JOURNAL OF BACTERIOLOGY, June 1994, p. 3148–3153
0021-9193/94/$04.00+0
Copyright © 1994, American Society for Microbiology
Vol. 176, No. 11

Organization of the *hao* Gene Cluster of *Nitrosomonas europaea*: Genes for Two Tetraheme *c* Cytochromes

DAVID J. BERGMANN,[1,2] DAVID M. ARCIERO,[2] AND ALAN B. HOOPER[1,2]*

Graduate Program in Biochemistry[1] and Department of Genetics and Cell Biology,[2] University of Minnesota, St. Paul, Minnesota 55108

Received 22 November 1993/Accepted 30 March 1994

The organization of genes for three proteins involved in ammonia oxidation in *Nitrosomonas europaea* has been investigated. The amino acid sequence of the N-terminal region and four heme-containing peptides produced by proteolysis of the tetraheme cytochrome c_{554} of *N. europaea* were determined by Edman degradation. The gene (*cycA*) encoding this cytochrome is present in three copies per genome (H. McTavish, F. LaQuier, D. Arciero, M. Logan, G. Mundfrom, J. A. Fuchs, and A. B. Hooper, J. Bacteriol. 175:2445–2447, 1993). Three clones, representing at least two copies of *cycA*, were isolated and sequenced by the dideoxy-chain termination procedure. In both copies, the sequences of 211 amino acids derived from the gene sequence are identical and include all amino acids predicted by the proteolytic peptides. In two copies, the *cycA* open reading frame (ORF) is followed closely (three bases in one copy) by a second ORF predicted to encode a 28-kDa tetraheme *c* cytochrome not previously characterized but similar to the *nirT* gene product of *Pseudomonas stutzeri*. In one copy of the *cycA* gene cluster, the second ORF is absent.

The chemolithotrophic soil bacterium *Nitrosomonas europaea* derives energy from the oxidation of ammonia to nitrite. Ammonia is oxidized to hydroxylamine ($O_2 + NH_3 + 2e^- \rightarrow NH_2OH + H_2O$) (where e^- is an electron) by ammonia monooxygenase in the plasma membrane (9). Hydroxylamine is oxidized to nitrite ($NH_2OH + H_2O \rightarrow HNO_2 + 4H^+ + 4e^-$) by hydroxylamine oxidoreductase (HAO) in the periplasm (2–4, 14).

Electrons from HAO are accepted by cytochrome c_{554}, a 25-kDa periplasmic (2) tetraheme (3) protein. Cytochrome c_{554} is thought to be central to a critical electron transfer branch point: two electrons must return to AMO, and two electrons pass to a second pathway which branches again to either a terminal electron acceptor (a cytochrome aa_3 oxidase [8] or to a nitrite reductase [10, 21]) or, once every 5.7 cycles (31), to ATP-dependent reverse electron transfer for the production of NAD(P)H (1). Cytochrome c_{554} is well suited for this central role because it is a two-electron carrier (4) and because its positive charge is in keeping with the formation of ionic complexes with possible electron donor/acceptors which are negatively charged: HAO (15), the outer face of the membrane (32), and cytochrome c_{552}. The electron carriers downstream from this branch point, which may include other cytochromes and ubiquinone (9, 22, 23), have not been identified.

Three copies of the gene for cytochrome c_{554} are present in the genome of *N. europaea* (20). Furthermore, each copy of the gene for cytochrome c_{554} lies within 2.7 kb downstream of a copy of the gene for HAO (20), although not in the same operon (29). Multiple copies of genes in prokaryotes are unusual. Different copies of a gene may encode polypeptides with slight differences in amino acid sequence which alter properties such as susceptibility to feedback inhibition of enzymatic activity, as observed in the *aroF*, *aroG*, and *aroH* genes of *Escherichia coli* (27). On the other hand, different copies of a gene may encode identical polypeptides but may have different promoter sequences and hence different transcriptional regulation, as in the *psbAII* and *psbAIII* genes of *Anacystis nidulans* (13).

Currently, amino acid sequences suggest the existence of four classes of tetraheme *c* cytochromes: cytochrome c_3 of *Desulfovibrio* species (24, 30), the cytochrome of the bacterial photosynthetic reaction center (11), cytochrome c_{552} nitrite reductase of *E. coli* (7), and the *nirT* cytochrome in the cd_1-dependent nitrite respiration in *Pseudomonas stutzeri* (16).

Here, we present the entire nucleotide sequence of one copy of the cytochrome c_{554} gene, *cycA*, and 94% of the sequence of a second copy. Cytochrome c_{554} is shown to represent a new class of tetraheme *c* cytochromes. We also have found the presence of an open reading frame downstream of two copies of *cycA* which is predicted to encode a tetraheme *c* cytochrome with homology to the predicted *nirT* gene product of *P. stutzeri*.

