Cloning and Sequencing of a Form II Ribulose-1,5-Bisphosphate Carboxylase/Oxygenase from the Bacterial Symbiont of the Hydrothermal Vent Tubeworm *Riftia pachyptila*

JONATHAN J. ROBINSON, 1 JEFFREY L. STEIN, 2 AND COLLEEN M. CAVANAUGH 1*

Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, 1 and Diversa Corp., San Diego, California 92121 2

Received 4 August 1997/Accepted 13 January 1998

The bacterial symbiont of the hydrothermal vent tubeworm fixes carbon via the Calvin-Benson cycle and has been shown previously to express a form II ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO). The gene cbbM, which encodes this enzyme, has been cloned and sequenced. The gene has the highest identity with the cbbM gene from *Rhodospirillum rubrum*, and analysis of the inferred amino acid sequence reveals that all active-site residues are conserved. This is the first form II RubisCO cloned and sequenced from a chemosynthetotrophic symbiont and from a deep-sea organism.

Hydrothermal vent environments are dominated by dense assemblages of invertebrates which harbor chemosynthetic sulfur-oxidizing bacteria within their tissues. This nutritional interaction between prokaryotic symbionts and various animal hosts is dependent upon the biological fixation of inorganic carbon by the symbionts and the subsequent supply of organic carbon to the host in a manner analogous to the chloroplasts of green plants and algae (reviewed in reference 2). Fundamental to the initial and subsequent characterization of these symbioses has been the detection of the key Calvin-Benson cycle enzyme, ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) (1, 8). The hydrothermal vent tubeworm *Riftia pachyptila* is of particular interest with regard to its carbon fixation abilities, as this animal completely lacks a mouth, gut, or anus (13) but is capable of extreme size and high growth rates due to its symbiotic association (16).

The primary carbon fixation step in the Calvin-Benson cycle is catalyzed by RubisCO, which carboxylates ribulose-1,5-bisphosphate with CO2 to yield two molecules of 3-phosphoglyceric acid. The enzyme is found in two forms, called form I and form II (30), which are distinct in primary and quaternary structure (26), reaction mechanism, and kinetic isotope effect (KIE) (22, 23). The form I RubisCO found in the vast majority of eukaryotic and prokaryotic autotrophs, consists of eight large subunits and eight small subunits, with the holoenzyme configuration, as reported for other species (30). The two forms are 25% identical to each other at the amino acid level (18). To date, the form II enzyme has only been characterized at the nucleic acid sequence level from five prokaryotes (31).

The bacterial symbiont of the vestimentiferan *R. pachyptila* has been shown to express a form II RubisCO (21). In this work we report the cloning and sequencing of the cbbM gene, which encodes a form II enzyme, from the *R. pachyptila* symbiont.

**Bacterial strains, plasmids, and polyclonal antisera.** The *Escherichia coli* construct pRR2119 (ATCC 37846) was used to generate probes for hybridization during library screening. This clone harbors the plasmid pXG9 containing the cloned form II RubisCO from *Rhodospirillum rubrum* (28). *E. coli* INFγF (Invitrogen) was used for cloning steps and grown in Luria broth supplemented with ampicillin (40 mg liter−1). Plasmid pCRII (Invitrogen) was used for subcloning, DNA sequence analysis, and protein expression studies.

The lambda DNA library (see below) was screened with polyclonal antiserum directed against the *R. rubrum* form II RubisCO (anti-RfII) (antiserum generously provided by George Lorimer [DuPont]), which has been shown to be specific to form II RubisCOs and to cross-react with the *R. pachyptila* enzyme (21). In all cases antiserum was used at a 1:3,000 dilution.

