Unusual ribulose-1,5-bisphosphate carboxylase/oxygenase genes from a marine manganese-oxidizing bacterium

Ron Caspi, Margo G. Haygood and Bradley M. Tebo

Author for correspondence: Bradley M. Tebo. Tel: +1 619 534 5470. Fax: +1 619 534 7313. e-mail: btebo@ucsd.edu

Marine Biology Research Division and Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0202, USA

The Gram-negative bacterium strain S185-9A1 is a novel marine α-proteobacterium that oxidizes manganese(II) to manganese(IV). Initial DNA hybridization screening showed that S185-9A1 possesses a gene similar to cbbL, the gene coding for the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO; EC 4.1.1.39). However, no RubisCO enzyme activity was found in cultures of S185-9A1. Genes coding for both large (cbbL) and small (cbbS) subunits of a RubisCO enzyme were identified, isolated and sequenced. When these genes were introduced into an Escherichia coli host strain, ribulose-1,5-bisphosphate-dependent CO₂ fixation occurred under control of a lac promoter, indicating that the protein is functional in E. coli. Although their function is unknown, this is the first direct evidence for the presence of RubisCO genes in a manganese-oxidizing bacterium. Phylogenetic analysis of the RubisCO genes of strain S185-9A1 showed that they are divergent, but are more related to those from non-chlorophyte algal chloroplasts than are those from other bacteria. The fact that the RubisCO sequence of strain S185-9A1 is not closely related to any other published RubisCO sequence suggests that the protein may be valuable for studies of the function and evolution of the RubisCO enzyme as well as its activity in the environment.

Keywords: ribulose-1,5-bisphosphate carboxylase/oxygenase, RubisCO, manganese oxidation, autotrophy

INTRODUCTION

Ribulose-1,5-bisphosphate carboxylase (RubisCO; EC 4.1.1.39) is one of two unique enzymes in the Calvin–Benson cycle, the most common pathway for autotrophic CO₂ fixation in plants and bacteria. Two forms (I and II) are known, and form I RubisCOs have been categorized into four types (Tabita, 1995). These forms and types are based on evolutionary relationships, and also reflect common biochemical features among the groups. Chloroplasts of terrestrial plants and green algae together with cyanobacteria contain RubisCOs that form a related group (Type IB), consistent with the accepted notion of the cyanobacterial origin of chloroplasts. Many marine algal chloroplasts (non-chlorophyte algae) contain RubisCOs (Type ID) that are only distantly related to the cyanobacterial type, and are more closely related to other bacterial RubisCOs (Type IC) than to other chloroplasts. The fourth type (IA) is a group of mostly proteobacterial RubisCOs allied with Type IB, although recently a prochlorophyte was found to have a RubisCO that belongs to this group (Shimada et al., 1995).

Manganese (Mn)-oxidizing bacteria have long been suspected of having autotrophic potential, but the existence of RubisCO genes in these organisms has never been investigated. Such genes may not only provide evidence consistent with the existence of autotrophic Mn oxidation, but may also be useful in understanding the evolution and biochemical function of RubisCO. It was proposed as early as 1913 (Beijerinck, 1913) that the oxidation of reduced Mn [Mn(II)] could provide energy for chemo-
lithoautotrophic growth. Although autotrophy supported by Mn oxidation is thermodynamically favourable, there is only one report that clearly demonstrates net carbon fixation by a Mn oxidizer (Kepkay & Nealson, 1987). Unfortunately, that organism, *Pseudomonas* sp. strain S-36, does not demonstrate consistent Mn oxidation in our hands. We have screened 45 Mn-oxidizing environmental isolates with probes constructed from *cbbL* (RubisCO large subunit) genes of both an *Anabaena* sp. and *Xanthobacter* strain H4-14, with the aim of finding potential autotrophic strains (Tebo & Haygood, 1989). About 20% of the isolates hybridized positively, and strain S185-9A1 was selected because of its consistent and reliable Mn oxidation.

Restriction analysis with two enzymes followed by DNA hybridization suggested that this strain has only one copy of the RubisCO form I genes, and hybridization with a probe derived from the *cbbM* gene of *Rhodopseudomonas viridis*, which codes for a form II RubisCO, yielded negative results, suggesting that S185-9A1 has only one copy of RubisCO genes in the genome.

Strain S185-9A1 was isolated in 1985 from Saanich Inlet, a stratified fjord on the west coast of Canada. The fjord is anoxic below ~130 m for about six months every year, resulting in an 

\[ O_2/H_2S \]

interface (Tebo & Emerson, 1985). Microbial catalysis of Mn oxidation has been demonstrated in this basin above the interface, where there is an accumulation of particulate Mn (Emerson et al., 1984). S185-9A1 was isolated from water collected at 125 m, within the particulate Mn maximum. When grown on an appropriate seawater-based medium containing Mn(II), the bacteria oxidize Mn(II) and precipitate Mn(IV) manganates on the cell surface, forming a coat around the cell. It is a Gram-negative rod-shaped bacterium, which, according to its 16S rRNA sequence, belongs to the α-subdivision of the proteobacteria. It is not closely related to any of the organisms in the Genbank and RDP databases (GenBank accession number U53824; Maidak et al., 1994). S185-9A1 grows very slowly, with a doubling time of about 1-2 d, depending on the medium. The final population density in liquid cultures depends on the content of organic nutrients in the medium, and can reach high values (more than 10⁸ cells ml⁻¹) when grown in a rich medium. However, the bacteria also grow with the carbon source: a compound of the formate of the sole organic carbon and energy source. Mn oxidation occurs only when grown in organic-poor media, and at the onset of the stationary phase of growth in batch cultures.

