Complete Genome Sequence of the Marine, Chemolithoautotrophic, Ammonia-Oxidizing Bacterium *Nitrosococcus oceani* ATCC 19707†

Martin G. Klotz, 1* Daniel J. Arp, 2 Patrick S. G. Chain, 3,4 Amal F. El-Sheikh, 1 Loren J. Hauser, 5 Norman G. Hommes, 2 Frank W. Larimer, 5 Stephanie A. Malfaiti, 3,4 Jeanette M. Norton, 6 Amisha T. Poret-Peterson, 1 Lisa M. Vergez, 3,4 and Bess B. Ward 7

*University of Louisville, Louisville, Kentucky 40292; Oregon State University, Corvallis, Oregon 97331; Lawrence Livermore National Laboratory, Livermore, California 94550; Joint Genome Institute, Walnut Creek, California 94598; Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831; Utah State University, Logan, Utah 84322; and Princeton University, Princeton, New Jersey 08544*

Received 26 February 2006/Accepted 16 June 2006

The gammaproteobacterium *Nitrosococcus oceani* (ATCC 19707) is a gram-negative obligate chemolithoautotroph capable of extracting energy and reducing power from the oxidation of ammonia to nitrite. Sequencing and annotation of the genome revealed a single circular chromosome (3,481,691 bp; G+C content of 50.4%) and a plasmid (40,420 bp) that contain 3,052 and 41 candidate protein-encoding genes, respectively. The genes encoding proteins necessary for the function of known modes of lithotrophy and autotrophy were identified. Contrary to betaproteobacterial nitrifier genomes, the *N. oceani* genome contained two complete *rrn* operons. In contrast, only one copy of the genes needed to synthesize functional ammonia monooxygenase and hydroxylamine oxidoreductase, as well as the proteins that relay the extracted electrons to a terminal electron acceptor, were identified. The *N. oceani* genome contained genes for 13 complete two-component systems. The genome also contained all the genes needed to reconstruct complete central pathways, the tricarboxylic acid cycle, and the Embden-Meyerhof-Parnass and pentose phosphate pathways. The *N. oceani* genome contains the genes required to store and utilize energy from glycogen inclusion bodies and sucrose. Polyphosphate and pyrophosphate appear to be integrated in this bacterium’s energy metabolism, stress tolerance, and ability to assimilate carbon via gluconeogenesis. One set of genes for type I ribulose-1,5-bisphosphate carboxylase/oxygenase was identified, while genes necessary for methanotrophy and for carboxysome formation were not identified. The *N. oceani* genome contains two copies each of the genes or operons necessary to assemble functional complexes I and IV as well as ATP synthase (one H⁺-dependent F₀F₁ type, one Na⁺-dependent V type).

The ammonia-oxidizing bacterium *Nitrosococcus oceani* ATCC 19707 (Fig. 1) (Bacteria, Proteobacteria, Gammaproteobacteria, Chromatiaceales, *Nitrosococcus, Nitrosococcus oceani*) was the first ammonia-oxidizing bacterium isolated by enrichment culture from seawater (*Nitrosocystis oceanus*) (53), and it resembles the original type strain, *Nitrosococcus winogradskyi* 1892, which was lost. As a member of the order *Chromatiaceales*, the purple sulfur bacteria, *N. oceani* is a member of the evolutionarily oldest taxonomic group capable of lithotrophic ammonia catabolism. To date, *N. oceani* and *N. halophilus* are the only recognized species of gammaproteobacterial ammonia-oxidizing bacteria (AOB). All other cultivated aerobic AOB are *Betaproteobacteria*, and their members have been detected in soils, freshwater, and sediments as well as marine environments (41). In contrast, gammaproteobacterial AOB have only been found in marine or saline environments. *Nitrosococcus oceanus* has been detected in many marine environments using immunofluorescence (72, 78) and more recently on the basis of cloned gene sequences from DNA extracted from natural seawater (61, 74). In addition to the truly marine environment, *N. oceani* was detected by immunofluorescence and fluorescent in situ hybridization in the saline waters of Lake Bonney, a permanently ice-covered lake in Antarctica (70). *Nitrosococcus halophilus* has been isolated only from saline ponds (39) and has not been detected in other environments by using molecular probes.

The general role of nitrifying bacteria in marine systems is to link the oxidizing and reducing processes of the nitrogen cycle by converting ammonium to nitrate. This conversion is responsible for maintaining nitrate, the major component of the fixed nitrogen pool in the oceans, which is present almost everywhere below a few hundred meters at concentrations approaching 40 μM. The deep nitrate reservoir of the oceans, believed to have come about by abiotic processes on the primordial Earth (48), is a huge pool of nitrogen whose availability to primary producers in the surface layer is still controlled largely by physical processes (27). The nitrification process produces oxidized forms of nitrogen that are lost via denitrification and anaerobic ammonia oxidation (anammox) (34, 65). By converting nitrogenous compounds released as waste products of metabolism into NOₓ intermediates that can act both as oxidants and reductants for the fixed N removal processes (14, 73), nitrification closes the global nitrogen cycle. Nitrification occurs in both the water column and in sediments of marine environments. In sediments, nitrification is often
tightly coupled with denitrification and can account for a significant fraction of the total oxygen consumption in sediments (80). It has been discovered only recently that aerobic ammonia oxidation is also carried out by some Crenarchaeota (58 and references therein). Könneke et al. (38) reported the isolation of a marine crenarchaeote, *Candidatus Nitrosopumilus maritima*, that was able to grow chemolithoautotrophically by aerobically consuming ammonia and producing nitrite. This physiological observation has been supported by the identification of DNA sequences similar to the genes encoding the three subunits of ammonia monooxygenase (AMO) from AOB (38). On the other hand, the recently completed genome sequence of the crenarchaeote *Cenarchaeum symbiosum*, a symbiont of marine sponges (28) predicted to be an ammonia-oxidizing archaeon (AOA), lacked all open reading frames (ORFs) with sequence similarity to genes known to be essential to ammonia oxidation in all AOB such as hydroxylamine oxidoreductase (HAO) and cytochromes c554 and c552 (7). Furthermore, the *C. symbiosum* genome lacked the genes for other cytochrome proteins known to control nitrosating stress in AOB (31), except for one (*norQ*) of the four required genes for functional NO reductase (28). In contrast, the *C. symbiosum* genome contained numerous genes that encode putative copper blue proteins (54), including one (CENSYa_1582) with significant sequence similarity to the NegA and NirK protein family, which is implicated in nitrite reduction by AOB (6). Since AMO activity is also copper dependent, it appears that ammonia oxidation in *Archaea* is entirely based on the function of copper-containing proteins. Because copper is known to be redox active only under oxic conditions and because the cytochrome-based core module of bacterial ammonia catabolism (encoded by the *hao* gene cluster) has evolved from bacterial inventory involved in (anaerobic) denitrification (7), it is highly likely that ammonia catabolism in the *Archaea* has evolved fairly late by incorporating an AMO-like function into an ammonia-independent metabolism. Recent phylogenetic analyses of known Amo subunit protein sequences (28, 58, and M. G. Klotz, unpublished results) suggest, indeed, that the AOA likely obtained their *amo* genes via lateral transfer from AOB. It will thus be interesting to identify genetically and biochemically how the AOA resolved the tasks of aerobic ammonia oxidation and detoxification of the resulting NOx compounds.

The susceptibility of ammonia oxidizers to inhibition by sunlight (due to the light sensitivity of ammonia monooxygenase
(AMO)) is probably responsible for the characteristic distribution of nitrification in the water column; maximal rates occur in surface waters near the bottom of the euphotic zone (74). Nitrification rates decrease with increasing depth as the rate of organic matter decomposition (and thus ammonium supply) decreases with depth. As a consequence of both nitrifier and denitrifier activities, the oceans emit large amounts of the greenhouse gas nitrous oxide (57). The total oceanic N₂O inventory is about two-thirds the size of the total atmospheric inventory, and the oceanic N₂O flux to the atmosphere is estimated to be 4 Tg N/year (57). It has been shown that nitrous oxide and nitric oxide can be produced by aerobic AOB either through the reduction of nitrite (NO₂⁻ [17, 66]) or the oxidation of hydroxylamine (NH₂OH [35]).

MATERIALS AND METHODS

Nitrosococcus oceani strain ATCC 19707 was obtained from the American Type Culture Collection and maintained at a temperature of 30°C in the dark on marine medium as described previously (1, 40, 59). For the isolation of genomic DNA, cultures were grown in 0.6- and 1-liter batches of medium in 2- and 4-liter Erlenmeyer flasks, respectively, titrated to pH 8.0 daily with K₂CO₃.

Library construction, sequencing, and sequence assembly. Genomic DNA was isolated from late-exponential-phase cultures of N. oceani strain ATCC 19707 as described by McTavish et al. (49), modified following the recommendations by the Department of Energy's Joint Genome Institute (DOE-JGI; Walnut Creek, CA). The genome was sequenced using the whole-genome shotgun method as previously described (16, 23). Briefly, random 3- and 8-kb DNA fragments were isolated and cloned into pUC18 and pMCL200 vectors, respectively, for amplification in Escherichia coli. A larger fosmid library was constructed containing approximately 40-kb inserts of shared genomic DNA clones into the pC17Fos cloning vector. Double-ended plasmid sequencing reactions were performed by the DOE-JGI using ABI 3730xl DNA Analyzers and MegaBACE 4500 Genetic Analyzers as described on the JGI website (http://www.jgi.doe.gov/).