**R. pachyptila genomic DNA library construction.** Tubeworm specimens used for genomic DNA library construction were collected from a depth of 2,600 m using the DSV *Alvin* from hydrothermal vents on the East Pacific Rise at the 13°N site (12°48′N, 103°56′W; November 1987). The worms were transported to the surface in a thermally insulated container and the symbiont-containing trophosome tissue was immediately dissected on board ship. Tissue was homogenized in a 1:1 (wt/vol) solution of ice-cold *Riftia* saline (46 mM imidazole, 0.46 M NaCl, 30 mM MgSO4, 2.5 mM KCl, 10 mM CaCl2, pH 7.1) at 30 to 40% speed in an Ultraturrax homogenizer for 2 min on ice. Symbionts in this solution were separated from host cells on 80% Percoll density gradients according to the method of Distel and Felbeck (6) with modifications.

DNA was extracted from the symbiont preparation by using a 5 M guanidinium isothiocyanate solution (15). DNA (75 μg) was sheared to an average size of 3 to 6 kbp by vigorous passage through a 25-gauge needle in a 1-ml syringe. The sheared DNA was blunt ended with mung bean nuclease and ligated to EcoRI linkers, and 3- to 6-kbp fragments were cloned into lambda gt11 (27). The library titer was estimated to be 1.5 × 1010 PFU (25).

**Library screening.** Phage were plated and screened by standard methods on a lawn of *E. coli* Y1090 (25). Plaques were screened for the expression of the form II RubisCO by incu-
bation with anti-RrFII antiserum (25). Plaques which were immunologically positive were rescreened with a 32P-labelled BglII/SmaI fragment of the R. rubrum form II RubisCO derived from plasmid pXG9 (28).

Two lambda clones were isolated. Inserts were amplified from purified lambda DNA by PCR with primers specific to the lacZ cloning region (Promega) and subcloned into the pCR II vector for transformation into E. coli host strain INF aF9 and subsequent DNA sequencing. The two clones are different sizes, with insert sizes of 3,300 bp for pRpR-1 and 2,200 bp for pRpR-2, and are oriented in opposite directions.

DNA sequence analysis. Oligonucleotide primer walking was used to generate a double-stranded sequence for the region encoding the form II RubisCO and immediate flanking regions for both clones. Sequencing was conducted with the Applied Biosystems Inc. (ABI) Dye Terminator Cycle Sequencing kit under standard conditions, an Ericomp thermal cycler, and an ABI model 373 automated sequencer. Sequences of regions flanking the EcoRI cloning sites were also determined, using the M13 universal primers designated reverse and forward.

Sequencing of a 1,678-bp region from both clones revealed open reading frames with high identity to previously sequenced cbbM genes (Fig. 1 and 2). cbbM is preceded by an in-frame TAG stop codon at position 29 and then begins with an ATG and proceeds 1,383 bp to a TAA stop codon, followed by a putative hairpin loop beginning 27 bp downstream (Fig. 1). The cbbM coding region is composed of 57.5 mol% G+C, and a 461-amino-acid protein with a calculated molecular weight of 50,552 Da is predicted. Efforts to express active recombinant form II RubisCO failed to yield enzyme with significant activity, suggesting that the recombinant does not fold properly in E. coli or is posttranslationally modified by the bacterial symbiont. Therefore, biochemical characterization of this RubisCO is currently being conducted on native enzyme.

Analysis of sequence flanking the EcoRI cloning sites revealed the presence of an open reading frame sharing identity to the LysR type regulator cbbR (not shown). This gene is upstream of cbbM and in the opposite orientation. The deduced amino acid sequence of the cbbR element has 61% identity with the cbbR of Chromatium vinosum (31) over the region sequenced, which spans 71 residues at the 5' end.

