

# Phylogeny and Functional Expression of Ribulose 1,5-Bisphosphate Carboxylase/Oxygenase from the Autotrophic Ammonia-Oxidizing Bacterium *Nitrosospira* sp. Isolate 40KI

Janne B. Utåker,\* Kjell Andersen, Ågot Aakra, Birgitte Moen, and Ingolf F. Nes

Laboratory of Microbial Gene Technology, Department of Chemistry and Biotechnology, Agricultural University of Norway, N-1432 Ås, Norway

Received 12 July 2001/Accepted 11 October 2001

The autotrophic ammonia-oxidizing bacteria (AOB), which play an important role in the global nitrogen cycle, assimilate CO<sub>2</sub> by using ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO). Here we describe the first detailed study of RubisCO (*cbb*) genes and proteins from the AOB. The *cbbLS* genes from *Nitrosospira* sp. isolate 40KI were cloned and sequenced. Partial sequences of the RubisCO large subunit (CbbL) from 13 other AOB belonging to the β and γ subgroups of the class *Proteobacteria* are also presented. All except one of the β-subgroup AOB possessed a red-like type I RubisCO with high sequence similarity to the *Ralstonia eutropha* enzyme. All of these new red-like RubisCOs had a unique six-amino-acid insert in CbbL. Two of the AOB, *Nitrosococcus halophilus* Nc4 and *Nitrosomonas europaea* Nm50, had a green-like RubisCO. With one exception, the phylogeny of the AOB CbbL was very similar to that of the 16S rRNA gene. The presence of a green-like RubisCO in *N. europaea* was surprising, as all of the other β-subgroup AOB had red-like RubisCOs. The green-like enzyme of *N. europaea* Nm50 was probably acquired by horizontal gene transfer. Functional expression of *Nitrosospira* sp. isolate 40KI RubisCO in the chemoautotrophic host *R. eutropha* was demonstrated. Use of an expression vector harboring the *R. eutropha cbb* control region allowed regulated expression of *Nitrosospira* sp. isolate 40KI RubisCO in an *R. eutropha cbb* deletion strain. The *Nitrosospira* RubisCO supported autotrophic growth of *R. eutropha* with a doubling time of 4.6 h. This expression system may allow further functional analysis of AOB *cbb* genes.

Ammonia-oxidizing bacteria (AOB) play an important role in the global nitrogen cycle by oxidizing ammonia to nitrite. They are aerobic and obligately chemoautotrophic organisms, gaining all of the energy that they need for growth and maintenance by oxidation of ammonia and relying on CO<sub>2</sub> as the main source of cell carbon. This metabolism is energetically unfavorable, considering the high redox potential of NO<sub>2</sub><sup>-</sup>/NH<sub>4</sub><sup>+</sup> (E<sub>0</sub>' = 340 mV) and the inefficient and energy-consuming CO<sub>2</sub> assimilation process used. However, the combination of NH<sub>3</sub> and CO<sub>2</sub> as growth substrates is unique (10), placing the AOB in an ecological niche in which they encounter little competition.

Phylogenetically, the AOB are divided into two main groups: (i) the marine nitrosococci in the γ subgroup of the class *Proteobacteria* and (ii) the nitrosomonads, the *Nitrosococcus mobilis* strains, and the nitrosospirads in the β subgroup of the *Proteobacteria*. The majority of AOB studied so far belong to the β subgroup. Since 1984 (51) extensive classification work has been performed with these bacteria, involving 16S ribosomal DNA (rDNA) sequencing, DNA-DNA hybridization, and 16S-23S intergenic spacer sequencing (1, 26, 32, 46). In general, phylogenies based on *rm* sequences show very close relationships among AOB belonging to the β subgroup and result in correspondingly low resolution in phylogenetic trees.

Therefore, for some time researchers questioned whether the clustering in phylogenetic trees reflected true differences between isolates. Recently, however, DNA-DNA hybridization experiments have largely confirmed the relationships among the AOB belonging to the β subgroup of the *Proteobacteria* (3).

Most autotrophic organisms assimilate CO<sub>2</sub> by the Calvin-Benson-Bassham (CBB) cycle, and ribulose 1,5-bisphosphate carboxylase/oxygenase (RubisCO) is one of the key enzymes in this cycle. Previously, no RubisCO gene sequence or gene expression studies have been performed with AOB, although RubisCO activity has been determined in a few strains (25, 45). Electron microscopy has revealed the presence of carboxy-some-like structures in some, but not all, strains of the AOB (11, 29, 32, 52). This suggests that CO<sub>2</sub> is fixed in the CBB cycle by a type I RubisCO.

In nature, there are at least two forms of RubisCO, type I and type II. The recently characterized RubisCO from members of the *Archaea* may represent an additional form (35, 50). Type I RubisCO is found in plants, algae, and most microorganisms. Type II RubisCO is less common and is exclusively microbial (49). Type I RubisCO is a hexadecamer of eight large and eight small subunits (L<sub>8</sub>S<sub>8</sub>); the *M<sub>r</sub>*s of the large and small subunits are ~55,000 and 15,000, respectively. The *M<sub>r</sub>* of the hexadecamer (the native enzyme) is about 550,000 (49). Type I enzymes can be subdivided into the following two main groups: green-like enzymes, which are present in plants and green algae, and red-like enzymes, which are present in red algae and a wide variety of other algae. Both types are widespread among prokaryotes (14, 42, 49). The levels of amino acid sequence identity between type I and type II large sub-

\* Corresponding author. Mailing address: Laboratory of Microbial Gene Technology, Department of Chemistry and Biotechnology, Agricultural University of Norway, P.O. Box 5051, N-1432 Ås, Norway. Phone: 47 64949471. Fax: 47 64948596. E-mail: janne-beate.utaker@ikb.nlh.no.

units are only 20 to 30%, but the three-dimensional structures of the large subunits are nevertheless quite similar (24). RubisCO catalytic efficiencies are quite low and often growth limiting, but they can differ considerably, even in closely related species (27). There is currently great interest in gaining a better understanding of the molecular basis of the CO<sub>2</sub>-O<sub>2</sub> substrate specificity of RubisCO. Comparative analyses of RubisCOs from AOB, which have demanding aerobic chemoautotrophic metabolism, are therefore of considerable interest.

Both type I RubisCO and type II RubisCO are highly conserved, even in different kingdoms (49), and the phylogeny of the RubisCO large subunits has therefore been used to describe the phylogeny of organisms, particularly plants and algae. The rather convoluted and intriguing evolutionary history of RubisCO, which includes several horizontal gene transfers, is becoming apparent. For example, closely related pairs of photosynthetic bacteria in which one bacterium has a green-like RubisCO and the other has a red-like RubisCO are known (37). Compared to the phylogeny of rRNA genes, RubisCO phylogeny thus reveals interesting discrepancies which may shed new light on the evolution of autotrophic organisms (14, 49).

The genes encoding RubisCO in microorganisms are usually clustered in operons containing other Calvin cycle-related genes, all of which are designated *cbb* genes (33, 41, 42). Type I RubisCO large and small subunits are encoded by the *cbbL* and *cbbS* genes, respectively. All known *cbb* operons are positively regulated by a LysR type of DNA-binding protein, CbbR (40), whose gene in most cases is upstream from the other *cbb* genes and is transcribed divergently (33, 41).

The facultative chemoautotroph *Ralstonia eutropha* (formerly *Alcaligenes eutrophus*) is a member of the  $\beta$  subgroup of the *Proteobacteria* that is closely related to the AOB in the  $\beta$  subgroup (26). The three-dimensional structure of the RubisCO of *R. eutropha* was determined recently (23), and this enzyme is now one of the most extensively characterized microbial RubisCOs. This provided the basis for our study of RubisCOs in the AOB. In this study, the genes encoding RubisCO (*cbbLS*) in *Nitrosospira* sp. isolate 40KI were isolated and sequenced. In addition, partial *cbbL* sequences from 13 other isolates of AOB were determined and analyzed phylogenetically. Furthermore, regulated expression of cloned *cbbLS* genes from *Nitrosospira* sp. isolate 40KI in an *R. eutropha* *cbbLS* deletion strain is described here.

## MATERIALS AND METHODS

**Bacterial strains and growth conditions.** *Nitrosospira* sp. isolates 40KI, III2, III7, L115, and AF (29, 46), *Nitrosospira multififormis* NI13 (= ATCC 25196), *Nitrosomonas europaea* Nm50 (= ATCC 25978), *Nitrosococcus halophilus* Nc4 (31), *Nitrosococcus mobilis* Nc3 (H.-P. Koops, unpublished data), and *Nitrosomonas* sp. isolate F5 (2) were grown as described previously (29, 31, 46). The *R. eutropha* (formerly *Alcaligenes eutrophus*) strains used were wild-type strain ATCC 17697 and a *cbb* deletion derivative of this strain, AE951. In AE951, almost the entire *cbbL* and *cbbS* coding region has been deleted; only the first 17 N-terminal amino acid residues of CbbL and the four C-terminal residues of CbbS are intact. Construction of AE951 will be described in detail elsewhere. AE951 produces no RubisCO protein, and autotrophic growth of this strain is completely blocked. *Escherichia coli* DH5 $\alpha$  (Bethesda Research Laboratories) was used to maintain plasmids and as a donor in conjugation experiments. The growth media and conditions used for *E. coli* and *R. eutropha* have been described previously (5). *R. eutropha* was grown autotrophically in liquid cultures at 30°C in 500-ml flasks closed with septum stoppers and filled with an atmosphere

containing 20% O<sub>2</sub>, 5% CO<sub>2</sub>, and 75% H<sub>2</sub>. The flasks contained 20 ml of medium and were shaken at 250 rpm. Syringes were used to withdraw samples or to replenish the gas.

**PCR primers and conditions.** The nucleotide sequences of the primers used for amplification of a conserved region in *cbbL* are available on request. The majority of the primers used are identical to *R. eutropha* ATCC 17707 *cbbLS* gene sequences (GenBank accession no. M17744). In addition, a few degenerate primers were constructed from a broad alignment of *cbbL* sequences with the CODEHOP program (38). PCRs were performed directly with cells lysed by a freeze-thaw method described elsewhere (26, 47). Therefore, the amount of template DNA in each reaction mixture was not determined, but in general ~1  $\mu$ l of lysed cells was sufficient for a 100- $\mu$ l PCR mixture. The PCR conditions used were as follows: 97°C hot start for 3 min and then 35 cycles of denaturation at 94°C for 1 min, annealing at 50 to 58°C (the optimum temperature varied with the primer and the bacterial strain) for 1 min, and elongation at 72°C for 1 min. A final step consisting of 72°C for 5 min was included to fill in incomplete ends. The primer concentration varied between 100 and 250 nM depending on the source of the DNA. For all PCRs the 100- $\mu$ l reaction mixture contained each deoxynucleoside triphosphate (Boehringer) at a concentration of 200  $\mu$ M, 1.5 mM MgCl<sub>2</sub> (Advanced Biotechnology), 50 mM KCl, 10 mM Tris-HCl (pH 8.5), 1% analytical-grade formamide (Merck, Darmstadt, Germany), and 2.5 U of *Taq* polymerase (Advanced Biotechnology).