The initial probing results suggested that strain S185-9A1 possesses a gene similar to *cbbL*, but we were unable to demonstrate *in vitro* ribulose-1,5-bisphosphate-dependent CO₂ fixation activity or autotrophic growth of the organism. We have proceeded by a molecular approach, cloning, sequencing, and expressing the genes, and thus confirming that the RubisCO genes of strain S185-9A1 do indeed code for a RubisCO enzyme that is functional, and might, under unknown conditions, be expressed to yield a functional protein in the parent organism.

**METHODS**

**Bacterial strains, media and growth conditions.** Strains and plasmids used in this study are summarized in Table 1. Strain S185-9A1 was maintained on solid media K (Krumbein, 1971) and M which consists of artificial sea water (ASW; see below) supplemented with 50 mg yeast extract l⁻¹, 50 mg peptone l⁻¹, 20 mM HEPES buffer (pH 7.8), 100 μM MnCl₂, 2 mM KHCO₃, and 15 g Bacto Noble agar (Difco) l⁻¹. ASW is prepared in our laboratory at double strength (2 x ASW) as follows: 247 g MgSO₄·7H₂O, 2.9 g CaCl₂·2H₂O, 35.1 g NaCl and 1.5 g KCl are dissolved individually in 250 ml distilled water each and then combined yielding 1:2 x ASW, which is diluted by 50% with distilled water and other components in the final medium.

For large scale DNA preparations the cells were grown in seawater complete (SWC) medium (Nealson, 1978) at room temperature for 1 week. For RubisCO activity assays cells were grown at room temperature for 1 week in J medium [ASW supplemented with 20 mM HEPES buffer (pH 7.8), 1.5 mM NH₄Cl, 2 mM KHCO₃, 10 ml vitamin mix (Kepkay, 1985), 73 μM KH₂PO₄ and 0.1 ml of a stock solution of 3 mg ferrous ammonium citrate (ml ASW)⁻¹] supplemented with 20 mM formate. *Escherichia coli* strains were grown at 37 °C in LB medium (Sambrook et al., 1989); *Thiothrix neapolitana* cells were grown in TMN-2 medium (Starr et al., 1981) at room temperature.

**Probing of environmental isolates.** Forty-five strains of Mn-oxidizing bacteria isolated from a variety of marine environments were screened with probes constructed from fragments of the *cbbL* genes from four organisms: a 1.5 kb *P. pestis*-EcoRI fragment of the plasmid pANP1155, which contains a 2.3 kb *P. pestis* fragment of the *cbbL* gene of *Anacystis nidulans* 6301 (Shinosaki & Sugura, 1983), a 0.95 kb *H. paludis*-HindIII fragment internal to *cbbL* from the plasmid pAn600, which contains a 17 kb fragment including the *cbbL* gene of *Anabaena* 7120 (Curtis & Haselkorn, 1983), a 14 kb EcoRI-BglII fragment of the plasmid pRR2119, which contains a 2.4 kb fragment of the *cbbM* gene of *R. rubrum* (Somerville & Somerville, 1984), and a 0.9 kb *S. malachitica*-SalI fragment from the plasmid pLM417R, which contains a 1.7 kb region of the *cbbL* gene of *Xanthobacter* strain H4-14 (Lehmke & Lister, 1985). Hybridization conditions were adjusted to the highest stringency that would still allow detection of a *T. neapolitana* positive control with the *Anabaena* probe.

**Construction of a genomic DNA library and subcloning the RubisCO genes.** Strain S185-9A1 cells are very difficult to lyse. In order to obtain large quantities of high molecular mass DNA we used the following procedure. Cells were grown in 1 l SWC medium without Mn to stationary phase (in the presence of Mn the cells become coated with Mn oxides, making them even more difficult to lyze). The cells were then spun down and resuspended in 95 ml TE buffer pH 8.0 (50 mM Tris buffer pH 8.0, 10 mM EDTA). The suspension was divided into 10 tubes and 200 μl lysozyme solution (50 mg ml⁻¹) was added to each. The cells were stored on ice for 10 min, then 0.5 ml 10% (w/v) SDS and 30 μl proteinase K solution (20 mg ml⁻¹) were added to each tube and the cells incubated at 37 °C for a few hours, until clear. The clear lysate was extracted with CTAB (hexadeyltrimethylammonium bromide, Sigma) and chloroform (Ausubel et al., 1987) and the DNA precipitated, dissolved in TE buffer, and purified on a CsCl density gradient. The DNA was partially digested with the restriction enzyme *Sal*3A, and fragments larger than 20 kb were purified on a sucrose gradient. The size-fractionated DNA was ligated to plpMMB33 cosmid arms as described by Frey et al. (1983), and packaged with a Gigapack XL packaging kit (Stratagene). The packaged library
was transduced into *E. coli* XL1-Blue cells. The resulting library had about 200000 c.f.u. and a mean insert size of 20 kb.