MATERIALS AND METHODS

Purification and digestion of cytochrome c_{554}. *N. europaea* was grown (19) and cytochrome c_{554} was extracted and purified (5) as described previously. Heme peptides were isolated following digestion with each of the enzymes indicated in Results and Discussion. Prior to proteolytic digestion, cytochrome c_{554} was denatured in 20 mM HCl for 60 min. All digestions were carried out in 100 mM Tris-acetate solution (pH 7.6) in the dark at 37°C. Thermolysin or chymotrypsin was added at 5% and trypsin at 2% (wt/wt, relative to cytochrome c_{554}). Digestions were carried out for 48 h (thermolysin or chymotrypsin) or 14 h (trypsin). Initial fractionation of heme-containing peptides was carried out as previously described with the trichloroacetic acid-acetonitrile solvent system (6). Subsequent purification of individual heme-containing peptides was carried out by gradient elution on the same column. Buffer A was 10 mM sodium phosphate (pH 3.2), and buffer B was 10 mM sodium phosphate (pH 3.2)–acetonitrile (10:90). A 20-min linear gradient was run between 15 and 45% buffer B. Heme-containing peptides were sequenced as described previously (6).

* Corresponding author. Phone: (612) 624-4930. Fax: (612) 625-5754. Electronic mail address: hooper@molbio.cbs.umn.edu.

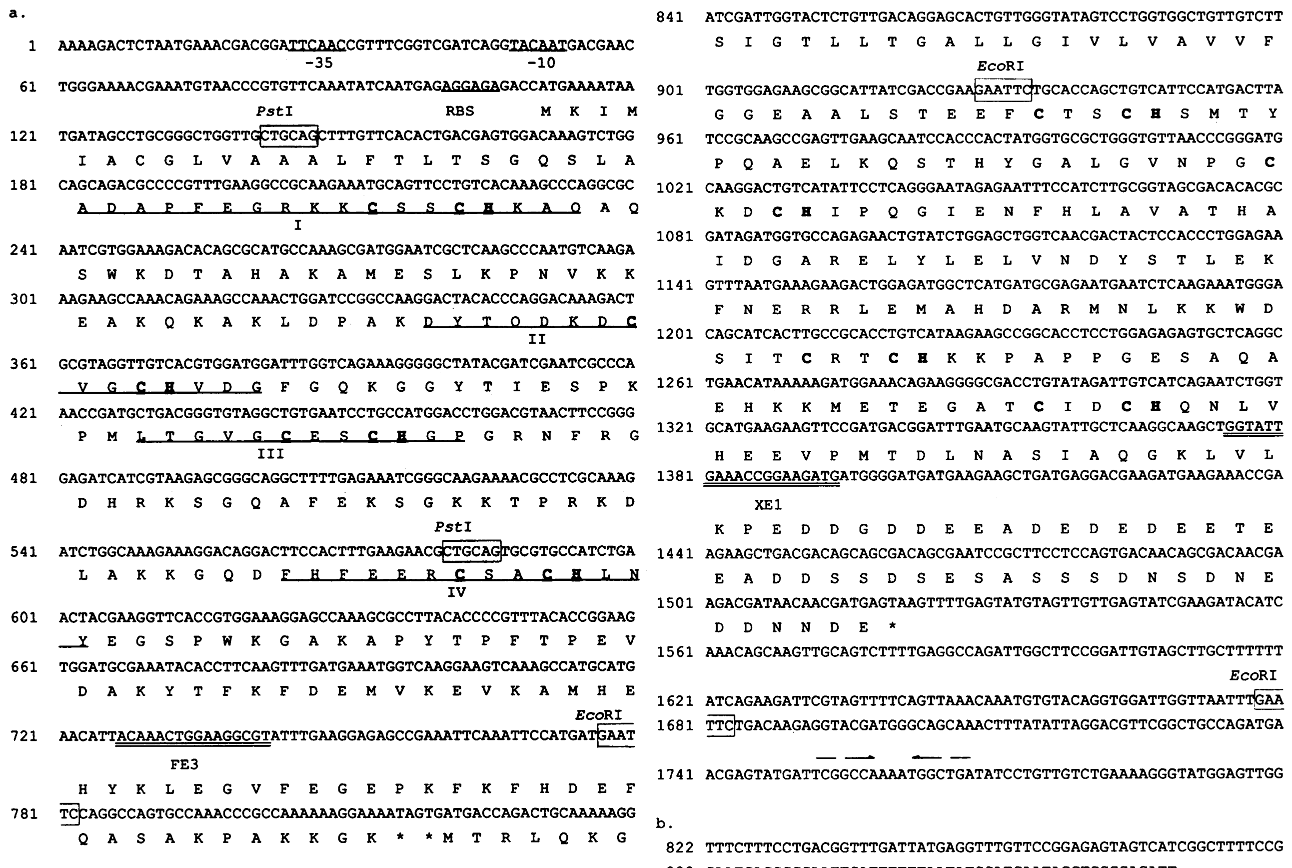

FIG. 1. (a) Nucleotide sequence of the *cycA*-ORF2 gene pair, derived from pE273 and pPST18. The predicted amino acid sequence is shown below the nucleotide sequence. The amino acid sequences of proteolytic, heme-containing fragments of cytochrome c554, determined by Edman degradation, are underlined and numbered with roman numerals as described in Results and Discussion. *Eco*RI and *Pst*I restriction sites are shown in boxes, and heme-binding sites (CXXCH) are shown in boldface. Possible −35 and −10 promoter sequences and the ribosome binding site (RBS) are underlined. The sequences of the FE3 and XE1 oligonucleotide probes are underlined twice. (b) Nucleotide sequence of pKPN43 beyond base 822. The nucleotide sequence of pKPN43 from bases 0 to 821 is identical to that from pE273 and pPST18 except for bases 632 (G) and 677 (T).

