWSDLDA RKKFAGED.V
AF119952 IVAEPELG PLHAAYDNKG NCTYTLFLDS QVAKWSDLDA RKKFAGED.F
AF119941 IVAEPELG PLHTAYDNKG NCTYTLFLDS QVAKWSDLDA RKKFAGED.V
AF119936 VVAEPELG PLHTAFDNKG NWYTLFLDS QVAKWSDLDA RKFAGED.V
Consensus

---LGLG PLH--D--G ---T--F--DS Q----W----A ----G----

AF119949 VVAEPGLGLG PLHTAYDGKG IAFTTLFLDS QVVKWDIQKA VDAYAGKD.V
AF016057 VVAEPGLGLG PLHTAYDTGQ IAYTTLFFDS QVVKWDIQKA IDAYAGKD.V
AF119945 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QVVKNWIQKA IDAYSKGD.V
AF119938 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QIVKWSMQKA IDAYSKGD.A
AF119943 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QAVKNWIQKA IDAYSKGD.V
AF119950 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QAVKNWIQKA VDAYAGKD.V
AF119954 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QAVKNWIQKA VDAYNGKD.V
AF119939 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QAVKNWIQKA IDAYNGKD.V
AF119935 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QAVKNWIQKA VDAYAGKK.V
AF119918 VVAEPGLGLG PLHTAFDGKG NAFTTLFLDS QVCKWNLEAA VKYAGVED.V
AF119951 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVAKWNMEAA IKQYAGED.A
AF119946 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVAKWNMEAA IKQYAGED.V
AF119937 IVAEPGLGLG PLHTAYDGKG NAFTTLFLDS QVVKWSMKEA IRAFKEGEK.V
AF119925 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QMVWNMLDLA LRKYGEE.V
AF119948 IVAEPGLGLG PLHTAFDGKG NAYTTLFLDS QVAKWSIDKA IRAKGEKEK.V
AJ002531 IVAEPGLGLG PLHTAYDGKG NAYTTLFLDS QVVKNIDLA KRAFKGEK.V
U47133 IVAEPGLGLG PLHTAYDGKG NAYTTLFLDS QVCKWNIEDA KRAYAGEK.V
AF016056 IVAEPGLGLG PLHTAFDGKG NAYTTLFLDS QVAKWNMAEA IEAYGAD.V
AF119927 IVAEPGLGLG PLHTAYDGKG NAFTTLFLDS QIVOKNLDKA IRAWAGEE.V
AF119926 IVAEPGLGLG PLHTAYDGKG NAFTTLFLDS QIVOKNLDKA IRAWYGE.V
AF119922 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNMEAA IKAFEGEK.V
AF119921 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNLEAA IKAFGEV.K
AF119920 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKWNMRA IKAFAKEK.V
AF119933 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNMEAA IKAFAGENV
AF119930 IVAEPGLGLG PLHTAFDGKG NCFTPLFLDS QVAKWNIDKA IQAYGVED.V
AF119928 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVCKWNIDKA KRAYAGEK.V
AF119932 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNIDKA LQKFAGED.V
AF119931 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNIDKA LQKFAGED.V
AF119934 IVAEPGLGLG PLHTAFDGKG NCSTTLFLDS QVVKNIDKA LQKFAGED.V
AF119929 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNIDKA LQKFAGED.V
AF119919 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNIDKA LQKFAGDD.V
AF119924 IVAEPGLGLG PLHTAFDGKG NCFTTLFLDS QVVKNIDKA LQKFAGED.V
AF119923 IVAEPGLGLG PLHTAFDGKG NAYTTLFLDS QVVKNLEAA IKAFGADV.V
X65278 IVAEPGLGLG PLHTAFDGKG NAYTTLFLDS QLVKWNLDAA IKFHKGDKNA