Ten thousand colonies representing the gene library were probed by Southern hybridization with the *Xanthobacter* sp. strain H-4-14 probe described above, and seven positive colonies were isolated. The cosmid of one of them, pMhfB33-7, which contained an insert of ~29 kb, was digested with different restriction enzymes and reprobed. Two positive bands were isolated, a 2013 bp *XhoI* fragment, which was cloned into *XhoI*-digested *pBluescript* SK -- both orientations, resulting in pRC1 and pRC1R, and a 1678 bp *PstI* fragment, which was cloned into *PstI*-digested *pBluescript* SK -- resulting in pRC3 and pRC3R (Fig. 2). A similarly digested chromosomal DNA yielded positive bands of the same size, suggesting that no rearrangement had occurred. Both pRC1 and pRC1R were subjected to nested deletions, using an Erase a Base kit (Promega), resulting in plasmids pRC1-1 to pRC1-10 and pRC1R-1 to pRC1R-10, respectively. Since it was found that plasmids pRC1 and pRC3 did not contain the whole *cbbL* gene, a synthetic oligonucleotide was prepared from the sequence at the end of the *SalI* fragment (cbbL1), and used to probe pMhfB33-7 again. A positive *SalI* fragment of 2.5 kb was isolated and cloned into *pBluescript* KS --, resulting in pRC4. This plasmid contained the rest of the *cbbL* gene, and also the *cbbS* gene.

To generate a clone containing the intact *cbbL/*cbbS region, the *SalI* fragment of pRC4 was cut and cloned into the 5.1 kb *SalI* fragment from pRC3. The resulting plasmid was named pRC10. In order to fuse this insert to the *lac* promoter of *pBluescript*, the whole insert of pRC10 was cut by simultaneous digestion with *KpnI* and *SacI*, which cut outside the insert but within the multiple cloning site of *pBluescript*, and cloned into a *KpnI*- *SacI*-digested *pBluescript* SK --. The resulting plasmid was named pRC11.

**Southern hybridization.** DNA probes were either prepared from double-stranded DNA using a Random Primed DNA Labeling kit (Boehringer Mannheim), or oligonucleotides were synthesized by a DNA synthesizer (Applied Biosystems, model 391) and tagged by a DNA Tailing kit (Boehringer Mannheim) using high specific activity (>111 TBq mmol⁻¹) [γ³²P]dATP (ICN Biochemicals). DNA was transferred to membranes using a downward transfer protocol (Koetsier et al., 1993).

**16S rRNA sequencing and analysis.** Overlapping fragments of the 16S rRNA gene were obtained by PCR amplification using standard methods and primers (Lane, 1990). Sequencing was done with an automated DNA sequencer (ABI model 373A) using a PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems). The sequence was aligned to other sequences using the RDP World Wide Web server (Maidak et al., 1994). Aligned sequences were imported into PAUP version 3.0s (Swofford, 1991 and a phylogenetic tree was generated by a heuristic search using 10 replications of random stepwise additions.

**RubisCO sequencing and analysis.** Sequencing was done either manually or by an automated sequencer. Single-stranded DNA for manual sequencing was obtained from *pBluescript* phagemids (Stratagene) according to the manufacturer's instructions, and both strands sequenced with Sequenase Labeling kit (Boehringer Mannheim), or oligonucleotides were prepared from double-stranded DNA using a Random Primed DNA Labeling kit (Boehringer Mannheim), or oligonucleotides were synthesized by a DNA synthesizer (Applied Biosystems, model 391) and tagged by a DNA Tailing kit (Boehringer Mannheim) using high specific activity (>111 TBq mmol⁻¹) [γ³²P]dATP (ICN Biochemicals). DNA was transferred to membranes using a downward transfer protocol (Koetsier et al., 1993).

**16S rRNA sequencing and analysis.** Overlapping fragments of the 16S rRNA gene were obtained by PCR amplification using standard methods and primers (Lane, 1990). Sequencing was done with an automated DNA sequencer (ABI model 373A) using a PRISM Ready Reaction DyeDeoxy Terminator Cycle Sequencing Kit (Applied Biosystems). The sequence was aligned to other sequences using the RDP World Wide Web server (Maidak et al., 1994). Aligned sequences were imported into PAUP version 3.0s (Swofford, 1991 and a phylogenetic tree was generated by a heuristic search using 10 replications of random stepwise additions.