After quality control of the 60,402 total initial reads of draft sequence, 51,334 were used for the final assembly, producing an average of 9.3-fold coverage across the genome. Processing of sequence traces, base calling, and assessment of data quality were performed with PHRED and PHRAP. Assembled se-

RESULTS AND DISCUSSION

Genome properties. The N. oceani ATCC 19707 genome is comprised of a single circular chromosome (3,481,691 bp; G+C content of 50.4%) and a plasmid (40,420 bp) that contain 3,052 and 41 candidate protein-encoding genes, respectively (Fig. 2). Most (91%) candidate genes were in orthologous clusters ORFs of published genomes, and a total of 76.5%, 68.7%, and 64.9% had hits with ORFs in the COG, Pfam, and InterPro databases, respectively (Table 1). The majority of genes (66.9%) could be assigned a function; however, only 11.4% of these genes were assigned to enzymes, and only 8.6% were connected to the KEGG (Kyoto Encyclopedia of Genes and Genomes) pathways. In contrast, 23.8% of the genes in the Nitrosococcus oceani genome were assigned to enzymes, and 18.6% were connected to the KEGG pathways (http://img.jgi.doe.gov) (16). The taxonomic breakdown of best BLASTP hits against the KEGG completed genomes database is as follows: Gammaproteobacteria (1,424 genes) followed by Betaproteobacteria (521), Cyanobacteria (183), Alphaproteobacteria (112), and Deltaproteobacteria/Epsilonproteobacteria (121). Individual top hits were with ORFs from Methyloccoccus capsulatus (545) followed by Nitrosomonas europaea (224), Pseudomonas aeruginosa PAO1 (189), Azoarcus sp. strain EbN1 (161), Pseudomonas putida (75), and Geobacter sulfurreducens (69).

In contrast to betaproteobacterial nitrifier genomes, the N. oceani genome contained two complete rrn operons that belong to different classes. These operons are located on different replicons, neither of which was near the origin of replication. The rrn operon on the plus strand belongs to the class that contains Ala-tRNA and Ile-tRNA genes between the 16S and 23S genes, whereas the rrn operon on the minus strand has no inserted genes. While the 16S-23S intergenic region in rrn1 contains 714 bp, including the two tRNAs, the 16S-23S intergenic region in rrn2 is only 224 bp, of which 45 bp downstream of the 16S and 107 bases upstream are identical to those positions in rrn1. However, the 16S, 23S, and 5S rRNA genes themselves are 100% identical.

The plasmid comprises mostly hypothetical and conserved hypothetical proteins. A transposase (Noc_A0021), phage integrase (Noc_A0015), and a small number of other phage-related genes along with restriction modification systems are also found on this replicon. A possible replication protein (Noc_A0039) whose only putative homologue (56%) similar over 80% of its length) is the RepA protein encoded by plasmid pRA2 from Pseudomonas alcaligenes (44), together with a possible partitioning system (Noc_A0013 to Noc_A0014), may help this plasmid be maintained within the N. oceani population.

Families and clusters of foreign and repeat sequences. Surprisingly, we identified several large blocks of genes that were identified as putatively phage related, indicating that N. oceani has been a frequent target for bacteriophages in the open ocean (Table 2). Ten regions, ranging in size from 6.4 to 44.3
genes, two were nearly identical frameshifted integrase genes (Noc_0080 and Noc_1095).

Two identical copies of the Tu translation elongation factor (Noc_2326 and Noc_2338) and two gene copies encoding a fatty acid desaturase and a metalloprotein were identified. Additionally, there are two copies of a gene annotated as ammonia permease (Noc_2700 and Noc_2701); however, both were much shorter thanAmtB proteins from other organisms and aligned only with their C termini. Furthermore, the region containing the membrane-spanning domains lacked amino acid residues that are conserved in other AmtB proteins. Hence, operation of AmtB-facilitated ammonia uptake by one or both products of the Noc_2700 and Noc_2701 genes needs to be experimentally verified. In addition to these examples, we have found several lines of evidence suggesting that genes or gene families are undergoing duplication and diversification. Examples include Noc_1310, which appears to be a truncated C-terminal version of Noc_1552, a full-length phosphoenolpyruvate-phosphate transferase, and the gene Noc_0725, which is only ~70% identical to one of three nearly identical copies (Noc_0343, Noc_0724, and Noc_0973) of a predicted hypothetical protein of 492 amino acids.

**Information processing and modification systems.** The genome of *N. oceani* contains the complete sets of genes necessary to encode DNA-directed DNA polymerases I (Noc_0554) and III (Noc_0002, Noc_0288, Noc_0846, Noc_1659, Noc_2593, Noc_2663, and Noc_2814) as well as multiple copies of the gene encoding the epsilon subunit of polymerase III. Genes encoding polymerase II were not identified. *Nitrosococcus oceani* is equipped with a full complement of genes to carry...
out repair of DNA lesions (uvrABCD, recFRO, mutSHT, radC, and recN) and recombination (recA, recD, and recG; ruvABC), which may be necessary as a consequence of exposure to mutagens and uptake of foreign DNA. The genome of N. oceani contains a large number of open reading frames (a total of 24) that code for type I and type III site-specific restriction endonucleases, as concluded from a comparison with available complete genome sequences of ammonia-oxidizing Betaproteobacteria and closely related Gammaproteobacteria, whereof the next highest number of ORFs was found in the genome of Xylella fastidiosa Ann-1 (a total of 13). The presumed function of these restriction modification systems is protection against phage infection and foreign DNA by recognizing specific methylation patterns and distinguishing between host and foreign DNA (11, 15). The reason for this high number of restriction endonucleases in the genome of N. oceani is presently unclear; however, a similarly high number of ORFs encoding site-specific endonucleases have been identified in the unfinished genome of Nitrosococcus halophilus (A. F. El-Sheikh and M. G. Klotz, unpublished results).

The genes encoding the subunits of DNA-directed core RNA polymerase (EC 2.7.7.6) did not reside in a single gene cluster in the N. oceani genome. Whereas the genes encoding the beta and beta’ subunits were arranged in tandem (Noc_2331 and Noc_2330, respectively), the gene encoding the alpha subunit resided upstream of this tandem as gene Noc_2300. A gene encoding an omega subunit (COG1758) was also found (Noc_1213). The genome also contained a variety of genes that encode alternative sigma factors, some of which were found in multiple copies (see below).

Metabolism and transport: genomic basis of ammonia lithotrophy. (i) Energy metabolism: acquisition of reductant from the environment. A cluster of three contiguous genes encoding the subunits of ammonia monooxygenase (Noc_2503 to Noc_2501) were found to be organized in overlapping operons as described previously (1, 59). No additional functional amo genes or amo pseudogenes were found in the genome. As reported previously, the terminator of the amo operon was succeeded by a transcriptional unit containing the orf5 gene (Noc_2500) (59). An additional orf5-like gene was found as an orphan in the genome (Noc_3006). The deduced Orf5 protein sequence revealed a signal peptide and an additional membrane-spanning domain at its C terminus. Such orf5 genes with a high degree of sequence identity and in conservation of synteny have also been identified in betaproteobacterial nitrifier genomes downstream of the amo operon (U92432 and AF016003); however, they have been identified as one of two sequence-related genes in an orf45 transcriptional unit (16). Interestingly, an orf5 homologue was also found in the whole-genome sequence of the gammaproteobacterial methanotroph Methyllococcus capsulatus (Bath) (MCA2130), where it was not in proximity of either of the two gene clusters that encode particulate methane monooxygenase (pMMO), a homologue of AMO (29, 59). Instead, the orf5 gene was resident in a transcriptional unit together with a gene that encodes a pan1-type multicooper oxidase (MCO; MCA2129), which is likely a homologue to the pan1-type MCO-encoding gene found upstream of the amo operon in the N. oceani genome (Noc_2506). Because both AMO and pMMO can oxidize ammonia to hydroxylamine and because of the sequence and genome organizational similarities between the associated mco and orf5 genes in the N. oceani and M. capsulatus genomes, a functional role of their membrane-associated expression products in transfer of electrons or intermediates to ammonia oxidation is proposed.

The oxidation of hydroxylamine is the core of ammonia catabolism, as it provides electrons for redox-dependent proton pumping (cytochrome bc complex, terminal cytochrome c oxidase). Cytochrome P460 (Noc_0899) is a likely ancient hydroxylamine (and nitric oxide) dehydrogenase suited to detoxify the mutagenic hydroxylamine at low turnover rates (9, 10). However, the electrons gained during the oxidation process cannot be transferred directly to the cytochrome bc complex in the membrane and are likely relayed via the periplasmic soluble cytochrome c552 (Noc_0751) to the terminal oxidase. While this pathway contributes to the proton-motive force, it is an inefficient use of reductant. A more efficient pathway occurs when hydroxylamine is catalytically oxidized by hydroxylamine oxidoreductase (HAO; Noc_0892) and electrons are funneled via cytochromes c554 (Noc_0894) and c552 (Noc_0895) to the ubiquinone pool (Q/OH2; Noc_1248 to Noc_1252). The structure and sequence conservation of the HAO-c554-c552 pathway in N. oceani was recently evaluated (7), based in part on the whole-genome sequence reported in this paper. The reduced quinone pool provides reductant to the AMO complex, to the cytochrome bc complex, to the proton-pumping complex (Noc_0297 to Noc_0299) in the electron transport chain, and to the Nuo complex I responsible for reverse electron transport needed to generate NADH (see below). In addition, numerous characterized MCO-encoding genes were found in the genome (Noc_0889 and Noc_2605, type 1 MCOs; Noc_1542, exported MCO; Noc_1741, copper resistance protein), which need further characterization to assess their involvement in catalytic and electron transfer processes of catabolic pathways in N. oceani. In addition, the finding of a gene encoding the red copper protein nitrosocyanin (Noc_1090) in the N. oceani genome suggests that this protein is unique to and important for ammonia catabolism.