Consensus

---LGLG PLH--D--G ---T--F--DS Q----W----A ----G----

AF119953 DYIRQKLDVR YQPGHNHASL TETSEADKw VVVLSKFSKD RFLLTPGPLHP
AF016059 DYIRQKLDVH YQPGHNHASL TETSEADKw VVVLSKFSKD RFLLTPGPLHP
X65277 NYIKQKLDVQ YQPGHNHASL CETSEADKw LVALSKFSKD RFLLTPGPLHP
M22628 NYIKQKLDVQ YQPGHNHASL CETSEADKw LVALSKFSKD RFLPVGPLHP
AF056319 NYIKQKLDVQ YQPGHNHASL TETSEADKw LVVLCFSKD RFLLTPGPLHP
AF119942 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119940 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119947 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119944 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119952 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119941 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119936 DPIULKDHDH YQPGHNHTSM GETKNADKW LISSLKFSKD RFIVGQPLKP
AF119949 DPILDKVDVQ YQPGHNHTSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF016057 SPILDKVDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119945 NPILEKIDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119938 NPILEKIDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119943 NPILEKIDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119942 NPILEKIDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119939 NPILEKIDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119935 NPILEKIDVQ YQPGHNHSSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119918 DPILDKLDVH YQPGHNHTSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119951 DPILQKIDVH YQPGHNHTSM GETKEADGNW LVSLNKFSKD RFINVGPLKP
AF119946 DPILQKIDVH YQPGHNHTSM GETKEADGKW LVSLNKFSKD RFINVGPLKP
AF119937 DPILHKVDVQ YQPGHNHSSM GETKEADGKW LNASLNKFSKD RFINVGPLKP
AF119925 DPILAKTDVQ YQPGHNHSSM GEAKDADGKW LSLINFSKD RFINVGPLKP
AF119948 NPVLKIDVH YQPGHNHSSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119948 NPVLKIDVH YQPGHNHSSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AJ002531 DPITQKLDVH YQPGHNHSMS GQTKEADGKW LSLINFSKD RFINVGPLKP
U47133 DPRHKLDVH YQPGHNHSSM GQTKEADGKW LSLINFSKD RFINVGPLKP
AF047429 NPIKDDKLVDQ YQPHLKTVM GETLDAANDW LVCLCKFSKD RFLNVGLKP
X774792 DPILDKLDVH YQPHLKTVM GETLDAANDW LVCLSKFSKD RFLNVGLKP
AF125260 DPILQKLDVQ YQPHLKTSM GETLEADGKW LVSMNKFSKD RFLNVGLKP
AF016055 DPITSKVDVH YQPHGNSTSM GETAEDAGKW LISMNKFSKD RFLNVGLKP
AF016056 DPITSKVDVH YQPHGNSTSM GETAEDAGKW LISMNKFSKD RFLNVGLKP
AF119927 DPITSKVDVA YQPHGNSTTM GETSEADGKW LISMNKFSKD RFLNVGLKP
AF119926 DPITSKVDVA YQPHGNSTTM GETSEADGKW LISMNKFSKD RFLNVGLKP
AF119922 DPITSKVDVH YQPHGNSTSM GETAEDAGKW LVSMNKFSKD RFINVGPLKP
AF119921 DPITSKVDVH YQPHGNSTSM GETAEDAGKW LVSMNKFSKD RFINVGPLKP
AF119920 DPITSKVDVH YQPHGNSTSM GETAEDAGKW LVSMNKFSKD RFINVGPLKP
AF119920 DPITSKVDVH YQPHGNSTSM GETAEDAGKW LVSMNKFSKD RFINVGPLKP
AF119933 DPITSKVDVH YQPHGNSTSM GETEADGKW LVSMNKFSKD RFINVGPLKP
AF119930 DPILKIDVH YQPGHNHSMS GETKEADGKW LVSMNKFSKD RFLNVGLKP
AF119928 DPILRKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119932 DPIVEKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119931 DPIVEKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119934 DPIVEKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119929 DPIVEKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119919 DPIVEKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119924 DPIVEKVDVH YQPHGNSTSM GETKEADGKW LVSMNKFSKD RFINVGPLKP
AF119923 DPIRKVDVH YQPHGNSTSM GETKEADGKW LISMNKFSKD RFINVGPLKP
X65278 KYVVRDLDLQ YQPHVNASQ SETVAADGKY LAVGCKFSKD RFLPVGLKP
Consensus ...........D-- YQPGH----- -----A----- -----KFSKD R----GPL-P

501 550
AF119953 ENDELID... ................... ................... ................... ............
AF016059 ENDQLLDISG DENKLVHDGP TFAEPHDcil ARRDQIKTRK I...WDRNDFP
X65277 ENDQLLDISG DMLKLVHDGP TFAEPHDCIM ARRDQIKTRK I...WDRNDFP
M22628 ENDQLLDISG DENKLVHDGP TFAEPHDCIM ARRDQIKTTK I...WDRNDFP
AF056319 ENDQLLDISG EENKLVHDGP AFAEPHDCIL ARRDQIKTKK I...WNRNDFP
AF119942 ENDELID... ................... ................... ................... ............
AF119940 ENDELID... ................... ................... ................... ............
AF119947 ENDELID... ................... ................... ................... ............
AF119944 ENDELID... ................... ................... ................... ............
AF119952 ENDELID... ................... ................... ................... ............
AF119941 ENDELID... ................... ................... ................... ............
AF119936 ENDQLLD... ................... ................... ................... ............
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF119949</td>
<td>ENQQLID</td>
</tr>
<tr>
<td>AF016057</td>
<td>ENEQQLIDISG DKMAVLHDGP TFAEPHDCII VRADIVNPRS V..WARDDAM</td>
</tr>
<tr>
<td>AF119945</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119938</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119943</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119950</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119954</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119939</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119935</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119918</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119951</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119946</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119937</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119925</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119948</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AJ002531</td>
<td>ENDQLIDISG DQMKLVHDGP SFAPHDATI VHRSKINPSI V..WRADPM</td>
</tr>
<tr>
<td>U47133</td>
<td>ENDQLIDISG DEMVLHDP TFAEHDATI VHASKINPVH V..WNRRDPF</td>
</tr>
<tr>
<td>AF047429</td>
<td>ENDQLIDISG DKVLHDP TFAEHDATI VSPSILPNIR SWDRKDPFW</td>
</tr>
<tr>
<td>X774792</td>
<td>ENDQLIDISG DKVLHDP TFAEHDATI VHPISILSDIK SWDRNDPFW</td>
</tr>
<tr>
<td>AF125260</td>
<td>ENDQLIDISG DKVLHDP TFAEHDATI VHASKMNIR SWDRNDPFW</td>
</tr>
<tr>
<td>AF016055</td>
<td>ENDQLIDISG DEMKVHDGP TFAEHSII VHASKVNPV. NISRDDPTW</td>
</tr>
<tr>
<td>AF016056</td>
<td>ENDQLIDISG DEMKVHDGP TFAEHSII VHASKVNPV. NISRDDPTW</td>
</tr>
<tr>
<td>AF119927</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119926</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119922</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119921</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119920</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119933</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119930</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119928</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119932</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119931</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119934</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119929</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119919</td>
<td>ENEQQLID</td>
</tr>
<tr>
<td>AF119924</td>
<td>ENDQLID</td>
</tr>
<tr>
<td>AF119923</td>
<td>ENDQLID</td>
</tr>
<tr>
<td>X65278</td>
<td>ENEQQLIDISG QKVLMADHP VRGEHDFII FRERLVRKQQ VYALDDFPL</td>
</tr>
<tr>
<td>Consensus</td>
<td>EN-QLIDIS- ---M-------P ---EPHD---</td>
</tr>
</tbody>
</table>