**RubisCO sequencing and analysis.** Sequencing was done either manually or by an automated sequencer. Single-stranded DNA for manual sequencing was obtained from *pBluescript* phagemids (Stratagene) according to the manufacturer's instructions, and both strands sequenced with Sequenase version 2.0 (United States Biochemicals). Double-stranded DNA for automated sequencing was obtained with Magic Mini

---

**Table 1. Bacterial strains and plasmids used in this work**

<table>
<thead>
<tr>
<th>Strain or plasmid</th>
<th>Genetic characteristic(s)</th>
<th>Source or reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Strains</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SI85-9A1</td>
<td>Manganese oxidizer</td>
<td>This study</td>
</tr>
<tr>
<td><em>Thiothrix neapolitana</em> ATCC 23638</td>
<td></td>
<td></td>
</tr>
<tr>
<td><em>Escherichia coli</em></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DH1</td>
<td><em>supE44 lacR17 recA1</em></td>
<td>Hanahan (1983)</td>
</tr>
<tr>
<td>XL1-Blue</td>
<td><em>lacI lacZΔM15</em></td>
<td>Stratagene</td>
</tr>
<tr>
<td>Plasmids</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pLL417R</td>
<td>Contains the <em>cbbL</em> gene of <em>Xanthobacter</em> strain H4-14</td>
<td>Lehnice &amp; Lidstrom (1985)</td>
</tr>
<tr>
<td>pANP115S</td>
<td>Contains the <em>cbbL</em> gene of <em>Anoxyatus nidulans</em> GS01</td>
<td>Shinozaki &amp; Sugura (1983)</td>
</tr>
<tr>
<td>pAN600</td>
<td>Contains the <em>cbbL</em> gene of <em>Anabaena</em> 7120</td>
<td>Curtis &amp; Haselkorn (1983)</td>
</tr>
<tr>
<td>pRR2119</td>
<td>Contains the <em>cbbL</em> gene of <em>Rhodospirillum rubrum</em></td>
<td>Somerville &amp; Somerville (1984)</td>
</tr>
<tr>
<td>pMMB33</td>
<td><em>IncQ</em> broad host range, Km&lt;sup&gt;c&lt;/sup&gt; cos&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Frey et al. (1983)</td>
</tr>
<tr>
<td>pMMB33-7</td>
<td><em>pMMB33</em> containing a 29 kb fragment of SI85-9A1 chromosome</td>
<td>This study</td>
</tr>
<tr>
<td><em>pBluescript</em> II KS --</td>
<td><em>SacI</em></td>
<td>Stratagene</td>
</tr>
<tr>
<td>pRC1</td>
<td><em>pBluescript</em> containing a <em>XhoI</em> fragment of pMMB33-7</td>
<td>This study</td>
</tr>
<tr>
<td>pRC1R</td>
<td>As <em>pRC1</em> with the <em>XhoI</em> fragment in the opposite orientation</td>
<td>This study</td>
</tr>
<tr>
<td>pRC3</td>
<td><em>pBluescript</em> containing a <em>PstI</em> fragment of pMMB33-7</td>
<td>This study</td>
</tr>
<tr>
<td>pRC3R</td>
<td>As <em>pRC3</em> with the <em>PstI</em> fragment in the opposite orientation</td>
<td>This study</td>
</tr>
<tr>
<td>pRC4</td>
<td><em>pBluescript</em> containing a <em>SalI</em> fragment of pMMB33-7</td>
<td>This study</td>
</tr>
<tr>
<td>pRC10</td>
<td><em>pBluescript</em> containing both <em>cbbL</em> and <em>cbbS</em></td>
<td>This study</td>
</tr>
<tr>
<td>pRC11</td>
<td><em>pBluescript</em> containing both <em>cbbL</em> and <em>cbbS</em> fused to the <em>lac</em> promoter</td>
<td>This study</td>
</tr>
</tbody>
</table>
Prep Columns (Promega). Resulting DNA sequences were analysed and translated by the MacVector Sequence Analysis Software version 3.5 (International Biotechnologies), and translated sequences were analysed by PAUP. Sequence alignments were performed manually using Microsoft Excel colour macros (Haygood, 1993). Nucleotide sequence identity was calculated using the alignment feature of the MacVector program. Pairs of SI85-9A1 and each of the other taxa were aligned, identical nucleotides were determined, and percentage identity calculated as the number of identical nucleotides divided by the length of the shorter sequence. Phylogenetic trees were generated from aligned amino acid sequences by heuristic searches, using 10 then 200 steps. Phylogenetic trees were generated by PAUP. Branch lengths indicate the inferred number of character state changes on each branch.

Sequences of Rubisco genes from other organisms used in this paper were chosen from those for which the sequence is available for both subunits, and so will represent all the major groups of Rubisco. The following sequences were used (GenBank accession numbers are given in parentheses): A. calidugoneus (M17744) (Andersen & Caron, 1987), Anacystis nidulans strain 6301 (X03220) (Shinozaki & Sugiyara, 1983), Chlamydomonas reinhardtii (X04471, J03997) (Dron et al., 1982; Goldschmidt-Clermont & Rahire, 1986), Cryptomonas (X14171, X62349) (Douglas & Durnford, 1989; Douglas et al., 1990), Cyanidium caldarium strain RK-15 (X55524) (Valentini & Zetsche, 1990a), Cytophaga sp. strain N1 (M59080) (Hwang & Tabita, 1991), Ectocarpus siliculosus (X52803) (Valentini & Zetsche, 1990b), Olisthodiscus luteus (M24288, X61918) (Boczar et al., 1989; Harrison et al., 1992), Porphyridium aerugineum (X17597) (Valentini & Zetsche, 1989), Rhizobium sp. (M64624) (Gilson et al., 1991), R. rubrum (X00286) (Nargang et al., 1984), Synechococcus sp. strain a-1 (D13539) (Yaguchi et al., 1993), Thiobacillus ferrooxidans (M85906) (Kusano et al., 1991), Xanthobacter flavus (X17252) (Meijer et al., 1991), and Zue minus (Y00322, V00171) (M. Lebrun, G. Waksman & B. Freyssinet, Genbank submission; McEntosh et al., 1980).