Cloning and sequencing the region coding for cytochrome c_{554}. Genomic DNA was isolated from 1 g of cells as described previously (19). Digestion of DNA with restriction endonucleases was performed as recommended by manufacturers (Promega Corp., Madison, Wis., or Gibco BRL, Gaithersburg, Md.). Agarose gel electrophoresis of DNA digests was performed in Tris-acetate-EDTA buffer, and Southern blots were performed as described by Sambrook et al. (28). Oligonucleotides were obtained from Oligos Etc. (Willsonville, Oreg.). Oligonucleotide probes were end labeled with [γ-^{32}P]ATP, and hybridizations and washes of Southern blots were performed as described previously (19).

A size-fractionated library of *Eco*RI fragments, approximately 2.7 kb in size, was created with a pBluescript plasmid vector, and libraries of 1.8-kb *Pst*I and 4.3-kb *Kpn*I fragments were created with a pUC119 vector as previously described (19). The plasmid libraries were screened, and plasmid DNA was purified as described by Sambrook et al. (28).

Dideoxy sequencing of double-stranded plasmid DNAs was performed with the Sequenase version 2.0 kit (U.S. Biochemicals, Cleveland, Ohio). Nested deletions were created for sequencing with an exonuclease III-mung bean nuclease kit (Stratagene, La Jolla, Calif.). Primers for sequencing included the M13 −40 primer, the M13 reverse primer, and custom-synthesized deoxyoligonucleotides. The Genetics Computer Group (GCG) Sequence Analysis Software Package (12) was used to analyze nucleotide sequence. The combined GenBank and EMBL data bases were searched for amino acid sequences homologous to those of *N. europaea* cytochrome c_{554} and open reading frame 2 (ORF2) with the GCG TFastA program, and the Swissprot data base was searched with the GCG FastA program.

Nucleotide sequence accession number. The nucleotide sequence for *cycA* and the ORF downstream has been deposited in the GenBank data base under accession number U08288.

RESULTS AND DISCUSSION

Amino acid sequence of heme-containing proteolytic peptides. The amino acid sequences of the four heme-containing peptides, determined by N-terminal analysis, chymotrypsin and thermolysin (I), chymotrypsin, thermolysin, and trypsin (II), thermolysin (III), and chymotrypsin and thermolysin (IV), are underlined in Fig. 1.

The amino acid sequence VLSEGEWELVHVWAK was previously reported as representing a tryptic fragment of cytochrome c_{554} (20). This sequence is incorrect and resulted from two errors made after isolation of two of the peptides: (i) the sequencing laboratory incorrectly reported sequence of a myoglobin peptide (above) for the tryptic (non-heme-contain-

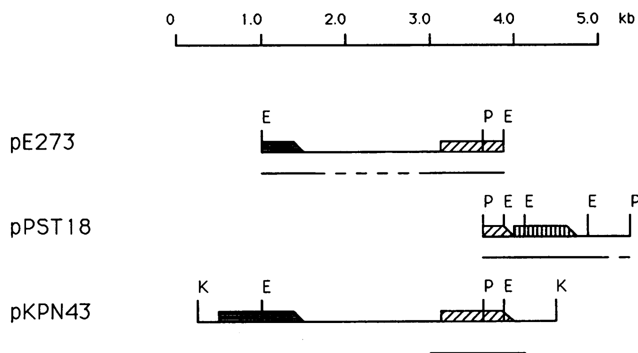

FIG. 2. Clones containing the cytochrome c_{554} coding region. The cytochrome c_{554} coding region (*cycA*) is indicated by diagonal hatches; *hao* is indicated by horizontal hatches, and ORF2 is indicated by vertical hatches. Restriction sites used in cloning are *Eco*RI (E), *Kpn*I (K), and *Pst*I (P). Regions of each clone sequenced on both strands of DNA are indicated by solid lines, and those sequenced on only one strand are indicated by dashed lines.

ing) peptide, and (ii) although the oligonucleotide 5′-GA (C,T)-TA(C,T)-ACI-CA(A,G)-GA(C,T)-AA(A,G)-GA(C,T)-TG(C,T)-GT-3′ (corresponding to the amino acid sequence DYTQDKCV of peptide II) was, in fact, used in the study reported in reference 20, it was mistakenly reported as the myoglobin sequence.

Cloning the region coding for cytochrome c_{554}. A degenerate oligonucleotide [5′-GA(C,T)-TA(C,T)-ACI-CA(A,G)-GA(C,T)-AA(A,G)-GA(C,T)-TG(C,T)-GT-3′] corresponding to the amino acid sequence DYTQDKCV of heme-containing peptide II was used to probe Southern blots of restriction digests of *N. europaea* genomic DNA. A single 2.7-kb *Eco*RI fragment hybridized to the probe (20). Thus, a 2.7-kb *Eco*RI fragment, which contains at least a portion of the cytochrome c_{554} gene, results from digestion of all three copies of the HAO-cytochrome gene cluster.

A library of *Eco*RI fragments approximately 2.7 kb long was created in pBluescript, and nearly 500 bacterial colonies containing the library were screened to yield one clone (pE273) which hybridized with the probe. This clone also hybridized with another degenerate oligonucleotide probe derived from the amino acid sequence of an internal, heme-containing tryptic peptide of HAO, MEYEGEYTHN (6), indicating that the clone contained the carboxy-terminal portion of the gene for HAO (*hao*) as well as at least a portion of the gene for cytochrome c_{554}, which we designate as *cycA*.