AF119953
AF016059
X65277
M22628
AF056319
AF119942
AF119940
AF119947
AF119944
AF119952
AF119941
AF119936
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF119949</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF016057</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119945</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119943</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119950</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119954</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119939</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119935</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119918</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119951</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119946</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119926</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119944</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119922</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119921</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119920</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119933</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119930</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119928</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119932</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119931</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119934</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119929</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>AF119924</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
</tr>
<tr>
<td>AF119923</td>
<td>WATTAAQAAK DGVKLESSE VVREGNKVRV YIWVAPTFS MEKFTVKQGD</td>
<td>X65278</td>
<td>AIIKDPKEG VFRNGKVTV KITSOAPAFS LEEFKLKKGD</td>
</tr>
<tr>
<td>Consensus</td>
<td>----------------- ------------------ ---R-G-KV-V ----AP----- F-----</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence</th>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF119953</td>
<td>ETVTTITNL DIEDVSHGFV MVNHGVSMEI RQQTASITF TAGKP.GCSG</td>
<td>AF016059</td>
<td>ETVTTITNL DIEDVSHGFV MVNHGVSMEI RQQTASITF TAGKP.GCSG</td>
</tr>
<tr>
<td>X65277</td>
<td>ETVTVTNMD QIEDVSHGFV MVNHGVSMEI SPQQTSSITF IADKFGLHWY</td>
<td>M22628</td>
<td>ETVTTITND QIEDVSHGFV MVNHGASMEI SPQQTSSITF VADKFGLHWY</td>
</tr>
<tr>
<td>AF056319</td>
<td>ETVTTITND QIEDVSHGFV MTNHGASMEI SPQQTSSITF TADKAG...</td>
<td>AF119942</td>
<td></td>
</tr>
<tr>
<td>AF119940</td>
<td></td>
<td>AF119947</td>
<td></td>
</tr>
<tr>
<td>AF119944</td>
<td></td>
<td>AF119952</td>
<td></td>
</tr>
<tr>
<td>AF119941</td>
<td></td>
<td>AF119941</td>
<td></td>
</tr>
<tr>
<td>AF119936</td>
<td></td>
<td>AF119936</td>
<td></td>
</tr>
</tbody>
</table>
AF119949 ...........................................................
AF016057 EVTLIVTNLD KVDDLTHGIT IANYGVAMEV GPQDTSSVTF NADRPGVHWF
AF119945 ..........................................................
AF119938 ..........................................................
AF119943 ..........................................................
AF119950 ..........................................................
AF119954 ..........................................................
AF119939 ..........................................................
AF119935 ..........................................................
AF119918 ..........................................................
AF119951 ..........................................................
AF119946 ..........................................................
AF119925 ..........................................................
AF119922 ..........................................................
AF119948 ..........................................................
AJ002531 QVTVYITNID AVEDLTHGFC IVNYGIMEV APQATASVTF SADKAGVYWY
U47133 EVTVYVTNID EVEDLTHGFC IYNYGINMEV APQATASVTF KASRPGVYWY
AF047429 EVTVYVTNLD EIDDLTHGFT MGNYGVAMEI GPQQTSSVTF VAANPGVYWY
X774792 EVTVYVTNLD EIDDLTHGFT MGNYGVAMEI GPQMTSSVTF VAANPGVYWY
AF125260 EVTVYVTNLD DVDDLTHGFT MGNHGVAMEI GPQATASVTF IADRPGVHWY
AF016055 EVTVYVTNLD DIDDVTHGFC LSNHGAVEV APQATASVTF IADRPGVHWY
AF016056 EVTVYVTNLD DIDDVTHGFC LSNHGAVEV APQATASVTF IADRPGVHWY
AF119927 ..........................................................
AF119926 ..........................................................
AF119922 ..........................................................
AF119921 ..........................................................
AF119920 ..........................................................
AF119933 ..........................................................
AF119930 ..........................................................
AF119928 ..........................................................
AF119932 ..........................................................
AF119931 ..........................................................
AF119934 ..........................................................
AF119929 ..........................................................
AF119919 ..........................................................
AF119924 ..........................................................
AF119923 ..........................................................
X65278  EVTLIIITNLID KIEDLTHGFA IPKYNVNFIV NPQETASVTF VADKPGVFWC
Consensus -VT---TN-D ---D--HG-- ----------------- -P----S-TF -A--G--W-

651
AF119953 ..........................................................
*AF016059 ILHWFCHALH ME
X65277 YCSWFCHALH MEMGRMMVE PA*
M22628 YCSWFCHALH MEMGRMMVE PA*
AF056319 ..........................................................
AF119942 ..........................................................
AF119940 ..........................................................
AF119947 ..........................................................
AF119944 ..........................................................
AF119952 ..........................................................
AF119941 ..........................................................
AF119936 ..........................................................
<table>
<thead>
<tr>
<th>Accession</th>
<th>Sequence Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AF119949</td>
<td>YCQWFCHALH ME</td>
</tr>
<tr>
<td>AF016057</td>
<td>YCQWFCHALH</td>
</tr>
<tr>
<td>AF119945</td>
<td></td>
</tr>
<tr>
<td>AF119938</td>
<td></td>
</tr>
<tr>
<td>AF119943</td>
<td></td>
</tr>
<tr>
<td>AF119950</td>
<td></td>
</tr>
<tr>
<td>AF119954</td>
<td></td>
</tr>
<tr>
<td>AF119939</td>
<td></td>
</tr>
<tr>
<td>AF119935</td>
<td></td>
</tr>
<tr>
<td>AF119918</td>
<td></td>
</tr>
<tr>
<td>AF119951</td>
<td></td>
</tr>
<tr>
<td>AF119946</td>
<td></td>
</tr>
<tr>
<td>AF119937</td>
<td></td>
</tr>
<tr>
<td>AF119925</td>
<td></td>
</tr>
<tr>
<td>AF119948</td>
<td>YCSWFCHAMH MEMKGRMFVE PKSV*</td>
</tr>
<tr>
<td>AJ002531</td>
<td>YCSWFCHAMH MEMKGRMFVE PKSV*</td>
</tr>
<tr>
<td>U47133</td>
<td>YCTWFCHAMH MEMKGRMLVE AQGA*</td>
</tr>
<tr>
<td>AF047429</td>
<td>YCQWFCHALH MEMRGRMFVE PKGA*</td>
</tr>
<tr>
<td>X774792</td>
<td>YCQWFCHALH MEMRGRMLVE PKEA*</td>
</tr>
<tr>
<td>AF125260</td>
<td>YCQWFCHALH MEMRGRMFVE PKAA*</td>
</tr>
<tr>
<td>AF016055</td>
<td>YCQWFCHALH ME</td>
</tr>
<tr>
<td>AF016056</td>
<td>YCQWFCHALH ME</td>
</tr>
<tr>
<td>AF119927</td>
<td></td>
</tr>
<tr>
<td>AF119926</td>
<td></td>
</tr>
<tr>
<td>AF119922</td>
<td></td>
</tr>
<tr>
<td>AF119921</td>
<td></td>
</tr>
<tr>
<td>AF119920</td>
<td></td>
</tr>
<tr>
<td>AF119933</td>
<td></td>
</tr>
<tr>
<td>AF119930</td>
<td></td>
</tr>
<tr>
<td>AF119928</td>
<td></td>
</tr>
<tr>
<td>AF119932</td>
<td></td>
</tr>
<tr>
<td>AF119931</td>
<td></td>
</tr>
<tr>
<td>AF119934</td>
<td></td>
</tr>
<tr>
<td>AF119929</td>
<td></td>
</tr>
<tr>
<td>AF119919</td>
<td></td>
</tr>
<tr>
<td>AF119924</td>
<td></td>
</tr>
<tr>
<td>AF119923</td>
<td></td>
</tr>
<tr>
<td>X65278</td>
<td>YCTWFCHALH LEMRTRMIVE A*</td>
</tr>
</tbody>
</table>