Rubisco assays. These were modified from a previously published assay (Glover & Morris, 1979). Strain SI85-9A1 cultures were grown for 1 week. T. neapolitanus cultures were grown for 3 days, and E. coli cultures were grown overnight. In the morning of the experiment two aliquots of 6 ml of each E. coli culture were transferred into sterile tubes. One millilitre of LB medium containing 14 mM IPTG was added to one of the aliquots (2 mM final concentration), while 1 ml of LB medium without IPTG was added to the other. The tubes were incubated at 37 °C for 2 h to allow expression of genes fused to the lac promoter. Assay mixtures were prepared by filtering two duplicates of each culture onto Whatman GFF filters. The volumes used were 1 ml for E. coli cultures, 10 ml for T. neapolitanus and 25 ml for strain SI85-9A1. The filters were added to scintillation vials, 400 μl 10% (v/v) Triton X-100 added and the filters incubated for 10 min. An exception was with strain SI85-9A1 cells, when 200 μl lysozyme in TE (5 mg ml⁻¹) were added first, the cells incubated on ice for 10 min, and then 200 μl 20% Triton X-100 was added and the cells incubated for 2 h at room temperature. When the assay mixtures were ready, 0.93 ml reaction buffer (70 mM Tris pH 8, 25 mM glutathione, 25 mM MgCl₂, 40 mM NaHCO₃, 3.07 μM Bq H¹¹CO₃⁻) was added to each vial and the vials were incubated at room temperature for 10 min to activate the enzyme. For each sample, 50 μl of a solution containing 4 mg ribulose 1,5-bisphosphate (RuBP) ml⁻¹ (dissolved in 1:5-diluted reaction buffer without radiotracer and adjusted to pH 6) was added to one duplicate (360 μM final concentration), while 50 μl of the same buffer without RuBP was added to the second duplicate.

The reactions were allowed to proceed for 1 h at room temperature, then terminated by adding 3 ml acetic acid/methanol (1:20, v/v) and dried at 65 °C overnight. Total radioactivity was measured by adding 5 μl of the stock solution to 0.5 ml β-phenylethylamine, then adding 10 ml Scintillation cocktail (ScintiVerse BD, Fisher Scientific) to all samples and counting for 5 min in a Beckman LS 6000TA scintillation counter.

Total protein concentrations. These were determined using the BCA Protein Assay kit (Pierce), according to the manufacturer's instructions. Samples were taken immediately before filtering cells for the Rubisco activity assay, centrifuged and the pellets frozen. The pellets were later resuspended in TE buffer, and NaOH and SDS were added to 0.1 M and 1%, respectively. Strain SI85-9A1 cells were incubated at 55 °C overnight; T. neapolitanus and E. coli cells lysed immediately.

RESULTS

16S RNA sequence

The 16S rRNA sequence analysis placed strain SI85-9A1 within the α-subdivision of the proteobacteria. It is not closely related to any known species, but the closest organisms as determined by the Similarity Rank test are members of the rhizobia group, such as Rhizobium ciceri (S₀ 0.763) and Rhizobium huakuii (S₀ 0.750) (Fig. 1).

\[ \text{Fig. 1. Relationships among different proteobacteria based on their 16S rRNA sequences, as determined by maximum parsimony (PAUP; Swofford, 1991). Strain SI85-9A1 belongs to the α-subdivision of that group. The closest organism to strain SI85-9A1 as determined by the Similarity Rank test is Rh. ciceri (S₀ 0.763). The arrow indicates position of root (determined with Bacillus subtilis as outgroup). Branch lengths indicate the inferred number of character state changes on each branch. Sequences other than strain SI85-9A1 were obtained from the RDP World Wide Web Server (Maidak et al., 1994).} \]
RubisCO genes from a manganese-oxidizing bacterium

RubisCO sequence

The region sequenced is shown in Fig. 2. The region containing the cbbL and cbbS genes was fully sequenced on both strands (Fig. 3). Two additional ORFs were identified from the partial sequence (Fig. 2). The first (partial) ORF contains 762 bp which encode the carboxyl-terminus 254 amino acids of a putative Calvin cycle aldolase (cbbA). The nucleotide sequence of this gene (not shown) closely resembles (75% identity) that of the cbbA gene of Rh. sphaeroides (Gibson et al., 1991). This putative gene ends with a TGA codon and is followed by a non-coding region of 173 bp.

The second ORF is 1461 bp long (bases 176–1636), and encodes a 486 amino acid CbbL subunit. This ORF is preceded by a Shine–Dalgarno sequence (GAAGGA) upstream from ATG, the start codon. The cbbL gene is followed by a non-coding region of 114 bp (bases 1637–1750).

The third ORF is 423 bp long (bases 1751–2173), and encodes a 140 amino acid CbbS subunit. This ORF is preceded by a Shine–Dalgarno sequence (GAAGAGGA) located upstream from the ATG starting codon. The cbbS gene is followed by a non-coding region of 129 bp.