The hydrolysis of urea to ammonia and carbon dioxide can be carried out in N. oceani by the ATP-independent heteromultimeric nickel enzyme urea-amidohydrolase (urease) (40). Ureolysis could be beneficial to the cell because it produces N. oceani’s sole sources for energy and reductant (ammonia) as well as carbon (CO2). Given the low concentration of dissolved urea in the oceans, the role of urease in this organism is unclear.

Some betaproteobacterial nitrifiers can utilize H2 (12). However, the genome of N. oceani did not contain any genes that encode subunits of a hydrogenase, which is in agreement with the fact that Nitrosococcus cannot grow on H2 as the sole source of energy and reductant.

(ii) Energy metabolism: electron flow, generation of universal reductant, generation of the proton gradient, and ATP production. The gene profile for N. oceani reveals complete sets of genes for electron transfer from NADH to O2 via NADH quinone oxidoreductase (complex I), cytochrome bc complex (complex III), and a Cu-a,b-type cytochrome c oxidase (complex IV). Coupled with genes for a complete tricarboxylic acid (TCA) cycle and glycolytic pathway, it seems that N. oceani has the potential to gain energy through the oxida-
tion of organic compounds. A complete pentose phosphate pathway provides an alternative mechanism for oxidizing sugars and generating NAD(P)H. It appears that the obstacle to an organotrophic mode of catabolism is N. oceani’s inability to import suitable organic substrates. When using ammonia as the energy source (lithotrophy), there is a need to generate NAD(P)H through reverse electron flow, and a complete TCA cycle is not needed. The complexes involved in electron flow, generation of redox potential and the proton gradient, and production of ATP are described.

Two complete yet different sets of genes encoding complex I (NDH-1/NADH quinone oxidoreductase) are present in the genome of N. oceani. Genes Noc_1115 to Noc_1127, which encode one copy of complex I, are most similar to complex I genes found in other Gammaproteobacteria. This operon contains only 13 genes; the c and d subunits are fused into a single gene. The second set of genes encoding complex I (Noc_2552 to Noc_2565) includes genes with top BLAST hits to Nitrosomonas europaea (6 genes), to other Betaproteobacteria (3 genes), and to Gammaproteobacteria (5 genes). The role of these distinct complexes in N. oceani is unknown. It may be that the complex with the strongest similarity to that found in Nitrosomonas europaea plays a role in reverse electron flow when ammonia is the sole electron donor, whereas the gammaproteobacterial complex may be important in forward electron flow associated with NADH oxidation. The N. oceani genome also contains the genes needed to encode a Na⁺-transporting NADH:ubiquinone oxidoreductase (Noc_0970 and Noc_1170 to Noc_1174).

The presence of candidate genes for a Na⁺-dependent complex I (Noc_0970 to Noc_1170 to Noc_1174), a Na⁺-dependent V-type ATPase (Noc_2081 to Noc_2089), and several Na⁺/H⁺ antiporters (Noc_0159, Noc_0521, Noc_1282, Noc_2134, and Noc_2952) raises the possibility of a sodium circuit in addition to the proton circuit in N. oceani, which is likely an adaptation to N. oceani’s high-salt environment. Under chemolithotrophic conditions, reverse operation of the sodium-dependent complex I in the plasma membrane could generate additional NADH; however, the sodium-dependent ATPase would be needed to remove excess sodium from the cytoplasm. Under mixotrophic conditions, whether lithoheterotroph or organosulfur troph, a sodium gradient might help provide at least some of the organic needs of the cell through import (4). Our finding of several sodium-dependent transporters in the genome (Noc_0779, Noc_1365, Noc_1575, Noc_1600, Noc_2446, and Noc_2711) support this possibility; however, operation of the sodium-dependent ATPase would, again, be needed to remove excess sodium from the cytoplasm. If import of certain organic compounds enabled N. oceani of organotrophy, additional NADH could be generated. The additional NADH produced by the sodium circuit (Na⁺-dependent complex I and transporters) on the expense of ATP generated by the proton circuit could serve the “forward” complex I as an extension of the bacterium’s electron transport chain, thereby effectively converting a sodium-motive force into a proton-motive force. This could be of utility for chemotaxis and proton-dependent transport.

The genes encoding a ubiquinol-cytochrome c reductase (the cytochrome bc₁ complex) are located in an operon (Noc_0297 to Noc_0299). Nitrosococcus oceani has genes encoding two complete terminal cytochrome _c_ oxidases (COX), both of the Cu-aa₃ type. One complex (Noc_3044 to Noc_3047) appears to be of gamma- and betaproteobacterial descent (Pseudomonas/Nitrosomonas), and the other (Noc_1244 to Noc_1247) of Bacteroidetes/Chlorobi descent. There are two additional copies of the genes for subunit 1 and subunit 2 but not for subunit 3. In addition, there are eight copies of genes encoding class I _c_-type cytochromes.

Additional electron sinks are also encoded in the genome, i.e., there are three di-heme cytochrome _c_ peroxidases (Noc_0488, Noc_1263, and Noc_2697). A partial denitrification pathway is present, including a Cu-type nitrite reductase (nirK; Noc_0089) and nitric oxide reductase (nor; Noc_1847 to Noc_1851). However, genes for nitrate reductase (Nar) and nitrous oxide reductase (Nos) were not identified.

Under lithotrophic conditions, the proton gradient is presumably generated by the action of the quinol-cytochrome _c_ oxidoreductase and cytochrome _c_ oxidase. Under organotrophic conditions, we assume that one or both of the NADH-ubiquinone oxidoreductases could also contribute to the generation of the proton gradient. In addition, gene Noc_1901 encodes a proton-translocating inorganic pyrophosphatase which, in the presence of a flux of pyrophosphate, could contribute to the generation of a proton gradient (Fig. 3). Alternatively, the enzyme could use the proton gradient to generate pyrophosphate necessary for glycolysis and other processes. A potential source of pyrophosphate could be polyphosphate, since a polyphosphate kinase is present (Noc_2388). ATP formed from degradation of polyphosphate could release pyrophosphate through the action of one of the many nucleoside diphosphate hydrolases (NUDIX hydrolases) encoded in the genome (Noc_0193, Noc_0306, Noc_2018, Noc_2420, Noc_2512, Noc_2643, and Noc_2749).

A typical proteobacterial H⁺-translocating FₙFᵥ-type ATP synthase is encoded by an operon (Noc_3073 to Noc_3080). Additionally, the genome also encodes a bacterial V-type ATP synthase (Noc_2081 to Noc_2089). Subunits A through I, with the exception of G, are encoded by this operon. A gene encoding subunit G does not appear to be present in the N. oceani genome; however, subunit G is not present in all bacterial V-ATPases (46). It appears to function with subunit D in formation of a peripheral stalk. Genes for subunits A and B, which form the active site of the enzyme, are highly conserved, while the remaining genes in the cluster are less well conserved.

(iii) Energy storage strategies: sucrose synthase and sucrose-phosphate synthase. Nitrosococcus oceani has genes encoding a sucrose synthase (SuSy; Noc_3068) and sucrose phosphate synthase (SPS; Noc_3069). Sucrose synthase in plants functions predominantly in the degradation of sucrose (47). Sucrose phosphate synthase functions in plants to synthesize sucrose phosphate, and sucrose phosphate phosphatase completes the synthesis of sucrose in plants. A separate gene for sucrose phosphate phosphatase is not present in the genome of N. oceani; however, gene Noc_3069 is a fusion of the genes for sucrose phosphate synthase and sucrose phosphate phosphatase. The halocald dehalogenase triad conserved in this superfamily of glycohydrolases is encoded by gene Noc_3069. UDP-glucose is the glucosyl donor for both enzymes, and gene Noc_2280 encodes a UTP glucose-1-phosphate uridylyltrans-
ferase. Gene Noc_3067 encodes a fructokinase that could make the fructose-6-phosphate required by sucrose phosphate synthase.

Sucrose synthesis in bacteria is not well documented or understood (47). Sucrose and sucrose-synthesizing activities were detected in two species of halotolerant methanotrophs, Methylobacter alcaliphilus 20Z and Methylobacter modestohalophilus 10S (37). Genes for sucrose synthesis have been identified, so far, in just a few Proteobacteria, including Acidithiobacillus ferrooxidans and Nitrosomonas europaea. The genes are also present in cyanobacteria, where their function in sucrose synthesis has been demonstrated. Their role in *N. oceani* is unknown. Perhaps sucrose can serve as an osmoprotectant, as has been suggested for other prokaryotes.