To localize and sequence the region of a  $\lambda$  clone containing *cbbLS* (see below), a ligation mixture containing the vector pGEM-7Zi(+) (Promega Corp., Madison, Wis.) and an *Xho*I fragment of the  $\lambda$  clone (see below) was used as a template. The insert was amplified by using the vector-specific primers T7<sub>rev</sub> and SP6<sub>rev</sub> (Promega Corp.). The conditions used for this PCR were as follows: 97°C hot start for 3 min and then 35 cycles of denaturation at 94°C for 1 min, annealing at 50°C for 1 min, and elongation at 72°C for 4 min. A final step consisting of 72°C for 5 min was included to fill in incomplete ends. One microliter of ligation mixture (total volume of reaction mixture, 10  $\mu$ l) was used as the template DNA. The primer concentration was 50 nM; all other conditions were the same as those described above.

For amplification of the whole *cbbLS* cluster in *Nitrosospira* sp. isolate 40KI, PCR primers encoding *Xba*I restriction sites were used (primer sequences are available on request). The conditions used for this amplification were as follows: 97°C hot start for 3 min and then 20 cycles of denaturation at 94°C for 1 min, annealing at 56°C for 1 min, and elongation at 72°C for 2 min. A final step consisting of 72°C for 5 min was included. DNA isolated from a *Nitrosospira* sp. isolate 40KI  $\lambda$  clone (see below) was used as the template at a concentration of approximately 5 ng (as determined by agarose gel electrophoresis). The primer concentration in the reaction mixture was 250 nM. The concentrations of deoxynucleoside triphosphates, MgCl<sub>2</sub>, KCl, and Tris were the same as those described above. No formamide was used. To reduce the error frequency in this PCR, a combination of *Taq* and *Pfu* (Stratagene, La Jolla, Calif.) thermostable DNA polymerases was used at a ratio of *Taq* to *Pfu* of 1:5 (D. B. Diep, personal communication). The total amount of enzyme used was 0.025 U/ $\mu$ l.

**$\lambda$  DNA library of *Nitrosospira* sp. isolate 40KI.** A  $\lambda$ -DASHII/*Bam*HI vector kit (Stratagene) and the recommended procedures were used. Briefly, total DNA from *Nitrosospira* sp. isolate 40KI was isolated (39) and partially digested with *Dpn*II to obtain a main size of 10 to 20 kb. DNA fragments were ligated to dephosphorylated  $\lambda$ -DASHII *Bam*HI arms (Stratagene) and subsequently packed into a Gigapack II XL packaging extract (Stratagene). *E. coli* XL1-Blue MRA (P2 lysogen) (Stratagene) was infected with the phage  $\lambda$ -40KI clones.

The  $\lambda$  library was screened as recommended by the manufacturer (Stratagene). Infected *E. coli* XL1-Blue MRA (P2) was plated on NZY (39) agar plates. Plaques were transferred to Colony/Plaque Screen nylon membranes (DuPont Biotechnology Systems, NEN Research Products, Boston, Mass.), and this was followed by denaturation, neutralization, and rinsing as recommended by the manufacturer. The membranes were baked for fixation at 80°C for 2 h. A ~480-bp PCR fragment from *cbbL* in *Nitrosospira* sp. isolate 40KI was labeled with <sup>32</sup>P by using a random primed DNA labeling kit (Boehringer, Mannheim, Germany), and this fragment was used as a probe. The conditions described by Church and Gilbert (13) were used for prehybridization, hybridization, and washing of the membranes. Positive plaques were isolated and preserved as recommended by Stratagene.

Positive  $\lambda$  phages were amplified as described in the  $\lambda$ -DASHII protocol (Stratagene). Briefly, plate lysates were prepared from NZY plates containing infected *E. coli* XL1-Blue MRA (P2) (Stratagene) on which confluent lysis had occurred.  $\lambda$  phages were then washed out with the recommended buffer. Amplified phage preparations were the starting point for large-scale liquid cultures in which the infection-at-low-multiplicity procedure was used (39). The number of PFU per 10<sup>10</sup> *E. coli* XL1-Blue MRA (P2) cells necessary to obtain the

maximum phage concentration was determined empirically;  $\sim 1.3 \times 10^9$  phage per  $10^{10}$  *E. coli* cells was optimal in this case. After growth and complete lysis, phage particles were isolated by the standard method for purification of bacteriophage  $\lambda$ , step 1 (39), including precipitation with polyethylene glycol 8000 and ultracentrifugation.

**DNA isolation from and mapping of positive  $\lambda$  clones.** DNA from positive  $\lambda$  clones, amplified and purified as described above, was isolated by formamide extraction (9). Briefly, the phage solution recovered from an ultracentrifuge tube was mixed with Tris-HCl (pH 8.5) and EDTA, and then analytical-grade formamide (Merck) was added. Disintegration was allowed to proceed for 30 to 60 min at room temperature, and this was followed by ethanol precipitation and washing. A high yield of pure  $\lambda$  clone DNA was obtained by this method.

Based on a restriction map of  $\lambda$ -DASHII (Stratagene), a restriction-Southern analysis was performed. Restriction enzymes and buffers were obtained from either Promega or New England Biolabs. Digested DNA was separated on 1% agarose gels and transferred to a GeneScreen Plus nylon membrane (New England Nuclear) with a Pharmacia LKB vacuum blotter, as described by Lillehaug et al. (34). The  $\sim 480$ -bp PCR product from *cbbL* of *Nitrosospira* sp. isolate 40KI (described above) was used as a probe to localize the *cbb* genes among the bands on the membrane. The probe was labeled with  $^{32}$ P with a random primed DNA labeling kit (Boehringer). The conditions used for prehybridization, hybridization, and washing of the membranes were the same as those described above (13).

**Plasmid construction.** The plasmid vector pRK310P was used as a general-purpose vector in *R. eutropha*. This vector is derived from the broad-host-range plasmid vector pRK310 (15), previously used for *R. eutropha* (7), by inserting the plasmid-stabilizing *parDE* region from RK2. The *parDE* region was obtained as a 0.7-kb *Cl*I-to-*H*indIII fragment from plasmid pRR120-0.7 (17). This plasmid was kindly provided by Don Helinski, University of California, San Diego. The *parDE* fragment was ligated into one of the two *Eco*RI sites of pRK310 following linearization of pRK310 with *Eco*RI and blunt ending of both fragments with T4 DNA polymerase. Screening of the resulting plasmids gave pRK310P, which retained the *Eco*RI site of the polylinker, while *parDE* was inserted into the other *Eco*RI site, eliminating it.

Plasmid pRub20-2 carries the *R. eutropha* ATCC 17707 chromosomally encoded *cbb* genes as a 2.1-kb fragment flanked by *H*indIII and *Eco*RI sites and inserted into the polylinker of pRK310P. This is the same *cbb* fragment carried by the previously described plasmid pAE312 (7), starting at bp 1 of GenBank database entry M17744.

The expression vector pAEX3 carries the *R. eutropha cbb* promoter-regulatory region cloned into the *H*indIII and *Eco*RI sites of the polylinker of pRK310P. It was constructed as follows. A 167-bp *Sau*3A *cbb* promoter fragment (corresponding to bp 1 to 165 of GenBank database entry M17744) was isolated from the insert of plasmid pRub20-2 and cloned into *B*amHI- and *B*glII-digested pMTL20 (12). The insert was isolated from the resulting plasmid as an *Eco*RI-to-*X*hoI fragment and was ligated to a *S*alI- and *B*amHI-digested synthetic fragment, and the whole construct was inserted into the *Eco*RI and *H*indIII sites of the pRK310P polylinker to obtain pAEX3 (see Fig. 4). The synthetic double-stranded DNA fragment was constructed from two oligonucleotides with complementary ends (the oligonucleotide sequence is available on request) by filling in with T4 DNA polymerase. The resulting construct was verified by sequencing.

**Subcloning of *Nitrosospira* sp. isolate 40KI *cbbLS* genes.** For easy handling and amplification, the  $\lambda$  DNA restriction fragments containing *cbbLS* (see above) were ligated directly into pGEM-7Zf(+) (Promega Corp.) or pUC18 (53) and transformed by electroporation (16) into *E. coli* DH5 $\alpha$ .

The *cbbLS* genes amplified with primers *cbbL-Xba*I<sub>fwd</sub> and *cbbS-Xba*I<sub>rev</sub> (see above) were digested with *Xba*I and ligated directly into the *Xba*I site in the expression vector pAEX3 (described above). T4 ligase (Promega Corp.) was used for ligation.

**Conjugal plasmid transfer.** The broad-host-range plasmid pRK310P and derivatives of this plasmid were transferred from *E. coli* to *R. eutropha* AE951 by triparental filter mating by using *E. coli* DH5 $\alpha$ (pRK2013) for plasmid mobilization as previously described (5).

**DNA sequencing and sequence analyses.** All sequences obtained in this study originated from PCR products. An ABI Prism 377 DNA sequencer (Perkin-Elmer, Applied Biosystems, Foster City, Calif.) was used for separation and registration of fluorescent DNA fragments obtained with a Thermo Sequenase II dye terminator cycle sequencing premix kit (Amersham Life Science/Pharmacia Biotech, Cleveland, Ohio). The cycle sequencing reactions were performed in 10- $\mu$ l (total volume) mixtures instead of 20- $\mu$ l mixtures, and one-half the recommended amount of dye terminator was used. The concentrations of primer and template were double the recommended concentrations, 0.5 pmol/ $\mu$ l and 5 ng/ $\mu$ l, respectively. DNA sequences were analyzed with the Genetics Computer

Group program package, version 8.1 (Program manual for the Wisconsin package, version 8, September 1994; Genetics Computer Group, Madison, Wis.), as well as the Auto Assembler sequence assembly software (release 1.4.0; MacApp version 3.1.2, Abisid version 1.0.5; ABI Prism, Applied Biosystems, Foster City, Calif.). Phylogenetic analyses were performed by using DNAML, PROML, SEQBOOT, and CONSENSE in the Phylip program package, version 3.572c (19), in combination with the TreeView software, version 1.4 (from R. D. M. Page [http://taxonomy.zoology.gla.ac.uk/rod/rod.html]).