The fourth (partial) ORF (orfX) is 100 bp long, and encodes the amino-terminal 33 amino acids (not shown) of an unidentified protein. This ORF is also preceded by a Shine–Dalgarno sequence (GAGGAG).

The DNA sequence of the cbb genes is most similar to RubisCO genes belonging to the Type IC group, namely Rh. sphaeroides, X. flavus and Az. euophorus. Table 2 shows the percentage identity of these genes to RubisCO genes from the various groups.

Parsimony analysis

In order to analyse these genes, we have aligned the amino acid sequences generated by translation of the cbbL and cbbS genes to published sequences from several organisms (Fig. 4). The amino acid sequence generated by translation of strain S185-9A1 cbbL gene possesses all the amino acids that were found to be implicated in activation and catalysis of the enzyme (Knight et al., 1990). The aligned sequences were analysed by maximum parsimony to generate phylogenetic trees. We generated trees for CbbS, CbbL, and CbbL and CbbS combined, but only the tree constructed for CbbL is shown (Fig. 5) since the information contained in the CbbS region alone was not sufficient to resolve phylogenetic relationships among some of the organisms, and the result for the combined sequence was similar to that obtained from CbbL alone. The tree was generated using the CbbM protein of Rs. rubrum (which belongs to the Type II group) as an outgroup. The tree has the following parameters: length
1 GAGGCGCGCTCCGCGGCGTCTGCCGCGCGCCGGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
101 GAGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
91 MDQGSOTXK  
100 GAGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
G K D R Y S G V M E Y R K M M G Y W E P D Y E P K E T D V I A C F  
110 AGGGCAAGGCGCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
R I T P Q D G V D P I R A A A A A V A G E S S T A T W T V V W T D R  
120 CCGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
L T A A E K Y R A S I T Q N V G F K P L K P L R R E D M R L  
110 CTGAGGCGCGGCTCTAGAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
F E N G S I A N L T A S I I N Q V G F K P L K P L R R E D M R L  
140 CTGAGGCGCGGCTCTAGAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
P T A V Y R K P Q P G P A T G I G V V R H R L D K P G C F R P L L G A T  
170 GCGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
V R P K L G L G S G Y R G V Y V B A L K G G L D F T K D D E N I N  
170 GCGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
S Q P F M D W W E R F L B Y A R E A N K G T E G I K G T Y L  
180 ACGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
190 ACGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
T A M Q S N A K W A R R N N X I L H L R H A G R S T Q K T H G  
200 ACGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
A 101 ACGGAGGAGCGCTCTCTCGAAGGGGCGGCGTCTGCCGCGCGCCGGCGCGCGCCGCAAGAACGAGATATTCCGAATGGACGGCGCGGACCGCCTCTCTCGC  
Fig. 3. Nucleotide sequence and deduced amino acid sequence of the cbbL and cbbS region of the insert in pRC10. Predicted amino acid residues are shown above the respective codons. Putative ribosome-binding sites are overlined and stop codons are indicated with asterisks. There is no restriction site at the end of the sequence.

Expression of SI85-9A1 RubisCO in E. coli

Because we were unable to detect RubisCO enzyme activity in cells of strain SI85-9A1 we thought that the SI85-9A1 RubisCO genes might be nonfunctional. To test this possibility we cloned the genes into an expression system in E. coli and assayed for activity (Table 3). The RubisCO expression experiment included the bacterium T. neapolitanus as a positive control, and the E. coli strain XLI-Blue carrying the plasmid pBluescript as a negative control. The control cells, together with XLI-Blue cells carrying the plasmids pRC10 and pRC11, were incubated for 1 h with and without the lac gene inducer IPTG, and with and without RuBP, in the presence of H14CO3-.

1253, consistency index (CI) 0.773, homoplasy index (HI) 0.227, retention index (RI) 0.746 and rescaled consistency index (RC) 0.576.
RubisCO genes from a manganese-oxidizing bacterium

Table 2. Nucleotide sequence identity of RubisCO genes from strain S185-9A1 with those from other organisms

<table>
<thead>
<tr>
<th>Organism*</th>
<th>RubisCO group</th>
<th>Identity (%)</th>
<th>cbbL</th>
<th>cbbS</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Rhodobacter sphaeroides</em></td>
<td>IC</td>
<td>77</td>
<td>72</td>
<td></td>
</tr>
<tr>
<td><em>Xanthobacter flavus</em></td>
<td>IC</td>
<td>76</td>
<td>65</td>
<td></td>
</tr>
<tr>
<td><em>Alcaligenes eutrophus</em></td>
<td>IC</td>
<td>75</td>
<td>66</td>
<td></td>
</tr>
<tr>
<td><em>Chromatium vinosum</em></td>
<td>IA</td>
<td>63</td>
<td>43</td>
<td></td>
</tr>
<tr>
<td><em>Periphyllum aerugineum cp</em></td>
<td>ID</td>
<td>61</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td><em>Cryptomonas sp.</em></td>
<td>ID</td>
<td>61</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td><em>Anaepithrix nidulans</em></td>
<td>IB</td>
<td>60</td>
<td>48</td>
<td></td>
</tr>
<tr>
<td><em>Olisthodiscus luteus cp</em></td>
<td>ID</td>
<td>59</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td><em>Cyanidium caldarium cp</em></td>
<td>ID</td>
<td>59</td>
<td>49</td>
<td></td>
</tr>
<tr>
<td><em>Ectocarpus siliculosus cp</em></td>
<td>ID</td>
<td>59</td>
<td>50</td>
<td></td>
</tr>
<tr>
<td><em>Thiocapsa roseopersicina</em></td>
<td>IA</td>
<td>59</td>
<td>39</td>
<td></td>
</tr>
<tr>
<td><em>Synechococcus sp.</em> strain a-1</td>
<td>IB</td>
<td>59</td>
<td>44</td>
<td></td>
</tr>
<tr>
<td><em>Cyanobium sp.</em> strain N1 cp *</td>
<td>ID</td>
<td>56</td>
<td>52</td>
<td></td>
</tr>
<tr>
<td><em>Chlamydomonas reinhardtii cp</em></td>
<td>IB</td>
<td>50</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>Zea mays cp</td>
<td>IB</td>
<td>50</td>
<td>30</td>
<td></td>
</tr>
<tr>
<td><em>Rhodospirillum rubrum (cbhM)</em></td>
<td>II</td>
<td>38</td>
<td>39</td>
<td>NA</td>
</tr>
</tbody>
</table>