The sequence of 776 bases in the 3′ portion of the clone encoded the amino acid sequences predicted from the four heme-containing, proteolytic peptides of cytochrome c_{554} (Fig. 1a) within an ORF which extended 665 bp to the 3′ end of the clone. This clone did not, however, contain a stop codon, indicating that the carboxy terminus of cytochrome c_{554} was not encoded within this clone. A *Pst*I site is located 195 bases from the 3′ end of the cloned 2.7-kb *Eco*RI fragment (Fig. 2), so that a *Pst*I fragment which overlapped the 2.7-kb *Eco*RI fragment and contained the 3′ end of the cytochrome c_{554} gene was cloned. Genomic DNA was digested with *Pst*I and probed with an oligonucleotide, FE3, from the nucleotide sequence between the *Pst*I site and the 3′ end of the 2.7-kb *Eco*RI fragment (Fig. 1a). Three fragments, 1.0, 1.8, and 6.0 kb long, hybridized with the probe. A library of 1.8-kb *Pst*I fragments was constructed in pUC119, and approximately 1,600 bacterial colonies were screened with the FE3 probe to yield two clones. The sequence of one of the clones (pPST18) included a stop codon marking the 3′ end of the *cycA* ORF and, 3 bp downstream, a second ORF (ORF2) extending for 702 bp (Fig. 1a and 2). Although the clones pE273 and pPST18 may be from different copies of *cycA*, the nucleotide sequences in the 201-base region of overlap between the clones were identical.

Because the sequence derived from the two overlapping *Eco*RI and *Pst*I fragments could represent a chimera, the *cycA* gene was cloned in its entirety. The FE3 probe hybridized to 6.5- and 4.3-kb fragments of genomic DNA digested with *Kpn*I. Because no *Kpn*I sites occurred within the previously cloned 2.7-kb *Eco*RI fragment, the 4.3-kb *Kpn*I fragment should contain the entire *cycA* gene. Screening 1,600 bacterial colonies containing a size-fractionated pUC119 library of 4.3-kb *Kpn*I fragments yielded one clone (pKPN43) hybridizing to the probe. The sequence of *cycA* obtained from pKPN43 is identical to that obtained from pE273 and pPST18 except for two bases: 632 (G) and 677 (T) (Fig. 1a). Neither of these differences changed the predicted amino acid sequence of cytochrome c_{554}. However, the nucleotide sequence of the region immediately downstream of the *cycA* ORF in pKPN43 was completely different from the region downstream of *cycA* in pPST18, indicating that ORF2 is absent in pKPN43 (Fig. 1b). A Southern blot of genomic DNA of *N. europaea* digested with six restriction endonucleases and probed with an oligonucleotide, XE1, derived from a sequence within ORF2 (Fig. 1a) also indicated that only two copies of ORF2 are present in the genome (Fig. 3).

Sequence of the gene for cytochrome c_{554}. The amino-terminal sequence of the mature protein, determined by Edman degradation as ADAPFEGRKKCSSCH, begins at base 183 of the nucleotide sequence shown in Fig. 1a. Three potential start codons (ATG) occur upstream of the amino terminus at bases 60, 111, and 120. The ATG at base 111 is the most likely start of translation, since a good match (AGGAG) to the consensus prokaryotic ribosome binding site (AGGA GG) lies 10 bases farther upstream. Likely −35 and −10 promoter sequences are indicated in Fig. 1a, although the transcription start site has not been mapped.

The region between bases 111 and 183 appears to encode a signal peptide (Fig. 1a) consisting of a short hydrophilic region with a positive charge (MK) followed by a long hydrophobic region rich in alanine (IACGLVAAALFTLTS), a glycine residue, and a carboxy-terminal region (QSLA). Similar amino-terminal signal peptides have been observed in many gram-negative bacteria (26). Three successive stop codons (TAG, TGA, and TGA), starting at base 816, mark the apparent 3′ end of the gene for cytochrome c_{554}. The mature cytochrome c_{554} is predicted to be 211 residues long and have a molecular mass of 25,938 Da (including the four hemes). This value agrees well with the values of 21,500 and 25,000 Da determined from gel filtration chromatography and polyacrylamide gel electrophoresis, respectively (3).

Interestingly, no heme-binding site motifs (CXXCH) occur within the carboxy-terminal third of the polypeptide. All cysteines of the molecule are bound covalently to the hemes in common with most *c*-type cytochromes. The cytochrome contains four histidines which are not in the CXXCH region and are available as the possible sixth axial ligand to the heme iron. Since one of the four hemes is high spin (3), at least one of the histidine residues must be unligated or loosely ligated to the iron.

