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

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters
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 921212

Received 4 August 1997/Accepted 13 January 1998

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 having a molecular weight of between 500 and 560 kDa (26). The form II enzyme (21). In all cases antiserum was used at a 1:3,000 dilution.

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 chemosynthetic 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 having a molecular weight of between 500 and 560 kDa (26). The form II enzyme is structurally less complicated, consisting of a having a molecular weight between 500 and 560 kDa (26). The large subunits and eight small subunits, with the holoenzyme of eukaryotic and prokaryotic autotrophs, consists of eight...
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 −9 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 (δ¹³C) of hydrothermal vent invertebrate-chemoautotrophic bacterial symbioses (3, 21). These symbioses fall into two groups based upon their δ¹³C values, with δ¹³C = -27 to -35% for mollusc symbioses and δ¹³C = -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 δ¹³C 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]) RubisCOs 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 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

FIG. 2. Alignment of R. pachyptila RubisCO with other form II enzymes and a representative form I from S. oleracea (shown for comparison). Asterisks refer to residues known to be involved in the active site and activating CO₂ binding site. Dots indicate residues identical to those in the R. pachyptila sequence, and dashes are added to preserve the alignment. The alignment of form I and II RubisCOs is based on the three-dimensional structure (26).