Consensus YC--FCHA--H --EM--R--VE ----
Genbank File of the Sequence of the \textit{nos} Cluster from \textit{A. cycloclastes}

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>AF047429</th>
<th>9120 bp DNA</th>
<th>BCT</th>
<th>28-MAY-1999</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>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.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ACCESSION</td>
<td>AF047429</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NID</td>
<td>g4914312</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td>AF047429.2 .GI:4914312</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td>.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>Achromobacter cycloclastes.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ORGANISM</td>
<td>Achromobacter cycloclastes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bacteria; Proteobacteria; beta subdivision; Alcaligenaceae; Achromobacter.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>1 (bases 1 to 9120)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>The nos (nitrous oxide reductase) gene cluster from the soil bacterium Achromobacter cycloclastes: Cloning, sequence analysis, and expression</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>2 (bases 1 to 9120)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>Direct Submission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Submitted (06-FEB-1998) Chemistry and Biochemistry, Montana State University/Bozeman, 108 Gaines Hall, Bozeman, MT 59717, USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>3 (bases 1 to 9120)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>Direct Submission</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Submitted (28-MAY-1999) Chemistry and Biochemistry, Montana State University/Bozeman, 108 Gaines Hall, Bozeman, MT 59717, USA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>REMARK</td>
<td>Sequence update by submitter</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMMENT</td>
<td>On May 28, 1999 this sequence version replaced gi:2935349.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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"
gene
gene
sig_peptide
cds
cds
cds
mat_peptide
mat_peptide
sig_peptide
gene
cds
functions="denitrification enzyme; nitrous oxide reduction to di-nitrogen"
functions="accessory protein for CuZ metallocenter assembly"
EC_number="1.7.99.6"
product="nitrous oxide reductase"
product="nitrous oxide reductase precursor"
product="nitrous oxide reductase"
product="nitrous oxide reductase"
mat_peptide

4465..5736
/gene="nosD"
/product="putative ATP/GTP binding protein NosF"
/protein_id="AAD09159.1"
/db_xref="PID:g2935353"
/db_xref="GI:2935353"
/translation="MTPTLTISRLTKRFSVEALKSVSLSALEPGRARRALLGNGAKS
TMKTLVLGISLPFDSGEVSQGASPSGSPARTQVAYLPENPVAFPALTGEQRLHYLAL
RGNKFAQTELARRVGLHHAARRRIGTYSKGMQRQVGLAQTLIGRPLLVLDEFTSGL
DPVSRREDYDDLLGLAEGTAALLLSHVLTEVEARTDSILLSSGELVEAGTIALDLR
RAALPVASFVRFCAPFAPLALTTPFEVGEEDGLLRLRSQSEKQLLLARIALGSAI
ADLVIPSLEDYSHFSRRDQ"
/gene="nosY"
/CDS
/translation="MSRIATAVSEFRIALRNNWSIAITGCMVLFLALVAAAGSAPTG
DVGVDRLSVATLSLAVYLVLALLMISDAVAGEVERTLEVLLTTYPVSRQLILL
GKLLAHAILGLVATLYGAAALAAAWFDPFRAGLQGLIWRHLSVULLGATFLGTKY
ALSALARRPSGAAGLVALVALWAVYLVYDIALALLALITDVGGAFTHAlPVALLANPAD
AFRVFNLSAQQVSAAGGLGGAAGTYPFPLWQSAASSLLAWFPARIHLAAAFRKVTP"
/gene="nosL"
/note="putative Type II signal peptidase leader"
/gene="nosL"

7419..7475
/gene="nosL"
/note="putative Type II signal peptidase leader"
/gene="nosL"

7419..8000
131

/CDS 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="MRTRLRFVLFVAAALALLSACKEDVAQSVIPQDMTPETLGHYQMNLLEHFGPKAQQIFLEGEPAPLFFSQVQDAAYARGPEQIALFILVYVMDGAAAGATWDQPGDINVIAAAKAFIVYVGSRGGMGAPEAVFSSRDEAAAFVLAEGGQQVIALADITDAMVLFVETGESEPRADDENLGRALRFHPAGG"

/mat_peptide 7476..7997
/gene="nosL"
/product="NosL"