The amino acid composition derived from the sequence indicates that apocytochrome c_{554} is a highly basic protein, with 35 lysine residues and a predicted charge of +13 at pH 7.0 and an isoelectric point (pI) of 9.61. This pI is less than the pI of the holocytochrome c_{554} (10.7) reported by Andersson et al. (3)

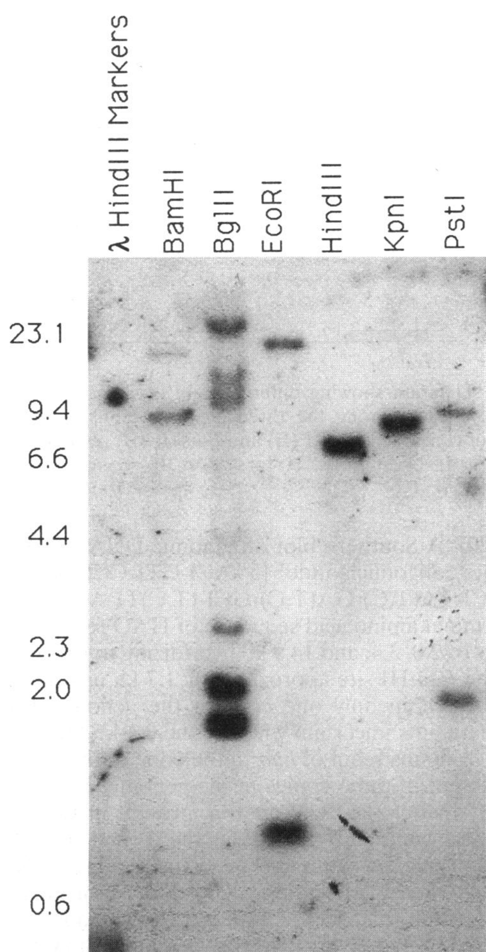

FIG. 3. Autoradiograph of a Southern blot of *N. europaea* genomic DNA, digested with six restriction endonucleases, hybridized with a radiolabeled oligonucleotide (XE1) derived from a sequence within ORF2, and washed for 30 min in high-stringency buffer (0.1% SSPE [1× SSPE is 0.18 M NaCl, 10 mM $NaPO_4$, and 1 mM EDTA {pH 7.7}]–0.2% sodium dodecyl sulfate) at 42°C. The units on the left are in kilobases.

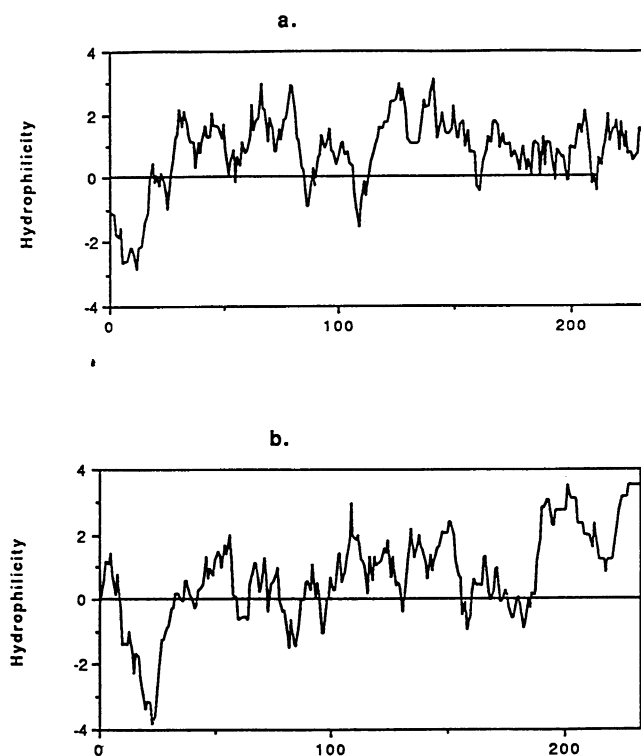

FIG. 4. (a) Hydropathy plot of the nascent cytochrome c_{554} polypeptide. (b) Hydropathy plot for the polypeptide encoded by ORF2. Hydropathy values (14) were calculated over a window of nine residues. Amino acid residue numbers are indicated on the *x* axes.

but close to the pI (9.5) we have recently estimated from isoelectric focusing. The presence of eight heme proprionate groups in the holocytochrome c_{554} might reduce the total positive charge and pI of the protein. The positively charged residues of cytochrome c_{554} may facilitate ionic interactions with HAO (3), c_{552} (3), and the plasma membrane (18).

The cytochrome c_{554} polypeptide is predicted to be hydrophilic along most of its length, as indicated by hydropathy values (17) (Fig. 4a), possibly accounting for its unusual solubility in highly concentrated solutions of ammonium sulfate (32). The hemes themselves will, of course, influence the hydrophobicity of the interior of the holocytochrome.

A search of the combined GenBank and EMBL data bases revealed only low-level similarities of uncertain homology between c_{554} of *N. europaea* and other bacterial cytochromes *c*. Residues 10 to 75 of cytochrome c_{554} are 27.7% identical with residues 102 to 165 of the tetraheme cytochrome c_3 of *Desulfovibrio vulgaris* Hildenborough (30), but much of this similarity is derived simply from the presence of the heme-binding sites (CXXCH). Further, when sequences of cytochrome c_3 and cytochrome c_{554} are aligned by use of the GCG Gap program (12, 25), the quality of alignment between the two sequences is not significantly greater than the average quality of alignment between cytochrome c_3 and 10 randomized cytochrome c_{554} sequences. Cytochrome c_{554} of *N. europaea* appears to be the first example of a new and fifth class of tetraheme *c* cytochromes.