**RubisCO assay and gel analysis.** RubisCO activity was determined by measuring ribulose 1,5-bisphosphate-dependent incorporation of  $^{14}$ C $_2$  (from NaH $^{14}$ CO $_3$ ) into acid-stable products by whole permeabilized cells as previously described (7). The reaction time was 60 s at 25°C in the presence of 25 mM NaHCO $_3$  (pH 7.8), and the O $_2$  concentration was the concentration in air. Denaturing polyacrylamide gel electrophoresis (PAGE) in the presence of sodium dodecyl sulfate (SDS) was performed as previously described (7). Gels were stained with BM Fast Stain (Boehringer) by the procedure recommended by the manufacturer. Protein concentrations were determined as previously described (7).

**Nucleotide sequence accession numbers.** The nucleotide sequence of *cbbLS* and the translated CbbL and CbbS sequences from *Nitrosospira* sp. isolate 40KI have been deposited in the GenBank database under accession no. AF426428. The *cbbL* partial nucleotide sequences obtained in this work have also been deposited in the GenBank database, together with the translated CbbL partial amino acid sequences. The accession numbers are as follows: *Nitrosospira* sp. isolate AF, AF426415; *Nitrosospira* sp. isolate III2, AF426416; *Nitrosospira* sp. isolate L115, AF426417; *Nitrosospira multiformis* N113, AF426418; *Nitrosospira* sp. isolate A4, AF426419; *Nitrosospira* sp. isolate F3, AF426420; *Nitrosospira* sp. isolate A3, AF426421; *Nitrosospira* sp. isolate O13, AF426422; *Nitrosospira* sp. isolate III7, AF426423; *Nitrosospira* sp. isolate Ka3, AF426424; *Nitrosococcus mobilis* Nc3, AF426425; *Nitrosococcus halophilus* Nc4, AF426426; and *Nitrosomonas europaea* Nm50, AF426427.

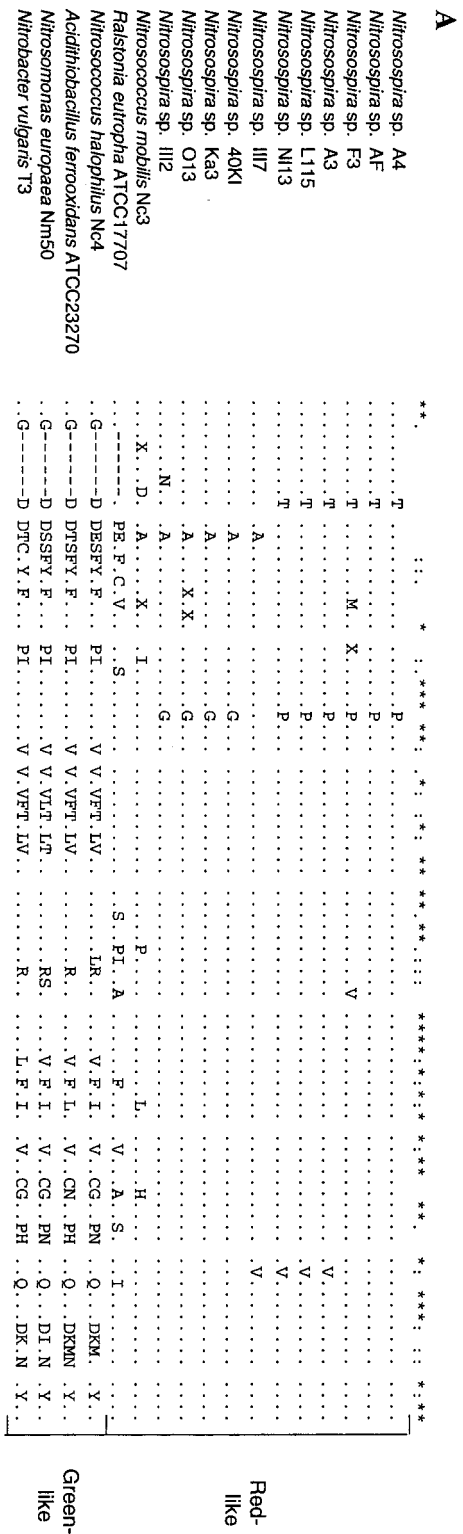
## RESULTS

### Sequence analyses of a conserved region of CbbL from AOB.

To detect *cbbLS* genes in *Nitrosospira* sp. isolate 40KI, PCR amplification reactions were performed with a selection of primers that recognize conserved regions of *cbbL* in *R. eutropha* ATCC 17707. Since *R. eutropha* is quite closely related to the ammonia oxidizers based on 16S rRNA sequence comparisons and RubisCO is a conserved protein, it appeared likely that some of the primers would recognize *cbbL* in *Nitrosospira* sp. isolate 40KI. It turned out that the majority of the primers tested produced fragments of the expected size when total DNA from *Nitrosospira* sp. isolate 40KI was used as the template. A PCR product covering a central region of *cbbL* was sequenced, and databank searches resulted in identification of *R. eutropha cbbL* as the most similar sequence. The level of DNA sequence similarity between *R. eutropha* and *Nitrosospira* sp. isolate 40KI in this part of *cbbL* was 72.2% (Fig. 1B).

The same region of *cbbL* was then amplified from 13 other AOB (belonging to both the  $\beta$  and  $\gamma$  subgroups of the *Proteobacteria*), including 1 *Nitrosomonas* sp. isolate, 10 *Nitrosospira* sp. isolates (1 of which has the basonym *Nitrosolobus*), and 2 *Nitrosococcus* sp. isolates (1  $\beta$ -subgroup member and 1  $\gamma$ -subgroup member).

An alignment of the resulting CbbL deduced amino acid sequences is shown in Fig. 1A. Except for the *Nitrosomonas europaea* type strain sequence, all of the sequences originating from the  $\beta$ -subgroup AOB showed >71% amino acid sequence identity to the *R. eutropha* CbbL sequence and >86% sequence identity to each other. When the gap in the alignment in Fig. 1A was excluded, the levels of amino acid sequence identity to the *R. eutropha* CbbL sequence were >77%. This indicates that these bacteria all have a red-like type I RubisCO, like *R. eutropha*. Compared to the *R. eutropha* CbbL



**B**

	Nsp A4	Nsp AF	Nsp F3	Nsp A3	Nsp L115	Nsp NI13	Nsp NI17	Nsp 40KI	Nsp Ka3	Nsp O13	Nsp III2	Nc3	<i>eutropha</i>	<i>Nc4</i>	<i>A. ferrooxidans</i>	Ntn50	<i>N. vulgarens</i>
Nsp A4	100	100	96.4	98.8	98.8	98.8	95.2	96.4	96.4	94.0	95.2	88.1	72.6	53.6	53.6	53.6	56.0
Nsp AF	91.3	100	96.4	98.8	98.8	95.2	96.4	96.4	94.0	95.2	88.1	72.6	53.6	53.6	53.6	53.6	56.0
Nsp F3	94.4	88.5	100	95.2	95.2	91.7	92.9	92.9	90.5	91.7	85.7	71.4	51.2	51.2	51.2	51.2	56.0
Nsp A3	87.7	84.1	91.7	100	100	96.4	95.2	95.2	92.9	94.0	86.9	71.4	52.4	52.4	52.4	52.4	54.8
Nsp L115	95.2	90.9	93.7	89.7	100	96.4	95.2	95.2	92.9	94.0	86.9	71.4	52.4	52.4	52.4	52.4	54.8
Nsp NI13	94.8	91.7	94.4	89.7	85.7	98.8	96.4	95.2	92.9	96.4	90.5	73.8	53.6	53.6	53.6	53.6	56.0
Nsp NI17	86.1	81.7	85.7	84.5	84.1	85.7	100	97.6	97.6	97.6	98.8	90.5	73.8	53.6	53.6	53.6	56.0
Nsp 40KI	83.3	83.3	84.9	85.7	85.7	84.1	100	100	100	97.6	98.8	90.5	73.8	53.6	53.6	53.6	56.0
Nsp Ka3	84.5	82.5	85.3	84.9	82.1	83.7	87.3	90.5	90.5	97.6	98.8	90.5	73.8	53.6	53.6	53.6	56.0
Nsp O13	85.3	81.0	83.7	82.1	82.1	83.3	92.9	84.5	84.5	100	96.4	88.1	71.4	52.4	52.4	52.4	54.8
Nsp III2	83.7	82.5	84.5	84.9	82.9	84.1	84.1	96.4	96.4	76.2	77.0	89.3	73.8	53.6	53.6	53.6	56.0
Nc3	77.8	75.8	77.8	75.4	77.0	77.4	77.0	79.0	78.2	76.2	77.0	100	73.8	53.6	53.6	53.6	56.0
<i>R. eutropha</i>	72.6	73.9	74.4	73.1	73.5	73.5	74.4	72.2	73.9	72.6	71.4	66.7	59.4	56.4	54.8	53.8	56.0
<i>Nc4</i>	57.7	60.3	60.3	60.3	59.4	59.4	59.4	62.0	61.5	63.2	61.5	59.8	59.4	56.4	54.8	53.8	56.0
<i>A. ferrooxidans</i>	60.3	59.0	60.3	60.3	59.8	60.7	63.2	61.1	61.1	62.0	62.8	64.5	61.1	56.0	54.8	53.8	56.0
Ntn50	58.5	60.3	59.4	58.1	58.5	60.7	62.0	59.8	60.7	60.7	59.8	62.8	56.0	54.8	53.8	53.8	56.0
<i>N. vulgarens</i>	61.1	63.2	62.4	63.2	61.5	63.7	63.7	64.1	62.4	62.4	62.8	60.7	66.2	76.9	79.5	73.9	100

FIG. 1. Comparison of partial RubisCO large-subunit sequences from 14 strains of AOB (for accession numbers see Materials and Methods). (A) Deduced amino acid sequence alignment compared to the corresponding regions in the CbbL proteins of *R. eutropha* ATCC 17707 (accession number M17744) (6), *Acidithiobacillus ferrooxidans* (formerly *Thiobacillus ferrooxidans*) (30) ATCC 23270 (AF129925) (J. M. Shively, unpublished data, 1999), and *Nitrosospira vulgarens* T3 (L22885) (M. Strecker, E. Sickinger, R. S. English, J. M. Shively, and E. Bock, unpublished data). The translated sequences were aligned by using ClustalW (44). Identical residues are indicated by asterisks, conservative substitutions are indicated by colons, and semiconservative substitutions are indicated by periods. The active site residue Asn126 (*R. eutropha* numbering) is indicated. (B) Percentages of identity for the CbbL amino acid sequences in panel A (upper right) and for the *cbbL* nucleotide sequences (lower left). Levels of identity were calculated by using the HOMOLOGIES program (Genetics Computer Group). Abbreviations: Nsp, *Nitrosospira*; Nc, *Nitrosococcus*; Nm, *Nitrosomonas*.