/gene 8005..8916
/gene="nosX"
/CDS 8005..8916
/note="possible regulatory partner of NosR"
/codon_start=1
/transl_table=11
/product="NosX"
/protein_id="AADO9162.2"
/db_xref="PID:g4914313"
/db_xref="GI:4914313"
/translation="MLLTRRRLIAISAVLAPATAHANAAATRFWGTQLGARASIRIAPHEAGTVRSMVEIDRLNVLSLYRANSLCRNRDGHLQPFFLELECILADARHLASGGRFDPTQPLNYWAQASANGPAPADVDRVNGWGVKKVLGAAATLTLPMLTLNGICQGYYVADRAMLVTEILDITIDGETGERLALGRRPGGGPWVLRTGERLPLRQAVATSAAPLTFDQAARDHILDPLGMPAPANWRSVSYSASSAAAAADALATAACLIPDAGISIVMAVRDFDGLRLEAASV"

BASE COUNT 1644 a 2923 c 2801 g 1752 t

ORIGIN

1 gtatggctgg gttgcgcgct gcctcgcgct cccctgcgcga taatcgcgggg aagggcagggc
61 aagggcgcgg cagatgcacct tcgaaggttat tgtgatcgcg ccggcgcgct ggcgcgctgga
dcgagttgct tgcgctgtttg tggagcgtat gaagcagatc cccctgcgtcgc gcctgccggcctgc
gcgtcggcgtc cggctagtga aacattcgga acccatcgtcc ctcactgtga tgcgctggctgc
ggcgcggccgg aataacgaat gcggcggctgg cgggctgggtgg cttcgaaatc gatcccgcggc
ctaccgcgcc ggccgactgg gtcgaggcag aaggtgacgg cacgctgcgc cgcctgtcgc tccggttacg
901 cacgctgcgc cgcctgtcgc tcgacgttgc gcaggtcaat gctgcctttcg ccggccatcc
cgatgagcgc gccgcgagc gcgcgattgc gcaagcgccg gaaacgacat tcatcgagat gcatggccggct
1021 gcggcggcca cgcctgttgct ttcggtgcca ttcggtgcca aatgctctgc cggcttcgac
cggcggcaga ctttccgca gggcggaggg ggcggatcga gtcgctatgg gggctggcgtg ggtttcgtg
1141 gttgctggcc gcgctgcggc tgtgcctgtct gtcgctgctgg cggcggctgg ggggcgtggtg
cggcggccgg aataacgaat gcggcggctgg cgggctgggtgg cttcgaaatc gatcccgcggc
cctaccgcgcc ggccgactgg gtcgaggcag aaggtgacgg cacgctgcgc cgcctgtcgc tccggttacg
1321 cgcgggaggg gcgcttgcgcc gtcgctgctgg cggcggctgg ggggcgtggtg ggtttcgtg

1644 a 2923 c 2801 g 1752 t
1321 ggctcgaaaa actctttcata cccttcatct cttccgtaac ctctcgcag ctttgaagct ctctcgtgct
d1441 tggctcgccg ccaccaacgg tggagaccgt tgcagcctgt gctatggttg acaggagac
d1501 tcacatcgtc cccggcatcg tggactgtca acggatcaag ctaatgccgt ctctcgtgct
d1561 cggcgcacat tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d1621 ccggtgctgt gctatcaggt tgggtccagt ctctcgtgct
d1681 cggcgcacat tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d1741 acgggttgac cggcgcacat tgggtccagt tgggggtacg tggagctgtg acgggttgac
d1801 cggcgcacat tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d1861 cggcgcacat tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d1921 tgcagcctgt tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d1981 tgcagcctgt tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d2041 taatgccgaa gaaatgctgc tgggtccagt tgggggtacg tggagctgtg acgggttgac
d2101 tgcagcctgt tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d2161 cggcgcacat tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat
d2221 gatgagatcc gccggagctt gacgctgggt cggcgcacat tgggtccagt tgggggtacg
d2281 gatgagatcc gccggagctt gacgctgggt cggcgcacat tgggtccagt tgggggtacg
d2341 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2401 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2461 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2521 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2581 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2641 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2701 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2761 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2821 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
d2881 tgggtccagt tgggggtacg tggagctgtg acgggttgac cggcgcacat tgggtccagt
8461 gcagcgcattc cgtttcattc gcggcatgcgg ctgacgttga acggcatcgg ccaggggtat
8521 gtgcgcgcacc gcgtagcagc aatgctttga gttcagggac tgacggacat tctgatcgtat
8581 acggcgcagtagctgcgtc gggggcaecga cgccgacggcg gcggatggcc cgtccggctg
8641 gaaacagcgc aagcgcctggc gtgccgcacc gcgggcggtcg ccctctcgcgc gggaatgccc
cggccagcgg cactggcgcct tcgctcctcg tcggcatatc cctgcgccgc cgcggacgcg
8701 ctgcgcacgcgc ctcagcctactgc gggcccggcg ggagcattcgc gacggacgcgt
8761 cccggggtgt aaggggtggt gcccggcttc cggagcctgc acctgcagcgc gggggctgatc
ggagctgtcc gctggtggtt ggtgtatgtt ccggtggtgc ccggatgcgg ccgtttgcgtctgcgcgaaaggta ggacggagtacggcagcagc gcaatgctttc gcggcatgcgg ctgacgttga acggcatcgg ccaggggtat
8821 gtgcgcgcacc gcgtagcagc aatgctttga gttcagggac tgacggacat tctgatcgtat
8881 ttcggagcgggc cccgcctgtac gcgggctgcgc gtttaaggct gcggacccag taaacctgttgcaccgcctgcgcgcctgcgtgttcggcatatc cctgcgccgc cgcggacgcg
9001 acgcagcgcag cagtcgccttc cggagtgtgtt cttcttttct ctgcggcgcgg ccgagggcca
9061 gtgcgcgcacc gcgtagcagc aatgctttga gttcagggac tgacggacat tctgatcgtat
 gcagcgat cacgtttgga cggcatgcgg ctgacgttga acggcatcgg ccaggggtat
## APPENDIX C