Sequence of ORF2. ORF2 begins only 3 bp downstream from the first stop codon at the 3′ end of *cycA* and thus is part of the *cycA* operon. The predicted gene product of ORF2 is 233 amino acids long. The presence of four putative heme-binding motifs (CXXCH) suggests that ORF2 encodes a tetraheme *c* cytochrome. With the exception of a hydrophobic region 32 amino acids long at the amino terminus, the polypeptide encoded by ORF2 is predicted to be hydrophilic (Fig. 4b). It is likely that the amino-terminal region constitutes a signal peptide for export to the periplasm: an amino-terminal region with two positive charges (MTRLQK-) is followed by a hydrophobic region (-GSIGTLLTGALLGIVLVAVVF-) and a possible peptidase recognition site (-GGEA-) (21). The 44-amino-acid carboxy-terminal region is predicted to be rich in aspartate and glutamate, and the polypeptide is predicted to have a net charge of −30 at pH 7.0 and a pI of 4.21. The molecular mass of the holocytochrome encoded by ORF2, lacking the putative signal peptide, is predicted to be 24,881 Da.

A search of the Swissprot data base revealed significant homology between the predicted ORF2 gene product and the predicted *nirT* gene product of *P. stutzeri* (16). When aligned by the GCG Gap program (12, 25), residues 7 to 201 of *nirT* and residues 2 to 198 of ORF2 had 32.6% sequence identity (Fig. 5). The quality of alignment between *nirT* and ORF2 was

```
  1 MTDKDGNKQQKGGILALLRRPSTRYSLGGILIVGIVAGIVFWGGFNTALE  50
          .: |||:|              |.|:.| : |||: :.. .: .
  1 .....MTRLQKGSI..............GTLLTGALLGIVLVAVVFGGEA  31

 51 ATNTETFCISCHEMGDNVYPEYKETIHYANRTGVRATCPDCHVPRDWTH.  99
    | .||.||.|||.|.  . :|.|:..||:.  ||.:.|.|||:|.: .:
 32 ALSTEEFCTSCHSMT.YPQAELKQSTHYGA.LGVNPGCKDCHIPQGIENF  79

100 ...KMVRKVEASKELWGKIVGTIDTAEKFEAKRLTLARREWARMRASDSR 146
        .: ::::.:||: .:|....| |||:.:||.:|: .: .:: .||
 80 HLAVATHAIDGARELYLELVNDYSTLEKFNERRLEMAHDARMNLKKWDSI 129

147 ECRNCHSLESMSSDMQKQRARKQHEMAREDNLTCIACHKGIAHH...... 190
    .||.||. .. .::    .|. :|..  .:. |||.||..:.|.
130 TCRTCHKKPAPPGE....SAQAEHKKMETEGATCIDCHQNLVHEEVPMTD 175

191 ............LPEGMTEEDED........................... 201
                 ||: .:::|.
176 LNASIAQGKLVLKPEDDGDDEEADEDEDEETEEADDSSDSESASSSDNSD 225
```

FIG. 5. Regions of homology between predicted polypeptides of *nirT* of *P. stutzeri* (upper line) and ORF2 of *N. europaea* (lower line) aligned by the GAP algorithm (7). Identical paired residues are indicated by vertical lines, paired residues with similarity values greater than or equal to 0.5 are indicated by double dots, and paired residues with similarity values greater than or equal to 0.1 are indicated by single dots. Putative heme-binding site motifs are boxed.

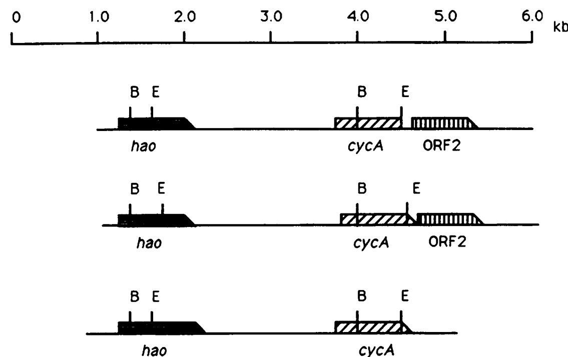

FIG. 6. Diagram showing differences in organization of the *hao*-*cycA* gene cluster among the three copies present in the genome. Positions of conserved *Eco*RI (E) and *Bam*HI (B) restriction sites are shown.

significantly greater than the average quality of alignment between *nirT* and 10 randomized ORF2 sequences. All four putative heme-binding sites of *nirT* share conserved residues with ORF2, although the unique proline present in the second heme-binding site of *nirT* is absent in ORF2. Furthermore, three histidine residues located outside of the putative heme-binding motifs are conserved in *nirT* and ORF2, with the position of H-99 of *nirT* close to that of H-80 of ORF2. These histidines would be good candidates for possible axial ligands to the hemes if the ORF2 and *nirT* gene products are tetraheme *c* cytochromes. The presence of the gene for this putative tetraheme *c* cytochrome within the *hao-cycA* gene cluster suggests that it has a central role in electron transfer. Two properties of the cytochrome derived from the gene sequence might make it well suited for this role: its charge is opposite to that of cytochrome c_{554}, thus allowing for ionic binding, and the number of hemes would accommodate a two- or four-electron transfer.