(iv) Energy storage strategies: carbon storage products. The genome of *N. oceani* also contains genes encoding the five enzymes necessary for the synthesis of glycogen from fructose-6-phosphate. The genes encoding glucose-1-phosphate adenyltransferase (Noc_0905) and 1,4-alpha glucan branching enzyme (Noc_0904) are contiguous. The gene for phosphoglucomutase was identified in *N. oceani* (Noc_1719), but genes necessary for the synthesis of poly-beta-hydroxybutyrate do not appear to be present in the genome of *N. oceani*.

(v) Central carbon metabolism. The gene profiles are consistent with complete pathways for glycolysis (from phosphorylated sugars or glucose) and gluconeogenesis. However, the mechanism for the interconversion of fructose-6-phosphate and fructose-1,6-bisphosphate is not clear. Gene Noc_0021 encodes fructose-1,6-bisphosphatase. A candidate for an ATP-dependent phosphofructokinase is not present, but gene Noc_2846 shows some similarity to genes encoding diphosphate-fructose-6-phosphate 1-phosphotransferase (EC 2.7.1.90; pyrophosphate-dependent phosphofructokinase). Both a membrane-bound, proton-translocating pyrophosphatase (Noc_1901) and a soluble pyrophosphatase (Noc_1134) are encoded, providing a mechanism to hydrolyze pyrophosphate generated during gluconeogenesis. The proton-translocating enzyme could also use energy from the proton gradient to synthesize pyrophosphate necessary for glycolysis.

The genes encoding a complete pentose phosphate pathway

---

**FIG. 3.** Proposed roles of polyphosphate and pyrophosphate in the *Nitrosococcus oceani* cell. The figure illustrates the mechanisms for phosphate uptake and the proposed flow of phosphate, pyrophosphate, and polyphosphate and their involvement in energy metabolism (ATP, pmf), central pathways (PP-Pfk), and stress tolerance (copper detox) in the *N. oceani* cell. The proposed phosphate-proton exchanger to remove polyphosphate-copper complexes has yet to be experimentally identified.
are also present. We cannot determine if glucose-6-phosphate dehydrogenase would couple to NADP$^+$ (as is most often the case) or NAD$^+$ (which occurs less frequently). A transhydrogenase, encoded by genes Noc_0261 and Noc_0262, would allow transfer of reductant from NADH to NADP$^+$ and from NADPH to NAD$^+$. All the genes necessary for a complete TCA cycle are present. Of particular interest was the presence of genes encoding alpha-ketoglutarate dehydrogenase, because this enzyme activity was missing in several obligate chemolithotrophs, including *Nitrosomonas europaea* (30). The absence of alpha-ketoglutarate dehydrogenase activity would create a break in the TCA cycle that would preclude organotrophy; however, if present, the role of alpha-ketoglutarate dehydrogenase in these exclusively or predominantly lithotrophic organisms is not known. As was the case for *Nitrosomonas europaea* (16), the genes encoding the three subunits (E1 to E3) of alpha-ketoglutarate dehydrogenase were identified in the *N. oceani* genome, and they were contiguous (Noc_0111 to Noc_0113). Because pyruvate dehydrogenase catalyzes a reaction mechanistically similar to that of alpha-ketoglutarate dehydrogenase and has a similar subunit structure, one gene (*lpd*) encodes the E3 subunit for both enzymes in many organisms. However, in *N. oceani*, pyruvate dehydrogenase is encoded by separate genes, Noc_1254 to Noc_1256, which include the gene for subunit E3.

**Amino acid and nucleotide metabolism.** The *N. oceani* genome contains genes encoding the biosynthesis of the 20 amino acids required for the synthesis of proteins. Unless otherwise indicated, all amino acids mentioned were the L form. The genes identified indicate most pathways are similar to previously identified synthesis pathways (5, 26, 68, 69, 77), and the genes encoding the enzymes have values highly similar to those found in other *Proteobacteria*, most of them with best matches to *Methylococcus capsulatus* or within the pseudomonads. While most biosynthetic pathway elements have been identified, missing enzymatic steps are typically involved with dual-function enzymes that may be difficult to identify based on sequence alone (for example, Noc_0176 encodes 3-phosphoshikimate 1-carboxyvinyltransferase with dual functions [EC 1.3.1.1 and EC 2.5.1.19]). In contrast, few amino acid degradative enzymes (catabolic enzymes) or specific transport genes were identified. Scavenging from leucine, valine, and isoleucine may be possible. Several transaminases were identified, but most could not be assigned to specific amino acids. Special attention was paid to amino acid biosynthetic functions found in the last common ancestor, as indicated by their universal distribution in the three domains of life (45).

Several large amino acid biosynthetic operons were identified, including a mixed-function supraperon similar to that found in *Pseudomonas* encoding aromatic amino acid biosynthesis (76) and the *his* operon (22). The aromatic supraperon is found encoded by Noc_0172 to Noc_0177. The dual-function chorismate mutase/prephenate dehydratase P protein is encoded by gene Noc_0174. The histidine biosynthetic genes are not in a single cluster and are located in the genome as genes Noc_2778 and Noc_2779 (*hisDG*) and Noc_3051 to Noc_3057 (*hisCBHAfIE* and *hisL4*). Split organizations of the *his* operon are relatively common and were also found in *Nitrosomonas europaea* (16, 22). The genes *hisL* and *hisE* overlap by 8 bp but do not appear to be fused. The gene encoding histidinol phosphatase (EC 3.1.3.15; Noc_0374) was found outside of the operon.

In the genome of *N. oceani*, 20 aminoacyl-tRNA synthetases (AARS) were identified, including two forms of LysRS (class I and II) and two distinct forms of GlxRS (EC 6.1.1.17), but AsnRS (EC 6.1.1.22) and GlnRS (EC 6.1.1.18) were missing. The most common organism for the top match for the AARS was *Methylococcus capsulatus*, as is the case for the overall taxonomic distribution of top matches. The class II aminoacyl-tRNA synthetases for Phe and Gly have two nonidentical subunits, and genes for both the alpha and beta subunits have been identified adjacent to each other in the genome. While no AARS was identified as the specific GlnRS type (EC 6.1.1.18), it is likely that this function is mediated by the product of gene Noc_0264 or Noc_3250, the nondiscriminatory-type GlxRS. Genes encoding a possible glutamyl-tRNA-Gln amidotransferase (gatAB) were found in the genome, although gatB (Noc_2014) is separated from gatCA (Noc_2635 and Noc_2636). So an indirect route for synthesis of glutaminyl-tRNA is probable. The sequence data alone are insufficient for specifically designating either GlxRS gene a discriminatory role. The presence of two genes does not appear to be a recent duplication event, as the peptides are more similar to their homologues in *Methylococcus* than to each other (only 44% identical to each other over 315 aligned residues versus 55 and 67% identical to the *Methylococcus* peptides).

No candidate for AsnRS was identified in the genome. The AspRS encoded by gene Noc_0302 does contain the GAD domain typically found in AspRS involved in the indirect transamination route to Asn-tRNA (EC 2.5.1.19). The gene Noc_2029, encoding the transamination function, is found outside of the aspartate transcarbamoylase operon. Although typically individual organisms contain only one class of LysRS, genes encoding both classes of LysRS have been identified in the genome of *N. oceani*. This case of LysRS existing in both class I and class II forms in the genome of *N. oceani* is the only known bacterial exception to the AARS “class rule” (60). The only known examples of the presence of both classes of LysRS in a single genome are in the archaea *Methanosarcina barkeri* and *M. acetivorans*. In *N. oceani*, the class I LysRS is encoded by gene Noc_1618, with the best match to *M. barkeri* LysRS. Both selective retention and horizontal gene transfer have played roles in the distribution of class I LysRS in bacteria (2). The gene Noc_2625, encoding 356 amino acids, represents the C-terminal region of class II LysRS with an intact core domain (the best match is to *Pseudalteromonas haloplanktis* TAC125 [CAJ85572]). Several other *Gammaproteobacteria* have a similar protein, as described for COG2269. In *N. oceani* this combination may be an example of gene displacement of the LysRS class II by a functional LysRS class I gene (33). Alternatively, in *M. barkeri* both forms are involved in the incorporation of the rare amino acid pyrrolysine into the enzyme monomethylamine methyltrans-
TABLE 3. Transporters in the genome of N. oceani ATCC 19707

<table>
<thead>
<tr>
<th>Transporter type</th>
<th>Number of genes (% total)</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATP dependent</td>
<td>98 (3.2)</td>
<td>Substrate transport driven by ATP hydrolysis</td>
</tr>
<tr>
<td>Ion channels</td>
<td>10 (0.3)</td>
<td>Energy-independent facilitated diffusion</td>
</tr>
<tr>
<td>PTS system</td>
<td>4 (0.1)</td>
<td>Phosphoenolpyruvate-dependent phosphotransferase</td>
</tr>
<tr>
<td>Secondary transporters</td>
<td>78 (2.5)</td>
<td>Electrochemical potential-driven transporters</td>
</tr>
<tr>
<td>Type II secretion</td>
<td>15 (0.5)</td>
<td>General secretory pathway</td>
</tr>
<tr>
<td>Type IV secretion</td>
<td>14 (0.5)</td>
<td>Conjugal DNA-protein transfer</td>
</tr>
<tr>
<td>Iron uptake</td>
<td>22 (0.7)</td>
<td>Permeases and TonB-dependent Fe-siderophore receptors</td>
</tr>
<tr>
<td>Unclassified</td>
<td>27 (0.9)</td>
<td></td>
</tr>
</tbody>
</table>

a For individual listing of repeats, refer to Table S2 in the supplemental material.

b Total of 3,093 protein-encoding genes.

ferase. This is the first report of the presence of class I and class II LysRS genes together in a bacterial genome (62).