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

<table>
<thead>
<tr>
<th>LOCUS</th>
<th>AF047430</th>
<th>2426 bp DNA BCT 06-MAR-1998</th>
</tr>
</thead>
<tbody>
<tr>
<td>DEFINITION</td>
<td>Sinorhizobium meliloti regulatory protein NosR (nosR) gene, complete cds.</td>
<td></td>
</tr>
<tr>
<td>ACCESSION</td>
<td>AF047430</td>
<td></td>
</tr>
<tr>
<td>NID</td>
<td>g2935357</td>
<td></td>
</tr>
<tr>
<td>VERSION</td>
<td>AF047430.1 GI:2935357</td>
<td></td>
</tr>
<tr>
<td>KEYWORDS</td>
<td>Sinorhizobium meliloti.</td>
<td></td>
</tr>
<tr>
<td>SOURCE</td>
<td>Sinorhizobium meliloti.</td>
<td></td>
</tr>
<tr>
<td>ORGANISM</td>
<td>Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group; Rhizobiaceae; Sinorhizobium.</td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>1 (bases 2109 to 2426)</td>
<td></td>
</tr>
<tr>
<td>AUTHORS</td>
<td>Holloway, P., McCormick, W., Watson, R.J. and Chan, Y.K.</td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>Identification and analysis of the dissimilatory nitrous oxide reduction genes, nosRZDYF, of Rhizobium meliloti</td>
<td></td>
</tr>
<tr>
<td>MEDLINE</td>
<td>96198150</td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>2 (bases 1 to 2426)</td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>The nos (nitrous oxide reductase) gene cluster from the soil bacterium Achromobacter cycloclastes: Cloning, sequence analysis, and expression</td>
<td></td>
</tr>
<tr>
<td>REFERENCE</td>
<td>3 (bases 1 to 2426)</td>
<td></td>
</tr>
<tr>
<td>TITLE</td>
<td>Direct Submission</td>
<td></td>
</tr>
<tr>
<td>JOURNAL</td>
<td>Submitted (06-FEB-1998) Chemistry and Biochemistry, Montana State University/Bozeman, 108 Gaines Hall, Bozeman, MT 59717, USA</td>
<td></td>
</tr>
</tbody>
</table>

**Features**

- Location/Qualifiers
  - source
    - /organism="Sinorhizobium meliloti"
    - /db_xref="taxon:382"
  - gene
    - /synonym="nosR"
    - /codon_start=1
    - /transl_table=11
    - /product="regulatory protein NosR"
    - /protein_id="AAC05094.1"
    - /translation="MYQAAIHKNLFKRLLRGLVALCIAAMLLAGPSMAAGQLSNYLQKVQPOQIPGARTRFGEVGFPIAPVFREGESILGAYLNVDVTSSVGYSGKPIHVVIDPKGVRGLKLVDHKEPIVLI1GEPEAKVSVNABALGKLGRVSAVGFPQVDIVSGA"
TVTBLVMDGSIVRSALGLKHGNRFQGADA-VPEQSSEIRKVDLLRTGDSLWETVLGDS
VSLRILTVGVEFAFAPQQAAADDHPEPSGDRFDYLAPTVAPFVIGRSLLGDSR
YEQRMKALKPGEEAIVVGDAFGSFKGSGYVRGSFFRIDLJQDSQGLRFRD
PTLAAAGAPRLREALFLVFPADGFDDIAAPWELQQLVQRTAV5KAVLFLPYN
LQSYFLVDAHVFDVAAAAPFSVPEAQIESAPTAAPFEGALWVPNW5MFRV5VY
TAVTNALAVLTL
LVFFQDNWKLVRALFENMVRVYLLFLTVLWCYNALQSVNVLTFTNSLTCFOMHE
FLASAPLVFLWNGSAAFLWFRQPGCWLPCFGALQELVNNARNKLKVQVPYFW
LHLSRIPTKYLTFLLGFLGLSLSYLSLALAEFAEVEFKTAIIILKFAREMPFV
VATVL
AALEFLEIRECYRLCPGAALAIPLGRIFMFEILKRWEPCGSPQCRAKECEPQVS
IHPE
GAINVNECYICMCQELYHDDQRCPHQLVRLKREFMALSTPASRGAPAKTVTH
GAPTRKADAPENV"
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.
FT Magnitude

Normalized Intensity

- oxidized Pea seedling AO, SADA (02/89)
- reduced Pea seedling AO, SADA (02/92)
- reduced Pig Kidney AO, PADC (06/94)
- NosL, CNL0B (12/98)

Energy (eV)

$k^2 \chi(k)$

- reduced Pea seedling AO, SADA (02/92)
- reduced Pig Kidney AO, PADC (06/94)
- NosL, CNL0B (12/98)

$k (\text{Å}^{-1})$

FT Magnitude

- oxidized Pea seedling AO, SADA (02/89)
- reduced Pea seedling AO, SADA (02/92)
- reduced Pig Kidney AO, PADC (06/94)
- NosL, CNL0B (12/98)

$R' (\text{Å})$
CNLOB Fit 11

$k^3 \chi(k)$

$k(\text{Å}^{-1})$

FT Magnitude

$R'(\text{Å})$
### Table 1. Curve fitting results for Cu EXAFS modeled as CuN,S.