na, Not applicable.

* cp, chloroplast.

Clone pRC10, which contains the cbbL/S genes in an opposite orientation from the lac promoter of pBluescript, did not show any RuBP-dependent CO₂-fixing activity, while clone pRC11, in which the cbbL/S genes are controlled by the lac promoter, demonstrated IPTG- and RuBP-dependent CO₂ fixation at a rate of 18 nmol CO₂ min⁻¹ (mg total soluble protein)⁻¹ (Table 3). The positive control cells (*T. neapolitana*) fixed CO₂ in the presence of RuBP at a rate of 51.3 nmol CO₂ min⁻¹ (mg total soluble protein)⁻¹. Strain S185-9A1 showed no RubisCO activity.

**DISCUSSION**

The RubisCO genes of strain S185-9A1 are unique, and are not closely related to any other published RubisCO sequences. The most parsimonious tree (Fig. 5) shows that the strain S185-9A1 genes are about equally closely related to the members of Type IC and Type ID groups. Although the grouping of S185-9A1 RubisCO with non-chlorophyte algal chloroplast RubisCOs (Type ID) was the most parsimonious with this data set, it was not strongly supported by bootstrapping (the alternative grouped S185-9A1 at the base of the Type IC group). Nonetheless, all of the analyses support the RubisCO genes of strain S185-9A1 being closest to Type ID RubisCO genes of non-chlorophyte algal chloroplasts among the bacterial RubisCO genes that have been sequenced to date. This result was obtained by analysing either the sequence of both genes together, or the sequence of cbbL alone.

The Type ID enzymes are the subject of intense current interest because they have a much higher specificity factor (lower susceptibility to the competing oxygenase reaction) than enzymes of Type IB from terrestrial plants (Tabita, 1995). The bacterial enzymes of Type IC have specificity factors in the range of terrestrial plants. The loop 6 region of the large subunit of RubisCO is believed to influence the specificity factor (Read & Tabita, 1994), and the S185-9A1 enzyme shares some amino acids with the Type ID enzymes, some with the Type IC and is unique at other positions. Clearly, the specificity factor and other biochemical properties of the S185-9A1 RubisCO are important subjects for future investigation.

The issue of autotrophy is particularly interesting in the case of strain S185-9A1 because of its ability to oxidize Mn. For many decades scientists have been debating whether Mn oxidation could provide the necessary energy for autotrophic growth (Ali & Stokes, 1971; Beijerinck, 1913; Ghiorse, 1984; Kepkay & Nealson, 1987; Nealson et al., 1988; van Veen, 1972). Although we have not yet been able to demonstrate autotrophic growth of strain S185-9A1, it is apparent that this strain has some of the required genes. This, and the fact that strain S185-9A1 oxidizes Mn only on organic-poor media, suggest that there might be a link between Mn oxidation and CO₂ fixation in this strain.

The sequence data, together with the expression of the genes in an E. coli host, suggest that strain S185-9A1 might have the capability for synthesizing an active RubisCO enzyme. It remains unclear, though, under what
<table>
<thead>
<tr>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
<th><strong>A. eutrophus</strong></th>
</tr>
</thead>
</table>

**Fig. 4.** Amino acid sequence comparison of the RubisCO large (a) and small (b) subunits from different organisms. Identical residues are shown in white lettering on a black background. Active site residues are labelled with an asterisk (Knight et al., 1990).
RubisCO genes from a manganese-oxidizing bacterium

**Type IA**

H$_2$S oxidizers

- C. vinsonum
- T. ferrooxidans

**Type IB**

Cyanobacteria

- *A. nidulans*
- *Synechococcus* sp.

Chlorophytes

- *Z. mays* cp
- *C. reinhardtii* cp

**Type IC**

H$_2$ oxidizers

- *X. flavus*
- *A. eutrophus*
- *R. sphaeroides*

Nonsulfur purple

- SI85-9A1

**Type ID**

Non-chlorophyte algal chloroplasts

- *C. eugametos* cp
- *P. aeruginosa* cp

- *C. f. thapsica* cp

- *Cylindrotheca* sp.