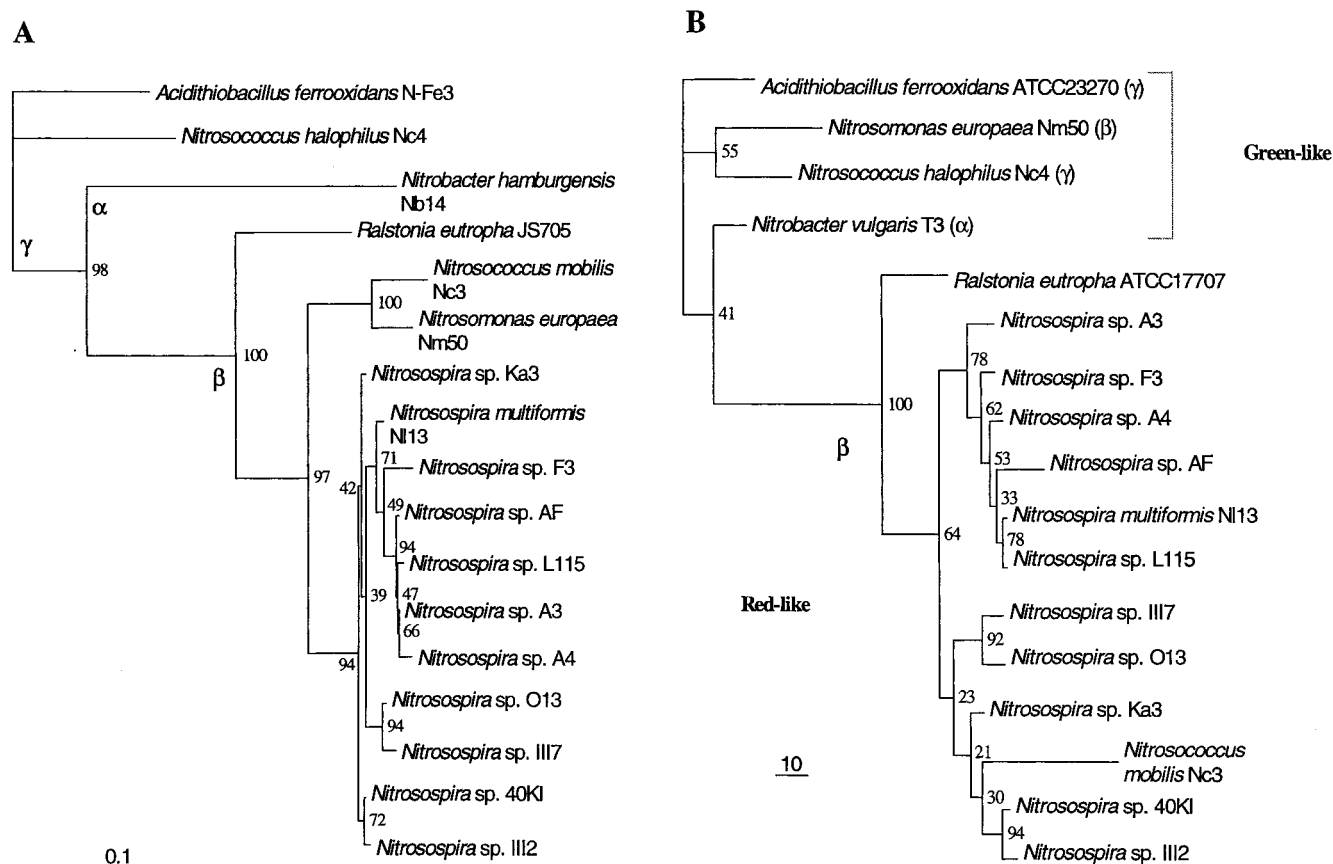


FIG. 2. Comparison of 16S rDNA (A) and RubisCO (CbbL protein) (B) phylogenetic trees derived from sequences for the ammonia oxidizers in this study. The positions of *R. eutropha*, *Acidithiobacillus ferrooxidans*, and *Nitrobacter* spp. are included for comparison. The phylogenetic trees are maximum-likelihood trees with bootstrap values resulting from an analysis of 100 data sets (19). The trees are based on ClustalW (44) multiple-sequence alignments. All CbbL amino acid sequences were cut to the same length (83 amino acids), whereas the 16S rDNA sequences in the analysis were between 1,300 and 1,500 bp long. Trees were visualized by using the TreeView program package, version 1.4 (from R. D. M. Page [http://taxonomy.zoology.gla.ac.uk/rod/rod.html]). The 16S rDNA sequences of the following organisms were used: *Nitrobacter hamburgensis* Nb14 (accession number L35502) (43), *Nitrosospira multiformis* N113 (L35509) (43), *R. eutropha* JS705 (AF027407) (48), *Nitrosomonas europaea* Nm50 (M96399) (26), *Nitrosospira* sp. isolate Ka3 (AJ012105) (2), *Nitrosospira* sp. isolate O13 (AJ010108) (2), *Nitrosospira* sp. isolate 40KI (X84656) (46), *Nitrosospira* sp. isolate AF (X84658) (46), *Nitrosospira* sp. isolate L115 (X84662) (46), *Nitrosospira* sp. isolate III2 (AJ000344) (47), *Nitrosospira* sp. isolate III7 (AJ000345) (47), *Nitrosospira* sp. isolate F3 (AJ005545) (1), *Nitrosospira* sp. isolate A4 (AJ005543) (1), *Nitrosospira* sp. isolate A16 (AJ005544) (1), *Acidithiobacillus ferrooxidans* N-Fe3 (X75268) (20), *Nitrosococcus mobilis* Nc3 (Å. Aakra, unpublished data), and *Nitrosococcus halophilus* Nc4 (AJ298748) (Aakra, unpublished data). The RubisCO CbbL sequences used were the sequences of *Nitrobacter vulgaris* T3, *R. eutropha* ATCC 17707, and *A. ferrooxidans* ATCC 23270 (see Fig. 1 for details). The remaining CbbL sequences were translated nucleotide sequences from this study.

sequence, the CbbL sequences of the ammonia oxidizers belonging to the *Nitrosospira* cluster have a six-amino-acid insertion at approximately Asn95 (*R. eutropha* sequence numbering) (Fig. 1A). This insert is not present in any other type I RubisCO sequenced so far. It is an interesting distinguishing feature of the AOB RubisCOs, as discussed below. According to the three-dimensional structure of the  $L_8S_8$  RubisCO from *R. eutropha* (23), this six-amino-acid insertion is located in a surface loop region, as discussed below (see Fig. 6).

Two representatives of the *Nitrosomonas* cluster, *Nitrosococcus mobilis* Nc3 and *Nitrosomonas europaea* Nm50, were studied. Strain Nc3 has a CbbL sequence that is the same type as the sequences of the nitrosospirads; i.e., six amino acids are inserted in the same position (Fig. 1A). The sequence of *Nitrosomonas europaea* Nm50 is surprisingly different, however, and exhibits a high level of similarity to the CbbL sequence of

the  $\gamma$ -subgroup member *Nitrosococcus halophilus* Nc4. Sequence comparisons with the database revealed that *Nitrosomonas europaea* Nm50 and *Nitrosococcus halophilus* Nc4 both have a green-like type I RubisCO. Their CbbL proteins exhibit <60% amino acid sequence identity to the *R. eutropha* CbbL (Fig. 1B) and >88% sequence identity to previously sequenced green-like CbbL proteins.

The CbbL fragment analyzed includes one known RubisCO active site residue (8), corresponding to Asn126 in the *R. eutropha* sequence (Fig. 1A). This residue is, as expected, conserved in all of the sequences.

The alignment in Fig. 1A was the basis for a phylogenetic study of the CbbL proteins, and in Fig. 2 the results are compared with a 16S rDNA-based phylogeny of the same organisms. The main branches of the CbbL-based phylogenetic tree are the same as those of the 16S rDNA-based phylogenetic

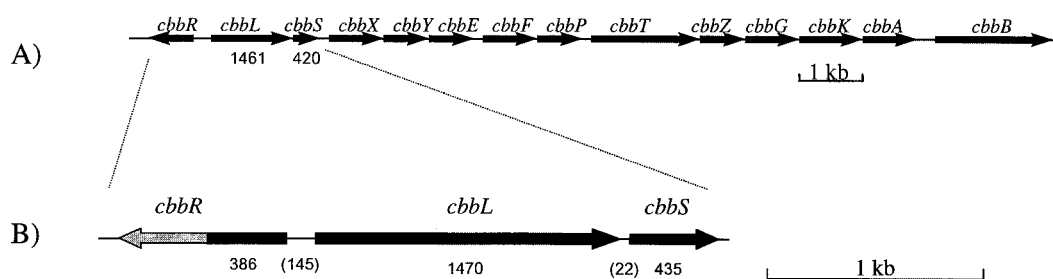


FIG. 3. Organization of *cbb* genes of the chromosomal *cbb* operon of *R. eutropha* H16 (A) (33) and the *cbbR* (partial sequence)-*cbbL*-*cbbS* region of *Nitrosospira* sp. isolate 40KI analyzed in this study (B). The lengths (in nucleotides) of the genes identified are indicated, and the lengths of the intergenic regions are indicated in parentheses.

tree, with the important exception of *Nitrosomonas europaea* Nm50, which clusters with the  $\gamma$ -subgroup AOB. The significance of this is discussed below. Within the  $\beta$ -subgroup AOB there is no distinction between the *Nitrosomonas* and *Nitrosospira* groups in the CbbL-based tree, and the various groups of nitrosospirads cluster somewhat differently. The stability is lowest in the CbbL tree, in which five branch points have bootstrap values of <50 and the mean bootstrap value is 58.9. In the 16S rDNA-based tree, there are four branch points with bootstrap values of <50, and the mean bootstrap value is 75.9.