Translation of the open reading frame and alignment with other form II enzymes and a representative form I RubisCO (Fig. 2) revealed strict conservation of residues known to form the enzyme active site (11, 26), e.g., the specific lysine residue which is carbamylated during enzyme activation and corresponds to position 191 of the R. rubrum sequence. N-terminal sequence analysis indicates that the first-position methionine residue is posttranslationally cleaved (3), a situation encountered in plant RubisCO enzymes (12). Amino acid identity with other form II RubisCOs ranges from a high of 76.2% with R. rubrum to a low of 69.1% with the dinoflagellate Gonyaulax polyedra. With regard to amino acid similarity, i.e., by comparison of major amino acid biochemical groupings, the R. pachyptila enzyme is most similar (89.2%) to the Rhodobacter sphaeroides form II enzyme and shows 78 to 89% similarity with all the other form II enzymes. In contrast, the R. pachyptila RubisCO shows only 22 to 32% amino acid identity with the gene encoding the large subunit of representative form I RubisCOs, including that of Spinacia oleracea.

The discovery of a form II RubisCO in a deep-sea organism indicates that this enzyme is found in diverse settings and is not as rare as once thought. Indeed, six other deep-sea symbionts and two bacterial mats have recently been shown to express this form of RubisCO (3, 20, 21). While both forms of RubisCO are expressed in some free-living bacteria (30, 31), the R. pachyptila symbiont appears to encode and express only a form II enzyme. In the R. pachyptila symbiosis, hybridization to a form I heterologous gene probe was not detected during library screening or Southern analysis of trophosome DNA, in contrast to earlier reports (29, 32), nor was a form I gene product detected (21). Indeed, other researchers have also failed to detect the gene encoding the form I enzyme in the R. pachyptila symbiont, detecting only the cbbM gene (14). Physiologically, the use of a form II RubisCO in this symbiosis is...
not surprising, given that form II enzymes typically have a low affinity for CO₂ and that concentrations of CO₂ are extremely high in the blood of the tubeworm, where concentrations of total dissolved inorganic carbon can be greater than 30 mM (5).

The expression of form I and II Rubisco has recently been suggested to account for the difference observed in stable carbon isotope ratios (d¹³C) of hydrothermal vent invertebrate-chemoautotrophic bacterial symbioses (3, 21). These symbioses fall into two groups based upon their d¹³C values, with d¹³C = -27 to -35% for mollusc symbioses and d¹³C = -9 to 16% for tubeworm and shrimp symbioses (references 21 and references within). Several hypotheses, such as carbon limitation (9, 19), a C₄-type pathway in the tubeworms (7), or the use of isotopically different source CO₂ utilized by the two groups (4), have been proposed to explain the differences in d¹³C values but have failed to be corroborated by experimental data. The KIEs of the few form I (for S. oleracea, 29%; for Anacystis nidulans, 22% [10, 22]) and form II (for R. rubrum, 17.8 to 23% [10, 23]) Rubisco examined by high-precision methods indicate that the two forms fractionate carbon isotopes to differing degrees. Given the high identity between the R. rubrum and R. pachyptila cbbM sequences, the expression of a form II Rubisco in R. pachyptila could account for the heavier isotopic composition if the extreme values for the KIE of the form II Rubisco are considered.

Chemoautotrophic symbioses and free-living chemoautotrophs represent a vast resource for examining different adaptations that have occurred in Rubisco biochemistry and evolution. These organisms promise to yield important new information regarding enzymological adaptation, regulation, and genetic diversity, as they inhabit many niches which are too inhospitable for photoautotrophs. The examination of a greater diversity of species for the form II Rubisco is necessary to determine the distribution of this enzyme among autotrophs.

Nucleotide sequence accession number. The R. pachyptila symbiont cbbM gene sequence has been deposited in GenBank under accession no. AF047688.

We thank the chief scientists and captains and crew of the RV Atlantis II and DSV Alvin for their excellent assistance in sample collection, George Lorimer for the generous gift of antiserum, and Marjory Snead for the R. pachyptila DNA library construction. This work was supported in part by NSF grants OCE-9317734 (J.L.S.) and OCE-9504257 (C.M.C.). J.J.R. was also supported by a Graduate Assistance in Areas of National Need fellowship from the Department of Education.

REFERENCES