<table>
<thead>
<tr>
<th>Sample filename (k range)</th>
<th>Fit</th>
<th>x,y</th>
<th>Shell</th>
<th>(N_s)</th>
<th>(R_{as}) (Å)</th>
<th>(\sigma_{as}^2) (Å²)</th>
<th>(\Delta E_0) (eV)</th>
<th>(f^b)</th>
<th>BVSc</th>
</tr>
</thead>
<tbody>
<tr>
<td>NosL + Cu (I)</td>
<td>1</td>
<td>2,0</td>
<td>Cu-N</td>
<td>2</td>
<td>2.18</td>
<td>0.0027</td>
<td>4.51</td>
<td>0.174</td>
<td></td>
</tr>
<tr>
<td>CNLOB (2-12.5 Å⁻¹)</td>
<td>2</td>
<td>3,0</td>
<td>Cu-N</td>
<td>3</td>
<td>2.19</td>
<td>0.0041</td>
<td>6.05</td>
<td>0.177</td>
<td></td>
</tr>
<tr>
<td>(\Delta k'\chi = 12.25)</td>
<td>3</td>
<td>4,0</td>
<td>Cu-N</td>
<td>4</td>
<td>2.18</td>
<td>0.0062</td>
<td>4.95</td>
<td>0.164</td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5,0</td>
<td>Cu-N</td>
<td>5</td>
<td>2.16</td>
<td>0.0085</td>
<td>3.54</td>
<td>0.167</td>
<td></td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6,0</td>
<td>Cu-N</td>
<td>6</td>
<td>2.15</td>
<td>0.0108</td>
<td>2.59</td>
<td>0.172</td>
<td></td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0,2</td>
<td>Cu-S</td>
<td>2</td>
<td>2.34</td>
<td>0.0046</td>
<td>-2.73</td>
<td>0.144</td>
<td></td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0,3</td>
<td>Cu-S</td>
<td>3</td>
<td>2.34</td>
<td>0.0070</td>
<td>-3.23</td>
<td>0.135</td>
<td></td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0,4</td>
<td>Cu-S</td>
<td>4</td>
<td>2.32</td>
<td>0.0093</td>
<td>-4.66</td>
<td>0.133</td>
<td></td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0,5</td>
<td>Cu-S</td>
<td>5</td>
<td>2.32</td>
<td>0.0115</td>
<td>-5.50</td>
<td>0.134</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0,6</td>
<td>Cu-S</td>
<td>6</td>
<td>2.31</td>
<td>0.0136</td>
<td>-6.21</td>
<td>0.139</td>
<td></td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>2,2</td>
<td>Cu-N</td>
<td>2</td>
<td>2.08</td>
<td>0.0032</td>
<td>-0.65</td>
<td>0.121</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>2</td>
<td>2.37</td>
<td>0.0032</td>
<td>[-0.65]a</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2,3</td>
<td>Cu-N</td>
<td>2</td>
<td>2.06</td>
<td>0.0069</td>
<td>-2.19</td>
<td>0.127</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>3</td>
<td>2.35</td>
<td>0.0066</td>
<td>[-2.19]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>3,2</td>
<td>Cu-N</td>
<td>3</td>
<td>2.09</td>
<td>0.0090</td>
<td>-0.34</td>
<td>0.125</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>2</td>
<td>2.37</td>
<td>0.0042</td>
<td>[-0.34]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>3,1</td>
<td>Cu-N</td>
<td>3</td>
<td>2.12</td>
<td>0.0044</td>
<td>1.15</td>
<td>0.123</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>1</td>
<td>2.39</td>
<td>-0.0007</td>
<td>[1.15]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1,3</td>
<td>Cu-N</td>
<td>1</td>
<td>2.06</td>
<td>-0.0001</td>
<td>-2.76</td>
<td>0.123</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>3</td>
<td>2.35</td>
<td>0.0060</td>
<td>[2.76]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>2,1</td>
<td>Cu-N</td>
<td>2</td>
<td>2.12</td>
<td>0.0010</td>
<td>1.32</td>
<td>0.122</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>1</td>
<td>2.39</td>
<td>-0.0009</td>
<td>[1.32]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>1,2</td>
<td>Cu-N</td>
<td>1</td>
<td>2.08</td>
<td>-0.0014</td>
<td>-1.13</td>
<td>0.121</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>2</td>
<td>2.36</td>
<td>0.0027</td>
<td>[-1.13]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

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. \(\sigma_{as}^2\) is a mean square deviation in \(R_{as}\). \(\Delta E_0\) is the shift in \(E_0\) for the theoretical scattering functions.

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

\[
f^b = \left( \frac{\sum_i \left( R_i^{\text{obs}} - R_i^{\text{calc}} \right)^2}{N} \right)^{1/2}
\]

\(N\) is the number of data points.

B V S C = \(\exp\left[\frac{\sum_i \left( R_i^{\text{obs}} - R_i^{\text{calc}} \right)^2}{N} \right]^{1/2}\)

b BVSc=\(\Sigma\exp\left[\left(r_{Cu(I)}-r_2\right)B, 9=0,37, r_{Cu(I)-N}=1.595, r_{Cu(I)-S}=1.898\right]\) (Liu & Thorp, 1993)

d Numbers in square brackets were constrained to be either a multiple of the above value (\(\sigma_{as}^2\)) or to maintain a constant difference from the above value (\(R_{as}, \Delta E_0\)).
CNLOC vs Fits 17 & 18

$k^3 \chi(k)$ vs $k(\text{Å}^{-1})$

- **CuS$_2$(O,N) Fit 17**
- **CuS$_2$(imid) Fit 18**
- **CNLOC, NosL + Cu(I) (04/99)**

FT Magnitude vs $R'(\text{Å})$

- **CNLOC, NosL + Cu(I) (04/99)**
- **CuS$_2$(imid) Fit 18**
- **CuS$_2$(O,N) Fit 17**
### Table 1. Curve fitting results for Cu EXAFS modeled as CuN,S.