**Cloning and sequencing of the *Nitrosospira* sp. isolate 40KI *cbbLS* region.** A 480-bp PCR fragment of *cbbL* from *Nitrosospira* sp. isolate 40KI was used as a probe to screen a *Nitrosospira* sp. isolate 40KI  $\lambda$ -DASHII library. Three positive  $\lambda$  clones were identified, and these clones were mapped to localize the region containing *cbbL*. Only one of the three positive  $\lambda$  clones contained the whole *cbbLS* cluster, which was located on a  $\sim$ 4.7-kb *Xba*I fragment. This fragment was ligated into the pGEM and pUC18 cloning vectors. The DNA sequence of *cbbLS* and the surrounding region in *Nitrosospira* sp. isolate 40KI was obtained by using vector-specific primers, as well as primers recognizing *R. eutropha* *cbbLS*. A schematic representation of the results of the sequencing analysis is shown in Fig. 3B. The organization of the *cbbL* and *cbbS* genes, as well as of a putative *cbbR* gene, of *Nitrosospira* sp. isolate 40KI is very similar to the organization that has been described for *R. eutropha* (Fig. 3A) (6, 33). In the promoter region of the *Nitrosospira* sp. isolate 40KI *cbb* genes there are two possible  $-10$  and  $-35$  regions. The *cbbL* gene is 1,470 bp long and encodes a CbbL protein that is 489 amino acids long, whereas the *cbbS* gene is 435 bp long and encodes a CbbS protein that is 144 amino acids long. Both genes are preceded by plausible ribosome-binding sites (RBS). All large-subunit active site residues identified in the high-resolution crystal structure of spinach RubisCO (8) were present in the 40KI sequence. The spacer region between *cbbL* and *cbbS* is 22 bp long. The overall levels of amino acid identity with *R. eutropha* sequences are 84.5% for CbbL and 75.2% for CbbS.

CbbR, the LysR-type regulator of the *cbb* genes, is known to be quite conserved (33, 41). The DNA sequence of the first 386 bp of *cbbR* from *Nitrosospira* sp. isolate 40KI (determined for only one DNA strand) was compared with the EMBL DNA sequence database. The *cbbR* gene of *Thiobacillus intermedius* strain K12 (J. M. Shively and F. Soyer, unpublished data, 1997) was most similar, with a level of DNA sequence identity of

58%. The level of DNA sequence identity with the *cbbR* gene of *R. eutropha* was only 35% in this region. The DNA sequence similarity between *Nitrosospira* sp. isolate 40KI and *T. intermedius* is restricted to the *cbbR* gene, since *T. intermedius* has a green-like type I RubisCO.

**Plasmid vector for regulated expression of foreign RubisCO genes in *R. eutropha*.** Cultivation of AOB is laborious because of slow growth and low cell yields. Therefore, expression of cloned genes from AOB in hosts better suited for generation of large quantities of cell material is desirable for studies of protein function. *R. eutropha* is a particularly convenient host for *cbbLS* gene expression experiments.

Plasmid pRub20-2 (22) carries a 2.1-kb fragment of the chromosomally encoded *cbb* operon of *R. eutropha* ATCC 17707, including the *cbbLS* structural genes and the regulatory region. The *cbbLS* upstream sequence in this plasmid is sufficient to direct regulated expression of *R. eutropha* RubisCO in the *cbb* deletion strain AE951 at levels similar to wild-type *R. eutropha* levels (Table 1). This *cbb* upstream sequence was therefore used in the construction of the gene expression cassette for *R. eutropha* shown in Fig. 4A. This expression cassette was designed with high-level production of heterologous proteins in mind, using principles used successfully for *E. coli* (36). It features a bicistronic construct encoding a short N-terminal fragment of the highly expressed CbbL protein, a stop codon, a second (presumably) strong RBS, and a second ATG, as well as downstream restriction sites for creating fusions. Care was taken to avoid potentially stable mRNA secondary structures

TABLE 1. Growth rates and RubisCO activities of *R. eutropha* wild type and *cbb* deletion strain AE951 complemented with RubisCO genes from *R. eutropha* and *Nitrosospira* sp. isolate 40KI

Growth conditions	Organism	Generation time (h)	RubisCO activity (% of <i>R. eutropha</i> wild-type activity) <sup>a</sup>
Heterotrophic	<i>R. eutropha</i> wild type	2.6	<1
	AE951(pRub20-2)	2.6	<1
	AE951(pRK310P)	2.6	<1
	AE951(pAEX3-40KI)	2.6	<1
Autotrophic	<i>R. eutropha</i> wild type	2.0	105
	AE951(pRub20-2)	2.0	100
	AE951(pAEX3-40KI)	4.6	34
	AE951(pRK310P)	NG <sup>b</sup>	<1

<sup>a</sup> Activity based on total cellular protein. The RubisCO activity of *R. eutropha* growing autotrophically was defined as 100%.

<sup>b</sup> NG, no growth.

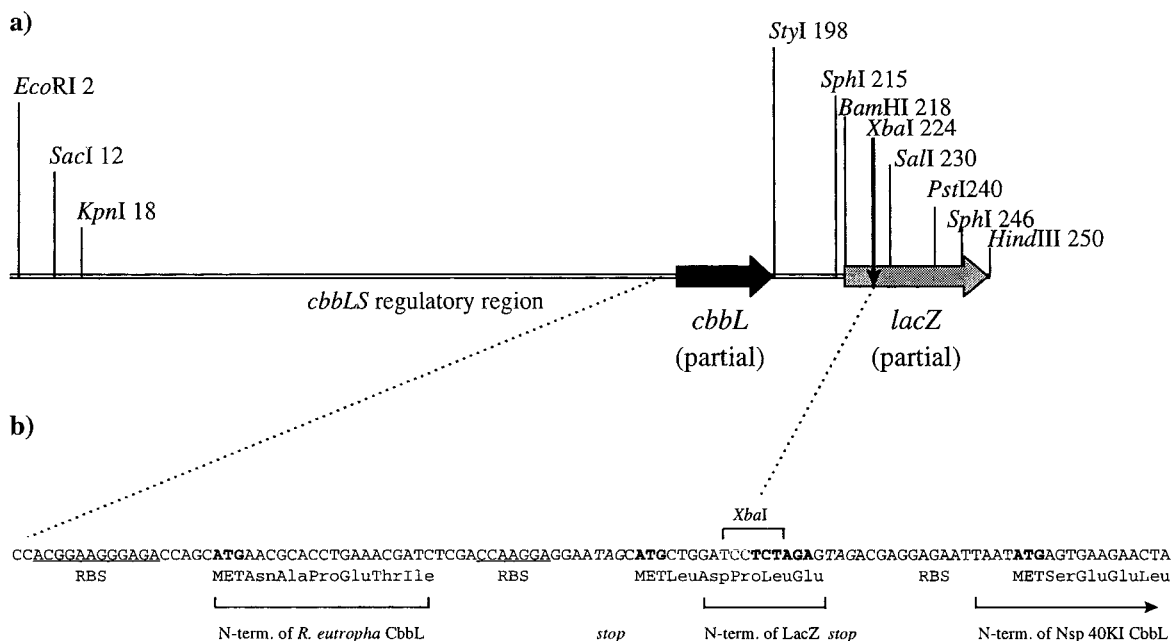


FIG. 4. Gene constructs for expression of *Nitrosospira* sp. isolate 40KI RubisCO in *R. eutropha*. (a) Gene cassette for regulated expression of foreign genes in *R. eutropha*. The cassette (the *EcoRI-HindIII* fragment of plasmid pAEX3) was constructed as described in Materials and Methods. Nucleotides 1 to 27 originate from the pUC18 polylinker, nucleotides 28 to 154 originate from the *R. eutropha cbb* promoter-regulatory region, nucleotides 155 to 192 are the RBS and N-terminal end of the *R. eutropha* wild-type *cbbL* sequence, nucleotides 193 to 217 are a synthetic DNA fragment containing an RBS, as well as a stop codon and a start codon, and nucleotides 218 to 250 are from the *lacZ* polylinker of pUC18. Genes inserted into the *lacZ* polylinker can be expressed under the control of the *R. eutropha cbbLS* promoter. (b) Gene construct for expression of *Nitrosospira* sp. isolate 40KI RubisCO in *R. eutropha*. The *cbbLS* genes from *Nitrosospira* sp. isolate 40KI, amplified from the  $\lambda$  clone, were inserted as an *XbaI* fragment into the *XbaI* site of pAEX3 shown in panel A. The translation start was directed to *Nitrosospira* sp. isolate 40KI *cbbLS* due to the stop codon (TAG) which was part of the forward amplification primer. N-term., N terminus.

in the translation initiation region. The expression cassette was inserted into the broad-host-range plasmid vector pRK310P to form the expression vector pAEX3. This vector allowed regulated expression of foreign genes in *R. eutropha*, as described below for *cbbLS* genes from AOB.

The *cbbLS* coding sequences of *Nitrosospira* sp. isolate 40KI are flanked by two *AseI* restriction sites. Efforts to perform blunt-end ligation of the *AseI* fragment into pAEX3 were not successful. Instead, the whole *cbbLS* region of *Nitrosospira* sp. isolate 40KI (~2 kb) had to be amplified from the  $\lambda$  clone with primers containing *XbaI* restriction sites (i.e., restriction sites that are not present in the cluster itself). To increase the fidelity of the PCR, a combination of the *Pfu* and *Taq* thermostable polymerases was used (see Materials and Methods). The PCR product obtained was ligated directly into pAEX3 after digestion with *XbaI* and was transferred into *E. coli* DH5 $\alpha$ . A detailed illustration of this construct is shown in Fig. 4B. To check for polymerase-generated errors, both DNA strands of the ligated PCR product were sequenced. No sequence ambiguities were observed.

**Complementation of a RubisCO deletion strain of *R. eutropha* by RubisCO genes from *Nitrosospira* sp. isolate 40KI.** *R. eutropha* strain AE951 (see Materials and Methods) is a deletion strain of *R. eutropha* wild-type strain ATCC 17697 (which contains only a chromosomal copy of *cbbLS* and no plasmid-encoded RubisCO). In the deletion strain, the chromosomal *cbbLS* genes have been deleted so that the organism expresses

no RubisCO at all and is therefore unable to grow autotrophically.

The *Nitrosospira* pAEX3-40KI *cbbLS* plasmid construct (Fig. 4B) was transferred from *E. coli* DH5 $\alpha$  to *R. eutropha cbbLS* deletion strain AE951 by conjugation. The plasmid was transferred from *E. coli* to *R. eutropha* at a high frequency (about  $1 \times 10^{-2}$ ), like the parent vector pRK310 (7). Transconjugants of *R. eutropha* AE951 were readily selected and purified on fructose plates with tetracycline. As a control, cloned *R. eutropha* ATCC 17707 *cbbLS* wild-type genes on plasmid pRub20-2 were transferred from *E. coli* to AE951. All transconjugants were capable of autotrophic growth in an atmosphere containing 5% CO<sub>2</sub>, 20% O<sub>2</sub>, and 75% H<sub>2</sub>. Complementation of AE951 with pRub20-2 (positive control) worked very well. The autotrophic growth rate both on plates and in liquid medium was indistinguishable from the growth rate of *R. eutropha* wild-type strain ATCC 17697. Autotrophic growth of AE951(pAEX3-40KI) both on plates and in liquid medium was, however, slower than growth of the *R. eutropha* wild-type control. In liquid cultures, the generation time of *R. eutropha* AE951 harboring *Nitrosospira cbbLS* was 4.6 h, which was slightly more than twice the generation time of the *R. eutropha* controls (Table 1).