Organization of the *hao-cycA* gene cluster. Although the complete sequences of all three copies of the *hao-cycA* gene cluster have not been obtained, some conclusions can be drawn regarding organization of this gene cluster and possible differences among the three copies. From Southern blots probed for sequences within *hao* (20, 29), *cycA* (20), and ORF2 (Fig. 1), it can be concluded that the gene cluster consists of *hao*, *cycA*, and ORF2 in two copies, whereas one copy consists of *hao* and *cycA* only (Fig. 6).

The presence of multiple copies of the gene cluster allows a reasonable estimate of the size of the cluster, if one assumes that the presence of conserved restriction sites indicates the likely identity of the region in different copies of the gene cluster. For the two copies of the gene cluster containing ORF2, the region containing conserved restriction enzyme sites in both copies is at least 6.6 kb long; a single 6.6-kb *Kpn*I fragment and a single 6.0-kb *Hin*dIII fragment in a genomic Southern blot hybridized to the XE1 oligonucleotide probe (Fig. 3). The positions of these restriction enzyme sites have not been established by sequencing. In the copy of the gene cluster lacking ORF2, the gene cluster includes at least *hao* and *cycA* (over 3.6 kb), but the region immediately downstream of *cycA* is not included.

The region containing conserved restriction enzyme sites in all three copies of the *hao-cycA* gene cluster does not extend more than 2.0 kb upstream of the 5′ end of *hao*. A *Bam*HI site within *hao* (29) is conserved in all three copies of the gene cluster (20). A Southern blot of genomic DNA probed with a degenerate oligonucleotide [5′-ACI-GTI-CCI-GA(T,C)-GA(G,A)-ACI-TA(T,C)-GA(T,C)-GCI-(T,C)TI-AA-3′] based on the N-terminal amino acid sequence of HAO revealed *Bam*HI fragments of 2.0, 3.4, and 14.9 kb hybridizing to the probe (18). Hence, the *Bam*HI site approximately 1.7 kb upstream of *hao* (29) is present in only one copy in the genome. From the existing data, it is uncertain whether the 0.4-kb ORF (ORFA) located 1.1 kb upstream of *hao* in the copy of the gene cluster (29) is present in three copies in the genome.

The putative *cycA*-ORF2 operon, present in two copies of the *hao-cycA* gene cluster, ends at the 3′ end of ORF2. No ORFs were detected within 500 bp of the 3′ end of ORF2, and an inverted repeat, indicating a possible factor-independent transcription termination site, is present 230 bp downstream of ORF2 (Fig. 1). However, it is not yet known how far downstream of ORF2 the region conserved in both copies of the gene cluster extends.

The sequences of the two copies of *cycA* which have been sequenced are extremely similar; the amino-terminal 94% of one copy of the gene (from pE273) and another copy (from pKPN43) differ only at two bases. Furthermore, the possible −10 and −35 promoter sequences of these two copies are identical, indicating that transcription may be regulated similarly. Interestingly, one copy contains ORF2 as well as *cycA*, almost certainly within the same operon, whereas the other copy contains only *cycA* and not ORF2. It is possible that a certain ratio of cytochrome c_{554} to ORF2 gene product must be maintained; hence, there are three copies of *cycA* and two of ORF2.

ACKNOWLEDGMENTS

We thank Hugh McTavish and Greg Simonsen for advice in planning this study, Candace Pilon for providing cells of *N. europaea*, and Dan Prestridge and Lynda Ellis for help in data analysis.

This work was supported by the National Science Foundation (grant DMB-9019687) and the Minnesota Sea Grant Program (NA86AA-D-56112).

REFERENCES

1. **Aleem, M. I. H.** 1966. Generation of reducing power in chemosynthesis. II. Energy linked reduction of pyridine nucleotides in the chemoautotroph, *Nitrosomonas europaea*. Biochim. Biophys. Acta **113:**216–224.
2. **Andersson, K. K., and A. B. Hooper.** 1983. O_2 and H_2O are each the source of one O in NO_2^-: ^{15}N-NMR evidence. FEBS Lett. **164:**236–240.