<table>
<thead>
<tr>
<th>Sample filename (k range)</th>
<th>Fit</th>
<th>$x,y$</th>
<th>Shell</th>
<th>$N_s$</th>
<th>$R_{as}$ (Å)</th>
<th>$\sigma_{as}^2$ (Å$^2$)</th>
<th>$\Delta E_0$ (eV)</th>
<th>$r^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NosL + Cu (I)</td>
<td>1</td>
<td>2.0</td>
<td>Cu-N</td>
<td>2</td>
<td>2.19</td>
<td>0.0002</td>
<td>9.60</td>
<td>0.152</td>
</tr>
<tr>
<td>CNLOC (2-13 Å$^{-1}$)</td>
<td>2</td>
<td>3.0</td>
<td>Cu-N</td>
<td>3</td>
<td>2.19</td>
<td>0.0027</td>
<td>19.29</td>
<td>0.153</td>
</tr>
<tr>
<td>$\Delta k^2 \chi = 9.55$</td>
<td>3</td>
<td>4.0</td>
<td>Cu-N</td>
<td>4</td>
<td>2.17</td>
<td>0.0063</td>
<td>7.78</td>
<td>0.167</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>5.0</td>
<td>Cu-N</td>
<td>5</td>
<td>2.13</td>
<td>0.0109</td>
<td>5.38</td>
<td>0.180</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>6.0</td>
<td>Cu-N</td>
<td>6</td>
<td>2.11</td>
<td>0.0146</td>
<td>3.68</td>
<td>0.187</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>0.2</td>
<td>Cu-S</td>
<td>2</td>
<td>2.32</td>
<td>0.0033</td>
<td>-4.70</td>
<td>0.100</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>0.3</td>
<td>Cu-S</td>
<td>3</td>
<td>2.32</td>
<td>0.0059</td>
<td>-5.01</td>
<td>0.101</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>0.4</td>
<td>Cu-S</td>
<td>4</td>
<td>2.32</td>
<td>0.0083</td>
<td>-5.58</td>
<td>0.107</td>
</tr>
<tr>
<td></td>
<td>9</td>
<td>0.5</td>
<td>Cu-S</td>
<td>5</td>
<td>2.31</td>
<td>0.0107</td>
<td>-6.19</td>
<td>0.117</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>0.6</td>
<td>Cu-S</td>
<td>6</td>
<td>2.31</td>
<td>0.0131</td>
<td>-6.82</td>
<td>0.131</td>
</tr>
<tr>
<td></td>
<td>11</td>
<td>2.2</td>
<td>Cu-N</td>
<td>2</td>
<td>2.00</td>
<td>0.0095</td>
<td>-3.77</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>2</td>
<td>2.33</td>
<td>0.0032</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>12</td>
<td>2.3</td>
<td>Cu-N</td>
<td>2</td>
<td>2.08</td>
<td>0.0295</td>
<td>-2.74</td>
<td>0.093</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>3</td>
<td>2.33</td>
<td>0.0062</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>13</td>
<td>3.2</td>
<td>Cu-N</td>
<td>3</td>
<td>2.07</td>
<td>0.0217</td>
<td>-0.85</td>
<td>0.084</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>2</td>
<td>2.34</td>
<td>0.0037</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>14</td>
<td>3.1</td>
<td>Cu-N</td>
<td>3</td>
<td>2.07</td>
<td>0.012</td>
<td>0.41</td>
<td>0.086</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>1</td>
<td>2.35</td>
<td>0.0002</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>1.3</td>
<td>Cu-N</td>
<td>1</td>
<td>1.95</td>
<td>0.0016</td>
<td>-6.43</td>
<td>0.079</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>3</td>
<td>2.322</td>
<td>0.0056</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>16</td>
<td>2.1</td>
<td>Cu-N</td>
<td>2</td>
<td>2.04</td>
<td>0.0074</td>
<td>-1.70</td>
<td>0.090</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>1</td>
<td>2.34</td>
<td>-0.0001</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>17</td>
<td>1.2</td>
<td>Cu-N</td>
<td>1</td>
<td>1.97</td>
<td>0.0025</td>
<td>-5.61</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-S</td>
<td>2</td>
<td>2.32</td>
<td>0.0030</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>18</td>
<td>1.2</td>
<td>Cu-S</td>
<td>2</td>
<td>2.32</td>
<td>0.0031</td>
<td>-5.49</td>
<td>0.077</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-N</td>
<td>1</td>
<td>1.99</td>
<td>0.0025</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-C</td>
<td>1</td>
<td>2.96</td>
<td>0.0037</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-C</td>
<td>1</td>
<td>2.97</td>
<td>0.0037</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-N</td>
<td>1</td>
<td>4.10</td>
<td>0.0051</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-C</td>
<td>1</td>
<td>4.11</td>
<td>0.0051</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>19</td>
<td>2.2</td>
<td>Cu-S</td>
<td>2</td>
<td>2.32</td>
<td>0.0031</td>
<td>-4.58</td>
<td>0.072</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Cu-N</td>
<td>2</td>
<td>2.00</td>
<td>0.0064</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bond</td>
<td>N</td>
<td>Distance (Å)</td>
<td>Error (Å)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>---</td>
<td>--------------</td>
<td>-----------</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu-C</td>
<td>2</td>
<td>2.98</td>
<td>0.0096</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu-C</td>
<td>2</td>
<td>3.01</td>
<td>0.0097</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu-N</td>
<td>2</td>
<td>4.13</td>
<td>0.0133</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cu-C</td>
<td>2</td>
<td>4.14</td>
<td>0.0133</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( a \) 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. \( \sigma_{as}^2 \) is a mean square deviation in \( R_{as} \). \( \Delta E_0 \) is the shift in \( E_0 \) for the theoretical scattering functions.

\( b \) \( f^2 \) is a normalized error (chi-squared):

\[
f^2 = \frac{\sum_i \left( \frac{\chi_i^{\text{calc}} - \chi_i^{\text{th}}}{\sigma_i} \right)^2}{N_f}
\]

\( c \) BVS = \( \sum_i \left( \frac{\sigma_i}{\beta} \right) \), \( \beta = 0.37 \), \( r_0(\text{Cu(I)-N}) = 1.595 \), \( r_0(\text{Cu(I)-S}) = 1.898 \). (Liu & Thorp, 1993)

\( d \) Numbers in square brackets were constrained to be either a multiple of the above value (\( \sigma_{as}^2 \)) or to maintain a constant difference from the above value (\( R_{as}, \Delta E_0 \)).