To determine the level of RubisCO in the cells, RubisCO activity in permeabilized whole cells was first assayed. When autotrophically grown cells were used, AE951(pAEX3-40KI) exhibited about one-third the specific activity (based on the

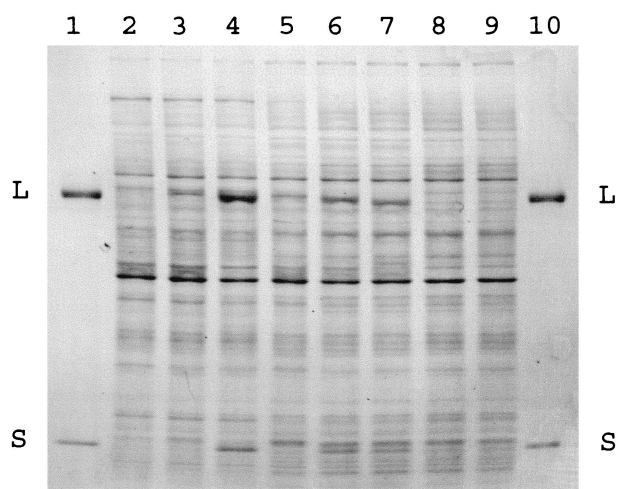


FIG. 5. Expression of cloned *Nitrosospira* sp. isolate 40KI RubisCO in *R. eutropha* analyzed by SDS-PAGE. Protein samples were electrophoresed in an 11 to 18% polyacrylamide gradient denaturing gel. Lanes 1 and 10 contained purified *R. eutropha* ATCC 17707 chromosomally encoded RubisCO. The positions of the large (L) and small (S) subunits are indicated. Lanes 2 to 9 contained total cellular protein, as follows: lanes 2 to 4, induced cells [lane 2, AE951(pRK310P) (negative control); lane 3, AE951(pAEX3-40KI); lane 4, AE951(pRub20-2)]; lanes 5 to 7, autotrophic cells [lane 5, AE951(pAEX3-40KI); lane 6, AE951(pRub20-2); lane 7, *R. eutropha* wild-type strain ATCC 17697]; lanes 8 and 9, heterotrophic cells [lane 8, AE951(pAEX3-40KI); lane 9, AE951(pRub20-2)]. The cell preparations were the same as those used to obtain the data shown in Table 1.

total cellular protein) of the *R. eutropha* controls. The lower activity is consistent with the lower growth rates observed with the *Nitrosospira* RubisCO (Table 1). No activity was detected in heterotrophically grown cells (Table 1). Thus, regulated expression of *Nitrosospira* RubisCO in *R. eutropha* was achieved. The pattern of growth medium-dependent expression was similar to that of *R. eutropha* wild-type RubisCO.

To assess the level of RubisCO protein in the cells, whole-cell protein was analyzed by SDS-PAGE. The band patterns for total proteins from the strains in Table 1 are shown in Fig. 5. Purified RubisCO from *R. eutropha* ATCC 17707, with large and small subunits consisting of 486 amino acids (molecular mass, 53.8 kDa) and 139 amino acids (16.1 kDa), respectively, was used as a reference. Bands corresponding to RubisCO large and small subunits were present in autotrophically grown cells of *R. eutropha*. The levels of these two polypeptides were similar in the two *R. eutropha* autotrophic controls, and the polypeptides were not present in uninduced fructose-grown cells (Fig. 5). These results are consistent with the RubisCO whole-cell activity results (Table 1). If no processing took place, the *Nitrosospira* sp. isolate 40KI small subunit (CbbS) was 144 amino acids long (molecular mass, 16.5 kDa), whereas the large subunit (CbbL) was 489 amino acids long (54.2 kDa); i.e., both subunits were slightly larger than the corresponding *R. eutropha* wild-type subunits. SDS-PAGE analysis revealed the presence of a polypeptide that was about the same size (but apparently slightly larger) than the *R. eutropha* RubisCO large subunit in autotrophically grown cells of AE951 complemented with pAEX3-40KI. This band presumably represented the *Nitrosospira* sp. isolate 40KI RubisCO large subunit. In agree-

ment with this, the band was not present in fructose-grown cells that exhibited no RubisCO activity (Fig. 5). Very high levels of *R. eutropha* RubisCO activity can be induced by using special induction conditions (4). This occurs when cells are incubated in L-Ile minimal medium under an atmosphere containing H<sub>2</sub> plus O<sub>2</sub> but no CO<sub>2</sub>. With *R. eutropha* RubisCO, these inducing conditions resulted in large increases in the levels of the large and small subunits compared with the levels seen in autotrophic cells. There was a corresponding increase in the whole-cell RubisCO level of activity (data not shown). With AE951 expressing *Nitrosospira* sp. isolate 40KI RubisCO, induction in L-Ile medium resulted in no significant increase in RubisCO activity or protein levels compared with the levels in autotrophic cells (Fig. 5). The *Nitrosospira* sp. isolate 40KI RubisCO small subunit was just barely distinguishable in the SDS-PAGE gel experiments. The gel analysis, however, clearly showed that the level of expression of the *Nitrosospira* sp. isolate 40KI RubisCO large subunit was much lower than the level of expression of the *R. eutropha* large subunit in parallel experiments. This is in general agreement with the differences observed in autotrophic growth rates and whole-cell RubisCO activity levels (Table 1 and Fig. 5).

## DISCUSSION

AOB are dependent on CO<sub>2</sub> fixation to obtain cell carbon. This is the first report analyzing and comparing carbon fixation genes from this group of chemoautotrophs. Type I and type II RubisCOs are conserved molecules with clear sequence differences, and a sequence analysis of the *cbbL* fragments obtained showed that the RubisCOs from *Nitrosospira* sp. isolate 40KI and all of the other AOB analyzed in this study are type I RubisCOs. SDS-PAGE of total proteins from *Nitrosospira* sp. isolate 40KI (Fig. 5) confirmed that the enzyme consists of small and large subunits that have molecular masses very similar to the molecular masses of the corresponding subunits in *R. eutropha*. A number of bacteria exhibit *cbb* gene duplications, as first demonstrated for *R. eutropha* (5). Although our PCR and sequencing results gave no indication of heterozygotes, the possibility that additional *cbb* genes are present in the AOB has not been eliminated.

**Red-like RubisCOs with an unusual sequence insert.** The levels of homology to *R. eutropha* RubisCO on both the DNA level and the protein level are striking for all of the ammonia oxidizer RubisCOs studied except the RubisCOs of *Nitrosomonas europaea* and *Nitrosococcus halophilus*. *R. eutropha* has a red-like RubisCO similar to the enzymes in rhodophytic and chromophytic algae, as well as in several members of the  $\alpha$  and  $\beta$  subgroups of the *Proteobacteria* (49). We therefore concluded that all but two of the AOB studied here possess a red-like type I RubisCO. The two exceptions have a green-like enzyme, as discussed below.

Until now, apart from quite variable N- and C-terminal ends, CbbL amino acid sequence alignments for type I RubisCOs have revealed very few gaps, none of which was more than two residues long (27; unpublished data). Therefore, the six-amino-acid residue gap identified in this study (Fig. 1A) was unexpected. So far, the insert is restricted to the organisms which we studied; it was not found in any other type I CbbL sequences in databank searches. It was also not found in the



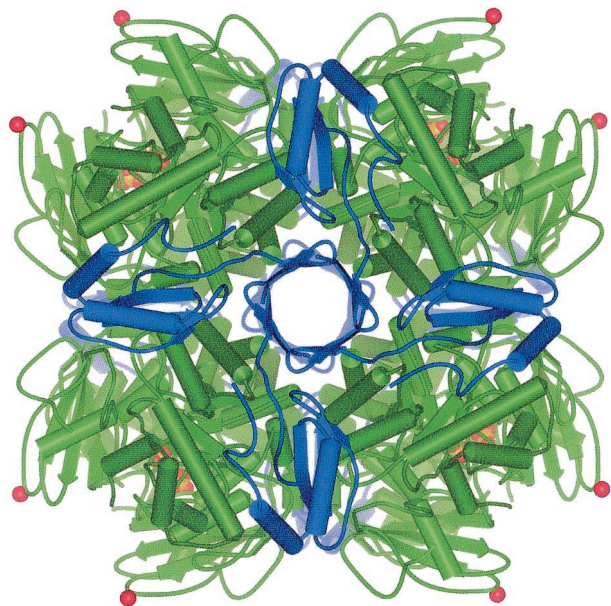


FIG. 6. Tertiary structure of RubisCO, showing the position of the six-amino-acid insert in the *Nitrosospira* sp. large subunit. The  $L_8S_8$  *R. eutropha* holoenzyme is shown, based on the structure of the unactivated enzyme, as determined by X-ray crystallography (23). Large subunits are green, and small subunits are blue. Phosphate ions bound in the active site are yellow. The C-alpha atom of Asn95 (corresponding to Asn91 in the *Nitrosospira* sp. isolate 40KI sequence) is represented as a red sphere to indicate the approximate position of the insert in the *Nitrosospira* large subunit. Asn95 is situated in a loop region between  $\beta$ -strands C and D. This loop extends from Val93 to Glu98 in the *R. eutropha* RubisCO (23). The image was produced with WebLab ViewerPro (Molecular Simulations Inc., San Diego, Calif.).

green-like CbbL proteins from *Nitrosomonas europaea* and *Nitrosococcus halophilus* characterized here. We observed that the six-residue insert is highly conserved in the different strains (Fig. 1).

The high levels of sequence identity between *Nitrosospira* sp. isolate 40KI and *R. eutropha* large and small subunits imply that these RubisCO proteins have very similar tertiary structures. The recently determined crystal structure of the *R. eutropha* enzyme (23) may therefore serve as a quite accurate structural model of the *Nitrosospira* sp. isolate 40KI RubisCO. The main structural feature distinguishing the red-like *R. eutropha* RubisCO from green-like RubisCOs is an eight-stranded  $\beta$ -barrel formed by the extended C terminus of the small subunits. This  $\beta$ -barrel is a central plug in the entrance to the central solvent channel of the holoenzyme. The *Nitrosospira* sp. isolate 40KI small subunit also has an extended C terminus that is the same length and exhibits a high level of sequence identity to the *R. eutropha* small subunit (alignment not shown). The *Nitrosospira* RubisCO is therefore likely to have a similar  $\beta$ -barrel structure.