Fig. 5. Relationships among different RubisCO large subunits based on amino acid sequences determined by maximum parsimony (PAUP; Swofford, 1991). The arrow indicates position of root (determined with *R. rubrum* as outgroup). Nodes significantly supported by bootstrap analysis (> 80%) are labelled with the percentage of bootstrap replicates that supported the node. Branch lengths indicate the inferred number of character state changes on each branch. cp, chloroplast.

**Table 3. Expression of the RubisCO enzyme**

Results are given as nmol CO$_2$ fixed min$^{-1}$ (mg total soluble protein)$^{-1}$. *T. neapolitanus* was used as a positive control and *E. coli* XL1-Blue carrying the plasmid pBluescript with no insert was used as a negative control.

<table>
<thead>
<tr>
<th></th>
<th>− IPTG</th>
<th>+ IPTG</th>
</tr>
</thead>
<tbody>
<tr>
<td>− RuBP + RuBP</td>
<td></td>
<td></td>
</tr>
<tr>
<td>pBluescript (no insert; negative control)</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>pRC10 (no lac promoter)</td>
<td>0.9</td>
<td>1.2</td>
</tr>
<tr>
<td>pRC11 (lac promoter)</td>
<td>0.6</td>
<td>1.4</td>
</tr>
<tr>
<td><em>T. neapolitanus</em> (positive control)</td>
<td>1.9</td>
<td>51.3</td>
</tr>
<tr>
<td>SI85-9A1</td>
<td>1.2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>− RuBP + RuBP</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>+ IPTG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>− RuBP + RuBP</td>
<td>1.3</td>
<td>1.2</td>
</tr>
<tr>
<td>pBluescript (no insert; negative control)</td>
<td>0.6</td>
<td>0.6</td>
</tr>
<tr>
<td>pRC10 (no lac promoter)</td>
<td>0.9</td>
<td>1.2</td>
</tr>
<tr>
<td>pRC11 (lac promoter)</td>
<td>0.6</td>
<td>1.4</td>
</tr>
<tr>
<td><em>T. neapolitanus</em> (positive control)</td>
<td>1.9</td>
<td>51.3</td>
</tr>
<tr>
<td>SI85-9A1</td>
<td>1.2</td>
<td>0.4</td>
</tr>
</tbody>
</table>

NA, Not applicable.

circumstances the genes might be expressed. Clone pRC10, in which the genes are not controlled by the lac promoter, did not demonstrate any RubisCO-dependent CO$_2$-fixing activity. However, in most organisms the *rbcL*/*S* genes are part of a larger operon that contains other CO$_2$-fixation-related genes (*cbbR, cbbF, cbbP* and *cbbA*). Since there clearly is a *cbbA* gene upstream of the *cbbL* gene in strain SI85-9A1, it is possible that the promoter for this operon is located upstream from the cloned fragment, and is thus missing in clones pRC10 and pRC11. Alternatively, the promoters might be present, but not recognized by the *E. coli* transcription machinery. The fact that the Mn-oxidizing strain SI85-9A1 has genes coding for a functional RubisCO enzyme and other carbon fixation genes, suggests that SI85-9A1 is either autotrophic for some energy source, or a heterotrophic descendant of a previously autotrophic organism that lost some of the genes required for autotrophy. In the latter case, the fact that functional RubisCO genes are still present could be explained if the RubisCO genes were
linked to other essential genes. A similar case has been reported for freshwater *Beggiaota* strains (Nelson et al., 1989) that have apparently lost the ability to grow autotrophically, while maintaining the RubisCO genes. Another similar case is that of *Thiothrix intermedia* in which two sets of RubisCO genes are present, coding for both form I and form II enzymes. However, all attempts to demonstrate the presence of a form II RubisCO enzyme in *T. intermedia* failed (Stoner & Shively, 1993). After cloning the genes of *T. intermedia* into *E. coli*, a low level of activity was present, and as in our case, a higher level (32 nmol CO$_2$ fixed min$^{-1}$ (mg protein)$^{-1}$) was observed under the control of an external lac promoter.

On the other hand, strain SI85-9A1 may be the first representative of a novel group of organisms that fix CO$_2$ in the environment. A sequence similar to the strain SI85-9A1 RubisCO was recently obtained from a water column mRNA sample from the Gulf of Mexico at a depth of about 90 m (Paul & Pichard, 1996), representing a different ocean basin to that from which strain SI85-9A1 was isolated. The wide distribution and relative abundance of this type of gene implied by the finding, and the fact that the gene was apparently transcribed, underscore the importance of further studies on strain SI85-9A1 and other representatives of this group.

**ACKNOWLEDGEMENTS**

We would like to thank Dr. D. Bartlett who kindly supplied cosmid PMM83, D. Edwards and L. Park for the 16S rRNA sequence analysis, and J. H. Paul and S. L. Pichard for allowing us to cite their work prior to publication. This work was supported by an Oceanic Biology grant from the Office of Naval Research (N00014-90-J-1097). Initial studies were also supported by a grant from the National Science Foundation (OCE-86-2089). R. Caspi was in part supported by a fellowship from the University of California Toxic Substances Research and Teaching Program.

**REFERENCES**