Genes for the synthesis of all five purine and pyrimidine nucleotides are present. In contrast, genes for degradation are very limited. There appears to be no capacity to catabolize nucleotides with the exception of uridine, which can only be converted to pseudouridine.

(vi) Transport. Approximately 263 ORFs in N. oceani, about 9% of the total, are involved with transport. Included are P-P-bond-hydrolysis-driven transporters, electrochemical-potential-driven transporters, and channels/porins from a large number of protein families (Table 3). Predominant among these are 23 ATP binding cassette (ABC)-type transporters (85 genes) for a variety of organic and inorganic substrates, the resistance-nodulation-cell division family (26 genes) of H+ antiport-driven efflux transporters, and a large number of genes (22) involved with iron transport.

The N. oceani annotation lists at least 22 genes involved with iron transport. Iron transport is particularly important due to the number of hemes integral to hydroxylamine oxidoreductase and other cytochromes in the energy-generating NH3 oxidation pathway. Eleven TonB-dependent iron siderophore receptors were identified (Noc_0321, Noc_0323, Noc_0326, Noc_0541, Noc_0859, Noc_1269, Noc_1430, Noc_1489, Noc_1820, Noc_1925, and Noc_2872), including two TonB-dependent receptors for ferrienterochelin (Noc_0859 and Noc_1269) and two for ferrichrome (Noc_0321 and Noc_0323). Unlike Nitrosomonas europaea, where many TonB-dependent receptors are adjacent to genes encoding FeoIR two-component sensor/regulatory proteins involved with iron uptake, this was not the case for the TonB-dependent receptors in N. oceani. No homologues of FecI or FecR were identified in the genome. However, two genes encoding the ferric uptake regulator (Fur) were present (Noc_1194 and Noc_2424), which may serve to regulate the iron siderophore receptor expression. Nitrosoccus oceani has genes for the synthesis of the hydroxamate-type siderophore aerobactin (Noc_1811 to Noc_1814) and an aerobactin receptor (Noc_1820). Additional iron transporters in N. oceani included an ABC-type Fe3+/cobalamin siderophore transport system (Noc_0838 to Noc_0840) and a high-affinity Fe2+/Pb2+ transporter (Noc_0164). While Fe2+ may not be abundant in marine environments, it may be formed in the periplasmic space by a multicopper oxidase. The components of the TonB/ExbB/ExbD-type membrane energy-transducing complex were present in multiple copies (TonB, Noc_0569 and Noc_0610; ExbB, Noc_0142, Noc_0607, Noc_0608, and Noc_2673; ExbD, Noc_0143, Noc_0609, and Noc_2674). No receptors for Fe3+, citrate or Fe3+/coprogen were found. In contrast to N. oceani, the betaproteobacterial AOB Nitrosporas europaea has over 100 genes involved in iron transport, including genes for receptors for Fe3+/coprogen, but does not synthesize any siderophores (16).

In addition to iron transport, a number of uptake systems (67 genes) for other inorganic ions were identified. Inorganic N may be imported either as nitrite via a formate-nitrite (FNT family) transporter (Noc_0109) or via a putative ammonia/ammonium permease (Amount) (Noc_2700 and Noc_2701). Sulfate could be imported via a sulfate permease (SulP family) (Noc_1626), which may function as an SO42−/H+ symport or an SO42−/HCO3− antiport, or by a sodium:sulfate symporter (Noc_1175). Phosphate transporters included two ABC transporters (Noc_2396 to Noc_2399 and Noc_0581 to Noc_0584) and two phosphate-selective porins (Noc_2417 and Noc_2418) (see below). A number of metal ion uptake and efflux systems were identified, including an ABC transporter system dedicated to Mn2+/Zn2+ transport (Noc_2421 to Noc_2423), a CorABC-type Mg2+/Co2+ ion channel of the mitochandrial tricarboxylate carrier family (Noc_0240, Noc_1416, and Noc_2263), three divalent cation transporters for Mg2+/Co2+ and Ni2+ of the MgE family of magnesium transporters (Noc_1840, Noc_1785, and Noc_2801), and three divalent heavy-metal cation transporters (Noc_0092, Noc_04234, and Noc_1342). Efflux systems include two small multidrug resistance (SMR) family transporters of cations and cationic drugs (Noc_0601 and Noc_2235), CopCD copper export proteins (Noc_1741 and Noc_1742), and four cation diffusion facilitator-type cation efflux proteins (Co2+/Zn2+/Cd2+) (Noc_0595, Noc_1534, Noc_1782, and Noc_2871). Other transporters include chloride channel protein ErIC (Noc_0358) and an major facilitator superfamily (MFS)-type cyanate transporter (Noc_1456). Two P-type cation-transporting ATPases were found (Noc_1406 and Noc_2130).

Likely of importance to N. oceani’s salty habitat in the oceans are a number of monovalent cation transporters. Several Na+/H+ antiporter systems were found, including an NhaD type (Noc_1492), an NhaC type (Noc_2134), and two NhaP types (Noc_0159 and Noc_0521). These play important roles in maintaining intracellular pH and conferring salt tolerance. Five genes of the small conductance mechanosensitive ion channel (MscS) family (Noc_0108, Noc_0602, Noc_1030, Noc_1853, and Noc_1914) were found which play a role in fast osmoregulatory responses. Other transport systems identified included the genes encoding both the NAD-binding component (Noc_3020 and Noc_3064) and the membrane component (Noc_0960, Noc_2194, and Noc_2952) of the KefB-type K+ transport systems, a Trk-type K+ uptake system (Noc_0242, Noc_1639, and Noc_1640), and two DASS (divalent anion/sodium ion symporter) family members (Noc_1175 and Noc_2446).

The genome of N. oceani revealed only a few transporters...
for importing organic compounds. ABC-type transporters may import dipeptides/oligopeptides (Noc_1344, Noc_1768, and Noc_2770 to Noc_2773), proline/glycine betaine (Noc_0539 and Noc_0540), and unspecified sugars (Noc_0279 and Noc_0282). Other transporters included an amino acid transporter (Noc_3063), a lactate permease (Noc_1578), and a urea transporter (Noc_2884). A number of potential transport systems for dicarboxylates were found, including a TRAP-type C4-dicarboxylate permease (Noc_0598, Noc_0709, and Noc_0710), a tellurite-resistance/dicarboxylate transporter (TDT) (Noc_0077 and Noc_0542) which may be involved in transporting dicarboxylic acid intermediates, and a DAACS (di/tricarboxylate/amino acid/cation symporter) family members (Noc_1175 and Noc_2446).


*Nitrosococcus oceani* possesses several protein export and secretion systems, including a preprotein translocase (Noc_2305), genes for exporting folded redox proteins via the sec-independent protein secretion system (TatABC; Noc_3058 to Noc_3060), genes encoding the type II general secretion/pilus synthesis pathway, and genes encoding the type IV conjugal DNA/protein transfer system.

*Nitrosococcus oceani* has a cluster of genes encoding parts of a phosphotransferase (PTS)-type sugar transport system. Specifically, genes for E1 (Noc_2800), HPt (Noc_2799), an Ntr-type IIa (Noc_2795), a mannose/fructose type IIa (Noc_2313), as well as an HPt kinase/phosphorylase (Noc_2796) were found. However, genes encoding components IIb, Iic, and IId were absent, making it unlikely that *N. oceani* expresses a functional sugar-transporting PTS system. Instead, as has been seen in other *Proteobacteria*, these genes may form part of a regulatory cascade involving RpoN (13).

**Growth, sensing, responses, and stress tolerance.** (i) **Carbon fixation.** The *N. oceani* genome encodes a form I ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBiSCO). The form I genes (*cbhL* and *cbhS*) occur in an operon with *cbhX* and a hypothetical gene (Noc_0330 to Noc_0333). The operon is preceded by a divergently transcribed *cbhR* gene (Noc_0334) encoding a LysR-type transcriptional regulator. The *cbhX* gene is required for efficient autotrophic growth in *Rhodobacter sphaeroides* (24) and is predicted to be an AAA family ATPase (which can be involved in chaperonin-like functions).

Both prokaryotic/plant-type (Noc_1341) and eukaryotic-type (Noc_1132) carbonic anhydrases are encoded by the genome. There is no obvious candidate for a bicarbonate acquisition system, however. Furthermore, the genome of *N. oceani* lacks genes for carboxysome formation.