If the highly homologous *R. eutropha* RubisCO is used as a structural model, it can be concluded that the six-amino-acid insert in the ammonia oxidizers is located on the surface of the enzyme molecule (Fig. 6). The insert site is in a loop region between  $\beta$ -strands C and D of the N-terminal domain of CbbL (nomenclature of Hansen et al. [23]), at approximately amino acid position 95 (*R. eutropha* numbering). It may therefore not

disrupt the secondary structural elements, which are conserved in all type I RubisCOs that have been examined so far by X-ray crystallography (23). The insert is too far from the active site to exert any direct influence on the catalytic properties. We note that the distantly related type II RubisCOs have a five-amino-acid deletion in the same region. It is possible, therefore, that this surface loop is more susceptible to variation than other parts of the conserved enzyme.

In the future, structural and functional consequences of the six-amino-acid insert can be investigated further by performing comparative studies in *R. eutropha* and in vitro of wild-type and mutant AOB RubisCOs in which the six residues have been deleted. As mentioned previously, many AOB, possibly including *Nitrosospira* sp. isolate 40KI, are believed to sequester their RubisCOs in carboxysomes (11, 29, 32, 52). Packaging of RubisCO in carboxysomes (or crystallization in vitro) may be influenced by such a surface-located insert.

**Functional expression.** In this work we demonstrated that the facultative chemoautotroph *R. eutropha* is a convenient host for studies of *cbb* gene function in the closely related AOB. Compared to the AOB, *R. eutropha* grows very rapidly (its doubling time is 2 h under autotrophic conditions) and to high cell densities.

Expression of wild-type *R. eutropha* RubisCO and other CBB pathway enzymes is subject to strong genetic control at the transcriptional level, and little or no expression occurs under most heterotrophic growth conditions (4, 33). The CbbR protein activates transcription of the *cbb* operon, and the promoter-operator sites have been characterized in detail (28). By using the *R. eutropha cbbLS* deletion strain AE951 and the expression vector pAEX3, which carries the *cbb* promoter-operator region, functional expression of *Nitrosospira* sp. isolate 40KI *cbbLS* genes was obtained in this host (Table 1). The expression was regulated. The *Nitrosospira* RubisCO supported autotrophic growth of *R. eutropha* AE951(pAEX3-40KI) with a doubling time of 4.6 h, which is slightly more than twice the doubling time of wild-type *R. eutropha* (Table 1). However, it is interesting that *R. eutropha* AE951(pAEX3-40KI) grew much faster autotrophically than *Nitrosospira* sp. isolate 40KI and other AOBs (doubling time, about 60 h). Analysis by SDS-PAGE indicated that the RubisCO protein level in AE951 cells expressing *Nitrosospira cbbLS* genes was lower than the levels in control experiments performed with *R. eutropha cbb* genes (Fig. 5). Therefore, slower growth with *Nitrosospira* RubisCO is apparently mainly due to a lower level of expression rather than to lower specific activity. Quantitation of RubisCO protein levels and growth rates may allow comparison of in vivo specific activities of different RubisCOs in *R. eutropha*.

Tabita and coworkers also successfully used functional expression in their studies of new RubisCOs (18, 27). This approach allowed purification and characterization of type I RubisCOs from *Xanthobacter flavus* and *Bradyrhizobium japonicum* following expression of the cloned genes in a *cbb* deletion strain of the purple nonsulfur bacterium *Rhodobacter sphaeroides*. Similarly, use of the *R. eutropha* genetic system for RubisCO expression should allow generation of large quantities of cell material for protein purification. At present, the level of expression of *Nitrosospira* RubisCO protein in *R. eutropha* is considerably lower than the level obtained for the *R.*

*eutropha* enzyme (Table 1). Improved gene constructs may lead to higher expression levels, which would simplify RubisCO purification.

Exploration of the catalytic properties of natural variants of RubisCOs from diverse sources is of considerable interest. Our experimental approach should facilitate future studies of structure-function relationships of RubisCOs from *Nitrosospira* and other AOB. Furthermore, our expression system makes RubisCOs from AOB amenable to genetic manipulation.

**Comparative phylogeny.** Except for *Nitrosomonas europaea*, the phylogeny of the most conserved part of the CbbL proteins from ammonia oxidizers is comparable, but not identical, to the phylogeny of the 16S rRNA gene (Fig. 2). The differences are clearest for the  $\beta$ -subgroup organisms, because *Nitrosococcus mobilis* Nc3 clusters in the *Nitrosospira* group in the RubisCO phylogeny. CbbL proteins from additional *Nitrosomonas* spp. will be sequenced to see if this is a general phenomenon. There are also some differences in the clustering of the various *Nitrosospira* groups in this region of the trees. However, the resolution is low in both trees, especially the CbbL tree, due to the high levels of similarity of the sequences compared. As a result of this, the stability of the trees is also low, and one must be careful when drawing conclusions about how the most similar sequences relate to each other.

The most parsimonious explanation for the presence of the unique conserved six-amino-acid insert in CbbL proteins from the *Nitrosospira* group of AOB is that it was introduced only once in a common ancestor. Conserved sequence insertions or deletions (indels) in proteins may be less likely than substitutions to result from independent mutational events. Such indels are therefore particularly useful in molecular phylogenies, where they can serve as signature sequences, as discussed by Gupta (21). The insert in CbbL described here has a defined length and is flanked by conserved regions, thus providing a useful signature sequence for RubisCO. This signature sequence could be useful for defining the *Nitrosospira* group and for distinguishing members of the AOB from related chemoautotrophs. Tracing this unique insert in *cbbL* may facilitate elucidation of the evolutionary history of the *cbb* genes in these bacteria.

**Green-like RubisCO and evidence of horizontal *cbb* gene transfer.** RubisCO has a rather idiosyncratic evolutionary history that includes several likely gene duplication-horizontal gene transfer events (14, 49). Incongruence of RubisCO- and 16S rRNA-based phylogenetic trees has led to postulation of at least four horizontal gene transfer events in the *Proteobacteria*. The finding that there is a green-like RubisCO in *Rhodobacter capsulatus* and a red-like RubisCO in the closely related organism *Rhodobacter sphaeroides* provides the most compelling evidence to date that such an event occurred (37). In this case the *cbbLS* neighboring genes were probably also transferred, as they are distinctly different in the two organisms.

The nitrite oxidizer *Nitrobacter vulgaris* (a member of the  $\alpha$  subgroup of the *Proteobacteria*) has previously been shown to contain a green-like type I RubisCO (41). The *Nitrosococcus halophilus* strain used in this study, a member of the  $\gamma$  subgroup of the *Proteobacteria*, also contains a green-like RubisCO (Fig. 1 and 2). This could be expected since so far no  $\gamma$ -subgroup bacterium has been found to have a red-like RubisCO. On the other hand, we did not expect that the

RubisCO from the  $\beta$ -subgroup bacterium *Nitrosomonas europaea* also would cluster with the green-like enzymes. This organism is in all other known respects clearly related to the other  $\beta$ -subgroup AOB, the *Nitrosomonas* group in particular. RubisCO from the other representative of the *Nitrosomonas* group in this study, *Nitrosococcus mobilis* Nc3, clusters with the red-like enzymes, as expected from the 16S rDNA-based phylogenetic tree (Fig. 1 and 2). The division of the *Nitrosomonas* group is most likely a new example of a horizontal *cbb* gene transfer event. When sequencing the area flanking *cbbL* in *Nitrosomonas europaea* Nm50, we found a gene organization typical of a green-like *cbb* region (*cbbR-cbbLS-cbbQO*) with a high level of similarity to the *cbb* genes of *Thiobacillus denitrificans* and *Acidithiobacillus ferrooxidans* (unpublished results). We speculate that a cassette containing at least these five genes was transferred into the *Nitrosomonas europaea* Nm50 genome. A search of the database resulting from the ongoing genome sequencing of *Nitrosomonas europaea* ([http://www.jgi.doe.gov/JGI\\_microbial/html/nitrosomonas/nitro\\_homepage.html](http://www.jgi.doe.gov/JGI_microbial/html/nitrosomonas/nitro_homepage.html)) supported this hypothesis, since there is a difference between the G+C content of the *cbbR-cbbLS-cbbQO* gene region (~47%) and the G+C content of the sequence contig in which these genes are found (51.4%). The genome of *Nitrosomonas europaea* has a G+C content of 51.4% (32). Comparative studies of the *cbbLS* flanking genes in the other AOB may further elucidate the extent of horizontal gene transfer in these bacteria.

#### ACKNOWLEDGMENTS

This work was supported by grants 412799 and 133858/130 from The Research Council of Norway.

We thank Hans-Peter Koops and Andreas Pommerening-Röser (University of Hamburg) for kindly providing *Nitrosococcus halophilus* Nc4 and *Nitrosococcus mobilis* Nc3.

#### REFERENCES

1. Aakra, Å., J. B. Utåker, and I. F. Nes. 1999. RFLP of rRNA genes and sequencing of the 16S–23S rDNA intergenic spacer region of ammonia-oxidizing bacteria: a phylogenetic approach. *Int. J. Syst. Bacteriol.* **49**:123–130.
2. Aakra, Å., J. B. Utåker, I. F. Nes, and L. R. Bakken. 1999. An evaluated improvement of the extinction dilution method for isolation of ammonia-oxidizing bacteria. *J. Microbiol. Methods* **39**:23–31.
3. Aakra, Å., J. B. Utåker, A. Pommerening-Röser, H.-P. Koops, and I. F. Nes. 2001. Detailed phylogeny of ammonia-oxidizing bacteria determined by rDNA sequences and DNA homology values. *Int. J. Syst. Evol. Microbiol.* **51**:2021–2030.
4. Andersen, K. 1979. Mutations altering the catalytic activity of a plant-type ribulose biphosphate carboxylase/oxygenase in *Alcaligenes eutrophus*. *Biochim. Biophys. Acta* **585**:1–11.
5. Andersen, K., and M. Wilke-Douglas. 1984. Construction and use of a gene bank of *Alcaligenes eutrophus* in the analysis of ribulose biphosphate carboxylase genes. *J. Bacteriol.* **159**:973–978.
6. Andersen, K., and J. Caton. 1987. Sequence analysis of the *Alcaligenes eutrophus* chromosomally encoded ribulose biphosphate carboxylase large and small subunit genes and their gene products. *J. Bacteriol.* **169**:4547–4558.
7. Andersen, K., and M. Wilke-Douglas. 1987. Genetic and physical mapping and expression in *Pseudomonas aeruginosa* of the chromosomally encoded ribulose biphosphate carboxylase genes of *Alcaligenes eutrophus*. *J. Bacteriol.* **169**:1997–2004.
8. Andersson, I. 1996. Large structures at high resolution: the 1.6 angstrom crystal structure of spinach ribulose-1,5-biphosphate carboxylase/oxygenase complexed with 2-carboxyarabinitol biphosphate. *J. Mol. Biol.* **259**:160–174.
9. Ausubel, F. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl (ed.). 1987. *Current protocols in molecular biology*, vol. 1, p. 1.13.3. John Wiley & Sons, New York, N.Y.
10. Barraclough, D., and G. Puri. 1995. The use of  $^{15}\text{N}$  pool dilution and enrichment to separate the heterotrophic and autotrophic pathways of nitrification. *Soil Biol. Biochem.* **27**:17–22.