3. **Andersson, K. K., J. D. Lipscomb, M. Valentine, E. Munck, and A. B. Hooper.** 1986. Tetraheme cytochrome *c554* from *Nitrosomonas europaea*: heme-heme interactions and ligand binding. J. Biol. Chem. **261:**1126–1138.
4. **Arciero, D. M., C. Balny, and A. B. Hooper.** 1991. Spectroscopic and rapid kinetic studies of reduction of cytochrome *c554* by hydroxylamine oxidoreductase from *Nitrosomonas europaea*. Biochemistry **30:**11466–11472.
5. **Arciero, D. M., M. Collins, J. Haladjian, P. Bianco, and A. B. Hooper.** 1991. Resolution of the four hemes of cytochrome *c554* from *Nitrosomonas europaea* by redox potentiometry and optical spectroscopy. Biochemistry **30:**11459–11465.
6. **Arciero, D. M., and A. B. Hooper.** 1993. Hydroxylamine oxidoreductase from *Nitrosomonas europaea* is a multimer of an octa-heme subunit. J. Biol. Chem. **268:**14645–14654.
7. **Darwin, A., H. Hussain, L. Griffiths, J. Grove, Y. Sambongi, S. Busby, and J. Cole.** 1993. Regulation and sequence of the structural gene for cytochrome *c552* from *Escherichia coli*, not a hexahaem but a 50 kDa tetraheme nitrite reductase. Mol. Microbiol. **9:**1255–1265.
8. **DiSpirito, A. A., J. D. Lipscomb, and A. B. Hooper.** 1986. A three-subunit cytochrome aa_3 from *Nitrosomonas europaea*. J. Biol. Chem. **261:**17048–17056.
9. **Dispirito, A. A., L. R. Taaffe, and A. B. Hooper.** 1985. Localization and concentration of hydroxylamine oxidoreductase and cytochromes *c552*, *c554*, c_m553, c_m552, and *a* in *Nitrosomonas europaea*. Biochim. Biophys. Acta **806:**320–330.
10. **Dispirito, A. A., L. R. Taaffe, J. D. Lipscomb, and A. B. Hooper.** 1985. A "blue" copper oxidase from *Nitrosomonas europaea*. Biochim. Biophys. Acta **827:**320–326.
11. **Dracheva, S., J. C. Williams, G. Van Driessche, J. J. Van Beeumen, and R. E. Blankenship.** 1991. The primary structure of cytochrome *c*-554 from the green photosynthetic bacterium *Chloroflexus aurantiacus*. Biochemistry **30:**11451–11458.
12. **Genetics Computer Group.** 1991. Program manual for the GCG package, version 7, April, 1991. Genetics Computer Group, Madison, Wis.
13. **Golden, S. S., J. Brusslan, and R. Haselkorn.** 1986. Expression of a family of psbA genes encoding a photosystem II polypeptide in the cyanobacterium *Anacystis nidulans* R2. EMBO J. **11:**2789–2798.
14. **Hollacher, T. C., M. E. Tate, and D. J. D. Nicholas.** 1981. Oxidation of ammonia by *Nitrosomonas*: definitive ^{18}O-tracer evidence that hydroxylamine formation involves a monooxygenase. J. Biol. Chem. **256:**10834–10836.
15. **Hooper, A. B., P. C. Maxwell, and K. R. Terry.** 1978. Hydroxylamine oxidoreductase from *Nitrosomonas*: absorption spectra and content of heme and metal. Biochemistry **17:**2484–2489.
16. **Jungst, A., S. Wakayabayashi, H. Matsubara, and W. G. Zumft.** 1991. The nirSTBM region coding for cytochrome cd1-dependent nitrite respiration of *Pseudomonas stutzeri* consists of a cluster of mono-, di-, and tetraheme proteins. FEBS Lett. **279:**205–209.
17. **Kyte, J., and R. F. Doolittle.** 1982. A simple method for displaying the hydrophobic character of a protein. J. Mol. Biol. **157:**105–132.
18. **McTavish, H. E.** 1992. Ph.D. thesis. University of Minnesota, St. Paul.
19. **McTavish, H., J. A. Fuchs, and A. B. Hooper.** 1993. Sequence of the gene coding for ammonia monooxygenase in *Nitrosomonas europaea*. J. Bacteriol. **175:**2436–2444.
20. **McTavish, H., F. LaQuier, D. Arciero, M. Logan, G. Mundfrom, J. A. Fuchs, and A. B. Hooper.** 1993. Multiple copies of genes coding for electron transport proteins in the bacterium *Nitrosomonas europaea*. J. Bacteriol. **175:**2445–2447.
21. **Miller, D. J., and D. J. D. Nicholas.** 1985. Characterization of a soluble cytochrome oxidase/nitrite reductase from *Nitrosomonas europaea*. J. Gen. Microbiol. **131:**2851–2854.
22. **Miller, D. J., and P. M. Wood.** 1982. Characterization of *c*-type cytochromes of *Nitrosomonas europaea* with fluorescent gels. Biochem. J. **207:**511–517.
23. **Miller, D. J., and P. M. Wood.** 1983. CO-binding *c*-type cytochromes and a high-potential cytochrome *c* in *Nitrosomonas europaea*. Biochem. J. **211:**503–506.
24. **Moore, G. R., and G. W. Pettigrew.** 1990. Cytochromes *c*: evolutionary, structural, and physiochemical aspects. Springer-Verlag, Berlin.
25. **Needleman, S. B., and C. D. Wunsch.** 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. **48:**443–453.
26. **Pugsley, A. P.** 1993. The complete general secretory pathway in gram-negative bacteria. Microbiol. Rev. **57:**50–108.
27. **Riley, M.** 1993. Functions of the gene products of *Escherichia coli*. Microbiol. Rev. **57:**862–952.
28. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
29. **Sayavedra-Soto, L. A., N. G. Hommes, and D. J. Arp.** 1994. Characterization of the gene encoding hydroxylamine oxidoreductase in *Nitrosomonas europaea*. J. Bacteriol. **176:**504–510.
30. **Voordoux, G., and S. Brenner.** 1986. Cloning and sequencing of the gene encoding cytochrome c_3 from *Desulfovibrio vulgaris* (Hildenborough). Eur. J. Biochem. **159:**347–351.
31. **Wood, P. M.** 1988. Chemolithotrophy, p. 183–230. *In* C. Anthony (ed.), Bacterial energy transduction. Academic Press, New York.
32. **Yamanaka, T., and M. Shinra.** 1974. Cytochrome *c*-552 and cytochrome *c554* derived from *Nitrosomonas europaea*: purification, properties, and their function in hydroxylamine oxidation. J. Biochem. **75:**1265–1273.