Genes for all enzymes to complete the Calvin-Benson-Bassham cycle are present. Transketolase, NAD-dependent glyeraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12), phosphoglycerate kinase, pyruvate kinase, and fructose-1,6-bisphosphate aldolase are encoded by an operon (Noc_2804 to Noc_2808), whereas fructose-1,6-/sedoheptulose-1,7-bisphosphatase (Noc_0021), ribose-5-phosphate isomerase (Noc_2667), and phosphoribulokinase (Noc_2826) are encoded by isolated genes. Genes encoding ribulose-5-phosphate 3-epimerase and phosphoglycolate phosphatase are grouped in an operon (Noc_2492 to Noc_1493).

Due to the aforementioned homology of ammonia monooxygenase and particular methane monoxygenases as well as the hypothesized lateral transfer of genes, the genome of *N. oceani* was investigated for the presence of *C*1 metabolic pathways (see Table 4). Whereas the *N. oceani* genome does not contain the genes needed to fix carbon directly from methane or methanol via the ribulose monophosphate or serine pathways, it contains the genetic inventory to funnel *C*1 carbon into the Calvin-Benson-Basham cycle (Table 4). Remarkably, two independent pathways for formaldehyde oxidation to formate (Noc_1394, Noc_1440, and Noc_2006) and two gene clusters encoding the capacity for formate dehydrogenation (Noc_1122 to Noc_1124 and Noc_2559 to Noc_2561) were identified. The identification of this inventory provides a theoretical explanation of earlier findings that labeled C1-carbon was assimilated into the biomass of *N. oceani* cultures growing in an ammonia-dependent fashion (36, 71).

(ii) **Phosphorus, nitrogen, and sulfur cycling.** *Nitrosococcus oceani* appears to be an efficient phosphorus sink in the oceans (Fig. 3). Some of the uptake and processing capacity for phosphorus is arranged in the *N. oceani* genome as a supercluster of genes, starting with gene Noc_2388, encoding polyphosphate kinase (EC:2.7.4.1), followed by a string of eight genes (Noc_2394 to Noc_2401) encoding exopolyphosphatase (EC 3.6.1.11), regulatory protein PhoU, the high-affinity binding protein-dependent ABC transporter PstBACS, and the two-component system PhoRB. The genome also encodes an inorganic pyrophosphatase (Noc_1134; EC 3.6.1.11; COG0221) and a V-type H+-translocating pyrophosphatase (Noc_1901; COG3808). A gene homologue encoding the inorganic phosphate transporter PitA was not found in the *N. oceani* genome. Therefore, like in several archaea and yeasts, polyphosphate-copper complexes formed to control copper concentrations at levels above toxicity may be removed from the cytoplasm by a phosphate:proton symport protein in the major facilitator superfamily unrelated to PitA (63). Taken together, *N. oceani* appears to have the capacity to utilize stored polyphosphate molecules as phosphagens for ATP synthesis, for substrate phosphorylation and the regulation of enzyme activity, for the direct generation of the proton-motive force via pyrophosphate, and to detoxify copper (Fig. 3).

As a nitrifying bacterium, *N. oceani* affects the nitrogen cycle in its marine environment by assimilatory and dissimilatory activities (Fig. 4). While the ammonia oxidation capacity (AMO and HAO; see above) of AOB links the oceanic pools of reduced (ammonia) and oxidized (nitrite) nitrogen, their
classical denitrification capacity through dissimilatory nitrite reductase (NirK; EC 1.7.2.1; Noc_0089) and nitric oxide reductase (Nor; Noc_1847 to Noc_1851) is likely a major source of nitrous oxide emitted from the oceans (20, 21, 57). Additionally, there are various genes that encode putative cytoplasmic (Noc_0889 and Noc_2605) and exported (Noc_1542) multicopper oxidases (MCO), as well as other MCOs that have been implicated in oxidation of NO, such as the pan1-type MCO (Noc_0889), an alternative NO reductase (6). In contrast to its organization in the genome of Nitrosomonas europaea, where it is clustered in a four-gene operon with the NirK-type nitrite reductase (6, 16), the pan1-type MCO gene Noc_0889 is clustered, but not necessarily in the same transcriptional unit, with a gene encoding cytochrome P460 (Noc_0890) directly upstream of the hao gene cluster (7). Cytochrome P460 has been implicated in hydroxylamine detoxification in several bacteria, including N. oceanii (9, 10, 31, 79).

Because cytochrome P460 can be reduced by NO and because of the physical proximity of the two genes in the N. oceanii genome, P460 and pan1-type MCO may be components of an alternative N-oxidation pathway that, in contrast to HAO, produces nitrite with NO as an intermediate. Considering the toxicity of NO and the fact that AOB have a high O2 consumption rate but tend to live near the oxic/anoxic interface, additional NO detoxification mechanisms that are operational under low-O2 conditions are likely necessary for survival of AOB.

### TABLE 4. C-1 carbon metabolism in the genome of N. oceanii ATCC 19707 in comparison with the genomes of the methanotroph M. capsulatus (Bath) and the beta-AOB N. europaea

<table>
<thead>
<tr>
<th>Process</th>
<th>Protein</th>
<th>Noc_0089 gene(s)</th>
<th>M. capsulatus gene(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Methane oxidation to methanol</td>
<td>AMO/pMMO (EC 1.13.12.−)</td>
<td>Noc_2501 to Noc_2503</td>
<td>MCA1796 to MCA1798, MCA0295, MCA2853 to MCA2855</td>
</tr>
<tr>
<td>Methanol oxidation to formaldehyde (FA)</td>
<td>Soluble methane monoxygenase (EC 1.14.13.25)</td>
<td>Not present</td>
<td>MCA1194 to MCA1205</td>
</tr>
<tr>
<td>Methanol dehydrogenase cluster (EC 1.1.99.8)</td>
<td>Not present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Protein with pyrrolo-quinoline repeat domain (EC 1.1.99,−)</td>
<td>Not present</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pyrrolo-quinoline quinone proteins</td>
<td>Noc_0821</td>
<td>MCA2891</td>
<td></td>
</tr>
<tr>
<td>FA oxidation to formate</td>
<td>Dye-linked FA DH* (EC 1.2.99.93)</td>
<td>Noc_2006</td>
<td>MCA2155</td>
</tr>
<tr>
<td>Glutathione (GSH)-dependent pathway</td>
<td>GSH-dependent FA DH (EC 1.2.1.1)</td>
<td>Noc_1394</td>
<td>Not present</td>
</tr>
<tr>
<td>S-hydroxy methyl-GSH to S-formyl GSH</td>
<td>GSH S-transferase (EC 3.1.2.12)</td>
<td>Noc_1440</td>
<td>Not present</td>
</tr>
<tr>
<td>Tetrahydrofolate (THF)-dependent pathway</td>
<td>5,10-Methylene THF reductase (EC 1.7.99.5)</td>
<td>Noc_2680</td>
<td>MCA0137</td>
</tr>
<tr>
<td>Condensation of THF with formaldehyde</td>
<td>Methylene THF DH (EC 1.5.99.9)</td>
<td>Not present</td>
<td>MCA0508</td>
</tr>
<tr>
<td>Methylene THF to methenyl THF</td>
<td>Methylene THF DH (EC 1.5.1.5)</td>
<td>Not present</td>
<td>MCA3018, MCA3019</td>
</tr>
<tr>
<td>Methenyl THF to formyl THF</td>
<td>Methenyl THF cyclohydrolase (EC 3.5.4.9)</td>
<td>Not present</td>
<td>MCA0507</td>
</tr>
<tr>
<td>Formyl THF to formate</td>
<td>Formyl THF hydrolase (EC 3.5.1.10)</td>
<td>Noc_1789</td>
<td>Not present</td>
</tr>
<tr>
<td>Tetrahydromethanopterin (THMPT)-dependent pathway</td>
<td>5-Formyl THF cycliglase (EC 6.3.3.2)</td>
<td>Not present</td>
<td>MCA2773</td>
</tr>
<tr>
<td>Condensation of THMPT with FA</td>
<td>Formaldehyde activating enzyme (EC 4.3,−,−)</td>
<td>Not present</td>
<td>MCA2778</td>
</tr>
<tr>
<td>Methylene THMPT to methenyl THMPT</td>
<td>Methylen THMPT DH (EC 1.5.99.9)</td>
<td>Not present</td>
<td>MCA0508</td>
</tr>
<tr>
<td>Methenyl THMPT to formyl THMPT</td>
<td>Methylen THMPT cyclohydrolase (EC 3.5.4.27)</td>
<td>Not present</td>
<td>MCA3018, MCA3019</td>
</tr>
<tr>
<td>Formyl THMPT to formate</td>
<td>Formaldehyde activating enzyme /hydrolase complex (EC 1.2.99.5 and EC 2.3.1.101)</td>
<td>Not present</td>
<td>MCA2863</td>
</tr>
<tr>
<td>Formate oxidation to CO2</td>
<td>Formate dehydrogenase (EC 1.2.1.2)</td>
<td>Noc_1122 to Noc_1124</td>
<td>MCA1391 to MCA1393</td>
</tr>
<tr>
<td>C-1 assimilation</td>
<td>Formate dehydrogenase (EC 1.2.1.43)</td>
<td>Noc_2559 to Noc_2561</td>
<td>MCA2576 to MCA2577</td>
</tr>
<tr>
<td>Ribulose monophosphate pathway</td>
<td>Hexulose 6-phosphate synthase (EC 4.1.2,−)</td>
<td>Not present</td>
<td>MCA3043, MCA3049</td>
</tr>
<tr>
<td>Serine cycle</td>
<td>Hexulose 6-phosphate isomerase (EC 5.3,−,−)</td>
<td>Not present</td>
<td>MCA3044, MCA3050</td>
</tr>
<tr>
<td>Malate dehydrogenase (EC 1.1.1.37)</td>
<td>Malyl-coenzyme A synthetase (EC 6.2.1.9)</td>
<td>Not present</td>
<td>MCA0610</td>
</tr>
<tr>
<td>Malyl-coenzyme A lyase (EC 4.1.3.24)</td>
<td>Malyl-coenzyme A lyase (EC 4.1.3.24)</td>
<td>Not present</td>
<td>Not present</td>
</tr>
</tbody>
</table>