11. Bock, E., H.-P. Koops, and H. Harms. 1986. Cell biology of nitrifying bacteria, p. 17–38. In J. I. Prosser (ed.) *Nitrification—1986*. IRL Press, Oxford, United Kingdom.
12. Chambers, S. P., S. E. Prior, D. E. Barstow, and N. P. Minton. 1988. The pMTL *nic*<sup>-</sup> cloning vectors. I. Improved pUC polylinker regions to facilitate the use of sonicated DNA for nucleotide sequencing. *Gene* **68**:139–149.
13. Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA* **81**:1991–1995.
14. Delwiche, C. F., and J. D. Palmer. 1996. Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids. *Mol. Biol. Evol.* **13**:873–882.
15. Ditta, G., T. Schmidhauser, E. Yakobson, P. Lu, X.-W. Liang, D. R. Finlay, D. Guiney, and D. R. Helinski. 1985. Plasmids related to the broad host range vector, pRK290, useful for gene cloning and for monitoring gene expression. *Plasmid* **13**:149–153.
16. Dower, W. J., J. F. Miller, and C. W. Ragsdale. 1988. High efficiency transformation of *E. coli* by high voltage electroporation. *Nucleic Acids Res.* **16**:6127–6145.
17. Easter, C. L., P. A. Sobecky, and D. R. Helinski. 1997. Contribution of different segments of the *par* region to stable maintenance of the broad-host-range plasmid RK2. *J. Bacteriol.* **179**:6472–6479.
18. Falcone, D. L., and F. R. Tabita. 1993. Expression and regulation of *Bradyrhizobium japonicum* and *Xanthobacter flavus* CO<sub>2</sub> fixation genes in a photosynthetic bacterial host. *J. Bacteriol.* **175**:866–869.
19. Felsenstein, J. 1993. PHYLIP (phylogeny inference package), version 3.572c. Department of Genetics, University of Washington, Seattle.
20. Goebel, B. M., and E. Stackebrandt. 1994. Cultural and phylogenetic analysis of mixed microbial populations found in natural and commercial bioleaching environments. *Appl. Environ. Microbiol.* **60**:1614–1621.
21. Gupta, R. S. 1998. Protein phylogenies and signature sequences: a reappraisal of evolutionary relationships among archaeobacteria, eubacteria, and eukaryotes. *Microbiol. Mol. Biol. Rev.* **62**:1435–1491.
22. Hansen, S., E. Hough, and K. Andersen. 1999. Purification, crystallization and preliminary X-ray studies of two isoforms of Rubisco from *Alcaligenes eutrophus*. *Acta Crystallogr. Sect. D* **55**:310–313.
23. Hansen, S., V. Burkow Volland, E. Hough, and K. Andersen. 1999. The crystal structure of rubisco from *Alcaligenes eutrophus* reveals a novel central eight-stranded  $\beta$ -barrel formed by  $\beta$ -strands from four subunits. *J. Mol. Biol.* **288**:609–621.
24. Hartman, F. C., and M. R. Harpel. 1994. Structure, function regulation and assemble of D-ribulose-1,5-bisphosphate carboxylase/oxygenase. *Annu. Rev. Biochem.* **63**:197–234.
25. Hatayama, R., R. Takahashi, M. Ohshima, R. Shibasaki, and T. Tokuyama. 2000. Ribulose-1,5-bisphosphate carboxylase/oxygenase from an ammonia-oxidizing bacterium, *Nitrosomonas* sp. K1: purification and properties. *J. Biosci. Bioeng.* **90**:426–430.
26. Head, I. M., W. D. Hiorns, T. M. Embley, A. J. McCarthy, and J. R. Saunders. 1993. The phylogeny of autotrophic ammonia-oxidizing bacteria as determined by analysis of 16S ribosomal RNA gene sequences. *J. Gen. Microbiol.* **139**:1147–1153.
27. Horken, K. M., and F. R. Tabita. 1999. Closely related form I ribulose bisphosphate carboxylase/oxygenase molecules that possess different CO<sub>2</sub>/O<sub>2</sub> substrate specificities. *Arch. Biochem. Biophys.* **361**:183–194.
28. Jeffke, T., N. H. Gropp, C. Kaiser, C. Grzeszik, B. Kusian, and B. Bowien. 1999. Mutational analysis of the *cbb* operon (CO<sub>2</sub> assimilation) promoter of *Ralstonia eutropha*. *J. Bacteriol.* **181**:4374–4380.
29. Jiang, Q. Q., and L. R. Bakken. 1999. Comparison of *Nitrosospora* strains isolated from terrestrial environments. *FEMS Microbiol. Ecol.* **30**:171–186.
30. Kelly, D. P., and A. P. Wood. 2000. Reclassification of some species of *Thiobacillus* to the newly designated genera *Acidithiobacillus* gen. nov., *Halotheiobacillus* gen. nov. and *Thermothiobacillus* gen. nov. *Int. J. Syst. Evol. Microbiol.* **50**:511–516.
31. Koops, H.-P., B. Böttcher, U. C. Möller, A. Pommerening-Röser, and G. Stehr. 1990. Description of a new species of *Nitrosococcus*. *Arch. Microbiol.* **154**:244–248.
32. Koops, H. P., and H. Harms. 1985. Deoxyribonucleic acid homologies among 96 strains of ammonia-oxidizing bacteria. *Arch. Microbiol.* **141**:214–218.
33. Kusian, B., and B. Bowien. 1997. Organization and regulation of *cbb* CO<sub>2</sub> assimilation genes in autotrophic bacteria. *FEMS Microbiol. Rev.* **21**:135–155.
34. Lillehaug, D., B. Lindqvist, and N. K. Birkeland. 1991. Characterization of LC3, a *Lactococcus lactis* subsp. *cremoris* temperate bacteriophage with cohesive single-stranded DNA ends. *Appl. Environ. Microbiol.* **57**:3206–3211.
35. Maeda, N., K. Kitano, T. Fukui, S. Ezaki, H. Atomi, K. Miki, and T. Imanaka. 1999. Ribulose bisphosphate carboxylase/oxygenase from the hyperthermophilic archaeon *Pyrococcus kodakaraensis* KOD1 is composed solely of large subunits and forms a pentagonal structure. *J. Mol. Biol.* **293**:57–66.
36. Makrides, S. C. 1996. Strategies for achieving high-level expression of genes in *Escherichia coli*. *Microbiol. Rev.* **60**:512–538.
37. Paoli, G. C., F. Soyer, J. Shively, and F. R. Tabita. 1998. *Rhodobacter capsulatus* genes encoding form I ribulose-1,5-bisphosphate carboxylase/oxygenase (*cbbLS*) and neighboring genes were acquired by a horizontal gene transfer. *Microbiology* **144**:219–227.
38. Rose, T. M., E. R. Schultz, J. G. Henikoff, S. Pietrovski, C. M. McCallum, and S. Henikoff. 1998. Consensus-degenerate hybrid oligonucleotide primers for amplification of distantly related sequences. *Nucleic Acids Res.* **26**:1628–1635.
39. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. *Molecular cloning: a laboratory manual*, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
40. Schell, M. A. 1993. Molecular biology of the LysR family of transcriptional regulators. *Annu. Rev. Microbiol.* **47**:597–626.
41. Shively, J. M., G. van Kanten, and W. G. Meijer. 1998. Something from almost nothing: carbon dioxide fixation in chemoautotrophs. *Annu. Rev. Microbiol.* **52**:191–230.
42. Tabita, F. R. 1999. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a different perspective. *Photosynth. Res.* **60**:1–28.
43. Teske, A., E. Alm, J. M. Regan, S. Toze, B. E. Rittman, and D. A. Stahl. 1994. Evolutionary relationships among ammonia- and nitrite-oxidizing bacteria. *J. Bacteriol.* **176**:6623–6630.
44. Thompson, J. D., D. G. Higgins, and T. J. Gibson. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**:4673–4680.
45. Tokuyama, T., N. Yoshida, T. Matsuishi, N. Takahashi, R. Takahashi, T. Kanehira, and M. Shinohara. 1997. A new psychrotrophic ammonia-oxidizing bacterium, *Nitrosovibrio* sp. TYM9. *J. Ferment. Bioeng.* **83**:377–380.
46. Utåker, J. B., L. Bakken, Q. Q. Jiang, and I. F. Nes. 1995. Phylogenetic analysis of seven new isolates of ammonia-oxidizing bacteria based on 16S rRNA gene sequences. *Syst. Appl. Microbiol.* **18**:549–559.
47. Utåker, J. B., and I. F. Nes. 1998. A qualitative evaluation of the published oligonucleotides specific for the 16S rRNA gene sequences of the ammonia-oxidizing bacteria. *Syst. Appl. Microbiol.* **21**:72–88.
48. van der Meer, J. R., C. Werlen, S. F. Nishino, and J. C. Spain. 1998. Evolution of a pathway for chlorobenzene metabolism leads to natural attenuation in contaminated groundwater. *Appl. Environ. Microbiol.* **64**:4185–4193.
49. Watson, G. M. F., and F. R. Tabita. 1997. Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation. *FEMS Microbiol. Lett.* **146**:13–22.
50. Watson, G. M. F., J. P. Yu, and F. R. Tabita. 1999. Unusual ribulose 1,5-bisphosphate carboxylase/oxygenase of anoxic *Archaea*. *J. Bacteriol.* **181**:1569–1575.
51. Woese, C. R., W. G. Weisburg, B. J. Paster, C. M. Hahn, R. S. Tanner, N. Krieg, H.-P. Koops, and E. Stackebrandt. 1984. The phylogeny of the purple bacteria: the beta subdivision. *Syst. Appl. Microbiol.* **5**:327–336.
52. Wullenweber, M., H.-P. Koops, and H. Harms. 1977. Polyhedral inclusion bodies in cells of *Nitrosomonas* sp. *Arch. Microbiol.* **112**:69–72.
53. Yanisch-Perron, C., J. Vieira, and J. Messing. 1985. Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene* **33**:103–119.