a DH, dehydrogenase.
It is thus not surprising to find in the genome of *N. oceani* a gene, *cycP*, that encodes a beta-sheet-structured cytochrome *c*’-beta (“c’-beta”; Noc_2696). Alpha-helical *c*’ cytochromes have been implicated in microaerobic NO sequestration and dehydrogenation by strains of the betaproteobacterial pathogen genus *Neisseria* (18, 19, 52, 64, 67). It has been proposed only recently that cytochrome *c*’-beta, which is evolutionarily related to and a putative redox partner of cytochrome P460 (8), has evolved from an alpha-helical monoheme cytochrome *c* that is ancestral to both *c*’- and *c*’-beta cytochromes (M. G. Klotz and A. B. Hooper, unpublished results). Both enzymes, *c*’-beta and P460, are also likely redox partners of the soluble periplasmic di-heme cytochrome *c* peroxidases. The P460-heme coordination site, found in HAO and cytochrome P460, is highly sensitive to hydrogen peroxide (32); therefore, it is interesting that the c’-beta-encoding gene is clustered with one of the three di-heme cytochrome *c* peroxidase genes (Noc_2697), suggesting a dedicated protective function of key periplasmic enzymes by this peroxidase. The genome of *N. oceani* also contains a gene encoding the red copper protein, nitrosocyanin (Noc_1090), a putative enzyme with a cupredoxin fold (3, 55). Because this gene and the encoded protein have so far been uniquely found in a beta-AOB (3, 16) and now also in a gamma-AOB, its putative catalytic function is likely involved in and specific to ammonia-oxidizing catabolism. It appears that the catabolic dependence on ammonia oxidation and the ultimate production of reactive and toxic NOX intermediates have imparted selective pressure on all AOB to maintain this suite of periplasmic enzymes, because a similar complement of genes, albeit with different genomic organization, was also identified in the genome of *Nitrosomonas europaea* (16). It appears that the strategy of maintaining a complement of NOX-detoxifying enzymes in the periplasm is similar to the...
strategy of active oxygen defense and is designed to avoid the formation and presence of nitrosating agents such as nitrous anhydride in the cytoplasm, where they have mutagenic activity (75).

Ammonia is also the source for nitrogen assimilation, and \textit{N. oceani} has the complete capacity for low (glutamate dehydrogenase)- and high (glutamine synthetase; glutamine oxoglutarate aminotransferase/glutamate synthase)-affinity ammonia assimilation. Small- and large-subunit glutamate synthase-encoding genes were contiguous (Noc\_1603 and Noc\_1604). In addition, two genes encoding putative NADPH-ferredoxin-dependent glutamate synthase large-chain proteins (Noc\_2957 and Noc\_0101) were identified. Two NADPH-specific forms of glutamate dehydrogenase were identified (Noc\_2054, EC 1.4.1.3; Noc\_0864, EC 1.4.1.4). The GS-GOGAT (glutamine synthetase–glutamine oxoglutarate aminotransferase) system presumably functions at lower concentrations of ammonia. A type I glutamine synthetase (GSI; EC 6.3.1.2) is encoded by the \textit{glnA} gene (Noc\_2652). To avoid futile cycling, GSI activity is likely regulated by adenylation; the adenyllyl-transferase encoded by gene \textit{glnE} was identified (Noc\_0135). Additional regulatory proteins encoded by \textit{glnB} (Noc\_0715) and \textit{glnD} (Noc\_0806) encoding a \textit{P} \textit{U}, uridylyl-transferase (EC 2.7.7.59) were also identified. Because these genes have been identified in the \textit{N. oceani} genome, the regulation of \textit{N} uptake is likely dependent upon the ratio of glutamine and glutamate and proceeds via adenyllylation of GS and uridylylation of proteins \textit{P} \textit{U} and \textit{P} \textit{U}, as has been experimentally determined for many other \textit{Gammaproteobacteria} (50).

As a member of the \textit{Chromatiaceae}, \textit{N. oceani} should participate in the sulfur cycle beyond acquiring sulfur for biosynthesis. The identified sulfate uptake capacity allows \textit{N. oceani} to acquire and process sulfate. Sulfate reduction may proceed via sulfate adenylyltransferase (\textit{cysND}; Noc\_2288 and Noc\_2289; EC 2.7.7.4) to adenosine phosphosulfate via adenylylsulfate kinase (\textit{cysC}; Noc\_2482; EC 2.7.1.25) to phosphoadenosine phosphosulfate and via phosphoadenosine phosphosulfate reductase (\textit{cysH}; Noc\_2290; EC 1.8.4.8) to sulfite, which may be further reduced to \textit{H}_2\textit{S} by an NADPH-dependent sulfite reductase, EC 1.8.1.2 (alpha-subunit CysI, Noc\_1305; beta-subunit CysI, Noc\_1306). Dihydrogen sulfide is required for cysteine biosynthesis, and genes encoding a thioredoxin-disulfide reductase (Noc\_0345; EC 1.8.1.9) and a thiol-disulfide interchange protein (Noc\_0551) were identified, as was the gene encoding thiosulfate sulfurtransferase (Noc\_0593; EC 2.8.1.11). The genome of \textit{N. oceani} also contains a gene cluster that encodes a putative polysulfide reductase (\textit{psr}; Noc\_1238 to Noc\_1240), a monoheme cytochrome (\textit{cccA}, Noc\_1241), a transporter (Noc\_1242), a cytochrome c oxidase (Noc\_1244 to Noc\_1247), and the five genes encoding the ubiquinone complex (Noc\_1248 to Noc\_1252). This cluster of genes is absent from the genomes of \textit{Nitrinosomas europaea}, \textit{Nitratosoma eu tropha}, \textit{Nitrospira multiformis}, and \textit{Nitrobacter winogradskyi} but is conserved (sequence and synteny) in the genomes of \textit{Nitrobacter hamburgensis}, \textit{Rhodopirellula baltica}, \textit{Cytophaga hutchinsonii} ATCC 3406, and \textit{Cupriavidus necator} (\textit{Ralstonia eutropha} JMP134). Functional polysulfite reductase is a molybdopterin oxidoreductase complex that has been experimentally described to act as a quinone oxidase in \textit{Wolinella succinogenes} (42, 43). Analysis of the \textit{N. oceani} genome did not reveal a molybdopterin guanine dinucleotide-binding protein subunit-encoding gene in the vicinity of the cluster. This putative polysulfide reductase activity awaits experimental verification in \textit{N. oceani}, which would indicate the residence of an alternative catalytic center for polysulfide reduction. If present, this molybdopterin oxidoreductase could theoretically also be involved in the anaerobic reduction of nitrate, chlorate, selenate, or other highly oxidized minerals, thereby accommodating electron disposal in a microaerophilic environment near the oxic/anoxic interface.

In contrast to many other purple sulfur bacteria, the genome of \textit{N. oceani} lacked genes for the formation of internal or external granules of sulfur compounds.

\textbf{(iii) Cellular growth and motility.} The genome of \textit{N. oceani} contains almost all of the typical complement of genes with an identified role in cell cycle and division of other \textit{Gammaproteobacteria}, such as \textit{Escherichia coli}. The genome clearly lacks the genes encoding FtsEX (involved in localization and stabilization of the septal ring), FtsN, and \textit{SulA}. On the other hand, the genome contains three genes (Noc\_0272, Noc\_1903, and Noc\_2569) encoding proteins with domains (conserved zinc-binding motif \textit{HEXXH}, ATPase domain, and peptidase domain) matching the cell division metalloprotease \textit{FtsH} (COG0465). An alignment of all three \textit{FtsH} proteins showed that they differ in sequence mostly at their \textit{N} and \textit{C} termini, which flank the ATPase and peptidase domains.

Genes encoding flagellation and motility are in the \textit{N. oceani} genome in two large clusters (Noc\_2354 to Noc\_2378 and Noc\_2155 to Noc\_2166) and several smaller clusters (Noc\_0833, Noc\_0834, Noc\_0124 to Noc\_0131, Noc\_2052, Noc\_2053, and Noc\_2683 to Noc\_2685). The master switch operon \textit{flhCD} was not identified and is likely absent, as it is from the genomes of other AOB. As a likely adaptation to life in the open ocean, \textit{N. oceani} appears to have only limited chemotactic capacity, because just one methyl-accepting chemotaxis protein (MCP) (Noc\_0128) of the \textit{PilJ} type was identified. In contrast, the genome of \textit{Nitrinosomas europaea} contained three MCPs (16). The presence of a sodium-driven polar flagellar motor protein (MotA; Noc\_0833), which can assemble with the product of the adjacent \textit{pomB} gene (Noc\_0834) in addition to the usual pmf-dependent flagellar rotation mechanism, may be an adaptation to \textit{N. oceani}’s marine lifestyle.

\textbf{(iv) Two-component systems.} Considering the reductive evolution of AOB as concluded from analysis of the \textit{Nitrinosomas europaea} genome (16), the genome of \textit{N. oceani} harbored an impressive complement of complete two-component systems. The genome contained 13 genes encoding histidine protein kinases (HPK), of which 12 were paired (mostly succeeded) by a response regulator (RR). In addition, 1 HPK- and 11 RR-encoding orphaned genes were identified (Table 5 and Table S3 in the supplemental material). One of the HPK genes (Noc\_1756) was succeeded by a tandem of two RR genes (Noc\_1757 and Noc\_1758). The \textit{N. oceani} genome also contained six hybrid genes whose deduced protein sequence contained HPK and RR domains (Table 5). One of these hybrid-HPK genes (Noc\_1700) was paired with a gene encoding an RR in the LuxR/FixJ family (Noc\_1701) and is adjacent to other HPK- and RR-encoding genes. This suggests that the RR domain in the hybrid kinase
has a regulatory phosphotransferase function in a phosphorylation cascade.

(v) Stress tolerance. In comparison with other AOB, the _N. oceani_ genome contains only a limited inventory that contributes to stress tolerance in general and oxidative stress tolerance in particular. The genome encodes a heme-containing monofunctional large-subunit catalase (KatE; Noc_1165) and an iron-containing superoxide dismutase (Fe-SOD; Noc_2428), both of which are supplied with iron and heme by bacterioferritin (Bfr; Noc_1411). The genome also contains genes encoding glutaredoxin (Noc_2427), thioredoxin (TRX; Noc_0603 and Noc_2583), and a thioredoxin-dependent peroxide reductase (AhpC; peroxiredoxin; Noc_1307), but it lacks genes for bacterioferritin-comigratory protein, NADH-peroxidoxin reductase (AhpF), glutathione oxidoreductase, and other isozymes of hydroperoxidases (KatG, KatA, and Mn-Cat) and SOD (Mn-SOD and Cu/Zn-SOD). Like _Nitrosomonas europaeca_, the genome also lacks genes for OxyR redox-autoregulatory protein, which regulates oxidative stress tolerance (KatG, AhpC, and SOD), iron and zinc transport proteins (Fur and Zur), and the stationary-phase-specific sigma factor RpoS in many bacteria. In contrast to _Nitrosomonas europaeca_, which lacks an RpoS gene, the _N. oceani_ genome contains two genes encoding RpoS (sigma 38; Noc_0183 and Noc_1702). RpoS is known to regulate hydroperoxidase (KatE) and the cell shape protein BolA (Noc_2387). _Nitrosococcus oceani_ seems minimally prepared to respond to other stresses. In addition to two genes that encode the minimal growth sigma factor RpoD (sigma 70; Noc_0045 and Noc_2066), the genome contains genes that encode alternative sigma factors involved in heat (RpoH; sigma 32; Noc_1935) and extreme heat (RpoE; sigma 24; Noc_2463) stresses, nitrogen starvation (RpoN; sigma 54; Noc_2793), and the need to move by flagellar motility (Flia; sigma 28; Noc_2155). Despite the absence of the flagellar master operon (_flagI_CD_), _Flia_-dependent regulation of a complete complement of flagella synthesis and chemotaxis gene clusters is aided by the presence of only one MCP of the pseudomonad PilJ type (with MA and HAMP domains; Noc_0128), whereas other MCPs in the Tar (CheM), Tsr (CheD), Tap, and Aer categories were absent from the genome. In addition, the _N. oceani_ genome did not contain genes with significant similarity to the two-component regulatory systems of LasRI/RhlR involved in homoserine lactone autoinducer synthesis (quorum sensing) as well as the regulation of motility, virulence, starvation response, and iron homeostasis in several _Gammaproteobacteria_. The ferric uptake regulation protein (Fur; Noc_1194) regulates, for instance, ferric citrate (FecIR) and ferrichrome (_fhu_ operon) transport, exotoxin synthesis, and the expression of hydroperoxidases in many proteobacteria. Interestingly, the zinc uptake regulation protein (Zur; Noc_2424) was found adjacent to a gene cluster that encodes a binding protein-dependent zinc ABC transporter system (Noc_2421 to Noc_2423) in the genome.

**Conclusions.** _Nitrosococcus oceani_ is one of only two known ammonia-oxidizing bacteria classified as _Gammaproteobacteria_, while the large majority of isolated ammonia-oxidizing bacteria are classified as _Betaproteobacteria_. The genome sequence of the betaproteobacterium _Nitrosomonas europaeca_ is available and facilitates a comparison of the genes most similar between these two bacteria. Of the 224 genes in _N. oceani_ that were most similar to genes from _Nitrosomonas europaeca_, 76 were classified as hypothetical or proteins of unknown function. These hypothetical and unknown proteins are slightly overrepresented (34%) in this subset of genes relative to genes without function prediction in the complete _N. oceani_ (31.5%) and _Nitrosomonas europaeca_ (29%) genomes. Whereas the number of genes without function and without similarity to other known genes (hypothetical ORFs) is 10 times higher in the _Nitrosomonas europaeca_ (4.7%) versus the _N. oceani_ (0.57%) genome, the number of genes without function but with similarity to genes in other genomes (conserved hypothetical ORFs) was significantly higher in the _N. oceani_ genome (30.9%) than in _Nitrosomonas europaeca_ (24.4%). For those genes assigned a function, the functions included central carbon and nitrogen metabolism, electron transport, gene regulation, and transport. The _Nitrosomonas europaeca_-like genes were distributed throughout the _N. oceani_ genome. Most of these genes were not flanked by additional _Nitrosomonas europaeca_-like genes, though in some cases two or three such genes were contiguous. A cluster of _Nitrosomonas europaeca_-like genes was present from Noc_1955 to Noc_1986, where 20 of these 31 genes were most similar to _Nitrosomonas europaeca_ genes. As additional genome sequences become available, it will be of interest to carry out more detailed comparisons of the genes in common among the ammonia-oxidizing bacteria. Of particular interest will be the uncharacterized genes, some of which may encode functions unique to the use of ammonia as a growth substrate, whereas others may be responsible for the difference between marine and freshwater sediment soil AOB in their abilities to form nitrification consortia (e.g., ammonia- and nitrate-oxidizing bacteria). Progress in ongoing and future research with _nitrosococci_ will be made available at the _Nitrosococcus_ Project website (http://nitrosococcus.org).

**ACKNOWLEDGMENTS**

Sequencing was funded by the U.S. Department of Energy’s (DOE) Office of Biological and Environmental Research and was carried out primarily at the Joint Genome Institute. Finishing was completed at the Lawrence Livermore National Laboratory under the auspices of U.S. DOE contract W-7405-ENG-48. A.F.S., A.T.P.P., and M.G.K. were supported, in part, by NSF grant EF-040621 and incentive funds provided by the University of Louisville.

Computational annotation was carried out at the Oak Ridge National Laboratory. Our thanks go to high school student Ariella Bar-
Cycles 17.
58. Nicol, G. W., and C. Schleper. 2006. Ammonia-oxidising Crenarchaeota: 
important players in the nitrogen cycle? Trends Microbiol. 14:207.
biochemical analysis of cytochrome 
60. Remonsellez, F., A. Oreil, and C. A. Jerz. 2006. Copper tolerance of the 
thermoacidophilic archaeon Sulfolobus metallicus: possible role of polyphos-
Camp, J. G. Kuenen, and M. S. M. Jetten. 2002. Anaerobic ammonia oxida-
tion in the presence of nitrogen oxides (NOₓ) by two different lithotrophs. 
63. Schmidt, L., R. J. M. van Spanning, and M. S. M. Jetten. 2004. Denitrifica-
tion and ammonia oxidation by Nitrosomonas europaea wild-type, and NirK-
and NorB-deficient mutants. Microbiology 150:4107–4114.
64. Turner, S. M., J. W. Moor, L. Griffiths, T. W. Overton, H. Smith, and J. A. 
New York, N.Y.
bacteria of the beta-subclass of the class proteobacteria in aquatic samples 
68. Ward, B. B. 1982. Marine ammonium-oxidizing bacteria: abundance and 
activity in the northeast Pacific Ocean. Ph.D. dissertation. University of 
Washington, Seattle, Wash.
D. L. Kirchman (ed.), Microbial ecology of the oceans. Wiley-Liss, New 
York, N.Y.
coccus oceanus, a marine ammonia-oxidizing gamma-proteobacterium, 
detected by PCR and sequencing of 16S rRNA and amoA genes. Appl. Envi-
suprasperon in Pseudomomas exhibits gene organization features of both 
of the tryptophan operon and the dynamics of evolutionary change. Micro-
oceanus in a Mediterranean lagoon by immunofluorescence. J. Appl. Bacte-
riol. 80:611–616.
a mine by cytochrome P-460 of the obligate methylotroph Methylococcus 
perspectives on processes and paradigms. Appl. Environ. Microbiol. 68: 
1015–1024.