Structure, Vol. 13, 779–789, May, 2005, ©2005 Elsevier Ltd All rights reserved. DOI 10.1016/j.str.2005.02.017

Crystal Structure of a RuBisCO-like Protein from the Green Sulfur Bacterium *Chlorobium tepidum*

Huiying Li,[1] Michael R. Sawaya,[1] F. Robert Tabita,[2] and David Eisenberg[1,*]
[1]Department of Chemistry and Biochemistry
Howard Hughes Medical Institute
UCLA-DOE Institute for Genomics and Proteomics
University of California, Los Angeles
Box 951570
Los Angeles, California 90095
[2]Department of Microbiology and
Plant Molecular Biology/Biotechnology Program
The Ohio State University
484 West 12th Avenue
Columbus, Ohio 43210

Summary

Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO) catalyzes the incorporation of atmospheric CO_2 into ribulose 1,5-bisphosphate (RuBP). RuBisCOs are classified into four forms based on sequence similarity: forms I, II and III are bona fide RuBisCOs; form IV, also called the RuBisCO-like protein (RLP), lacks several of the substrate binding and catalytic residues and does not catalyze RuBP-dependent CO_2 fixation in vitro. To contribute to understanding the function of RLPs, we determined the crystal structure of the RLP from *Chlorobium tepidum*. The overall structure of the RLP is similar to the structures of the three other forms of RuBisCO; however, the active site is distinct from those of bona fide RuBisCOs and suggests that the RLP is possibly capable of catalyzing enolization but not carboxylation. Bioinformatic analysis of the protein functional linkages suggests that this RLP coevolved with enzymes of the bacteriochlorophyll biosynthesis pathway and may be involved in processes related to photosynthesis.

Introduction

Ribulose 1,5-bisphosphate carboxylase/oxygenase (RuBisCO, EC 4.1.1.39) is the most abundant enzyme found on earth (Ellis, 1979). It catalyzes the addition of CO_2 to ribulose 1,5-bisphophate (RuBP) in the photosynthesis pathway. The resultant 6-carbon intermediate is then cleaved to generate two molecules of 3-phosphoglycerate (3-PGA). Carboxylation is actually the sum total of five steps: enolization, carboxylation, hydration, C-C cleavage, and protonation (Cleland et al., 1998; Schneider et al., 1992). In the first step, RuBP is enolized to the 2,3-enediol form. In the second step, one molecule of CO_2 reacts with the C-2 carbon atom of the enediol to form a six carbon intermediate, 3-keto-2-carboxyarabinitol 1,5-bisphosphate (3-keto-CABP). 3-keto-CABP is then hydrated in the third step, and this derivative exists mainly in the gem diol form. A proton is abstracted from the gem diol and the bond between C-2 and C-3 is cleaved in the fourth step, resulting in the formation of one molecule of 3-PGA and one molecule of the C-2 carbanion of 3-PGA. Finally, the C-2 carbanion is quickly protonated into another molecule of 3-PGA in the fifth step.

Of the four forms of RuBisCO (Tabita, 1999), form I is the most abundant class of RuBisCO. It is found in plants, algae, and bacteria, and is composed of eight large subunits and eight small subunits with 422 symmetry (L_8S_8) (Baker et al., 1977). Many form I RuBisCO structures from different organisms have been determined and show high similarity (Andersson, 1996; Andersson and Taylor, 2003; Chapman et al., 1987; Curmi et al., 1992; Hansen et al., 1999; Knight et al., 1990; Schreuder et al., 1993b; Shibata et al., 1996; Spreitzer and Salvucci, 2002; Sugawara et al., 1999; Taylor and Andersson, 1996; Taylor and Andersson, 1997a; Taylor and Andersson, 1997b; Taylor et al., 1996; Taylor et al., 2001; Zhang et al., 1994). Form II RuBisCO is found primarily in certain bacteria and is composed solely of large subunits that differ substantially in sequence from form I large subunits. Depending on the source, form II RuBisCO may be oligomerized to form dimers, tetramers, or even larger oligomers. The crystal structure of the form II RuBisCO from *Rhodospirillum rubrum* (L_2) has been determined (Lundqvist and Schneider, 1989a; Lundqvist and Schneider, 1989b; Lundqvist and Schneider, 1991a; Lundqvist and Schneider, 1991b; Schneider et al., 1990; Soderlind et al., 1992) and reveals high similarity to the large subunit structure of form I RuBisCO. Form III RuBisCO is found only in archaea, and has been shown to form either dimers (L_2) (Finn and Tabita, 2003; Watson et al., 1999) or decamers ($[L_2]_5$) (Maeda et al., 1999), depending on the organism. The crystal structure of the form III RuBisCO from *Thermococcus kodakaraensis* reveals that the protein is comprised of a pentamer of dimers (Kitano et al., 2001). Its dimeric interface is very similar to those observed in the large subunit of form I and form II RuBisCO. Form IV, also called the RuBisCO-like protein (RLP), was recently discovered to be a homolog of bona fide RuBisCO (Hanson and Tabita, 2001). This report offers the structure and hints of the function of the RLP from *Chlorobium tepidum*.

RLPs are found in physiologically and systematically diverse prokaryotes, including *Bacillus subtilis* (Kunst et al., 1997), *C. tepidum* (Eisen et al., 2002; Hanson and Tabita, 2001), *Rhodopseudomonas palustris* (Larimer et al., 2004), and the extremophilic archaeon *Archaeoglobus fulgidus* (Klenk et al., 1997), among others. Although these RLPs are related at the sequence level, it appears that different sources of RLP possess discrete functions (Ashida et al., 2003; Hanson and Tabita, 2003). *B. subtilis* RLP functions as a 2,3-diketomethylthiopentyl-1-phosphate (2,3-DK-MTP-1-P) enolase in the methionine salvage pathway of this organism (Ashida et al., 2003; Murphy et al., 2002; Sekowska and Danchin, 2002). In the *C. tepidum* genome, however, there are no recognizable genes for a methionine salvage pathway. *C. tepidum* RLP is shown to be involved

*Correspondence: david@mbi.ucla.edu

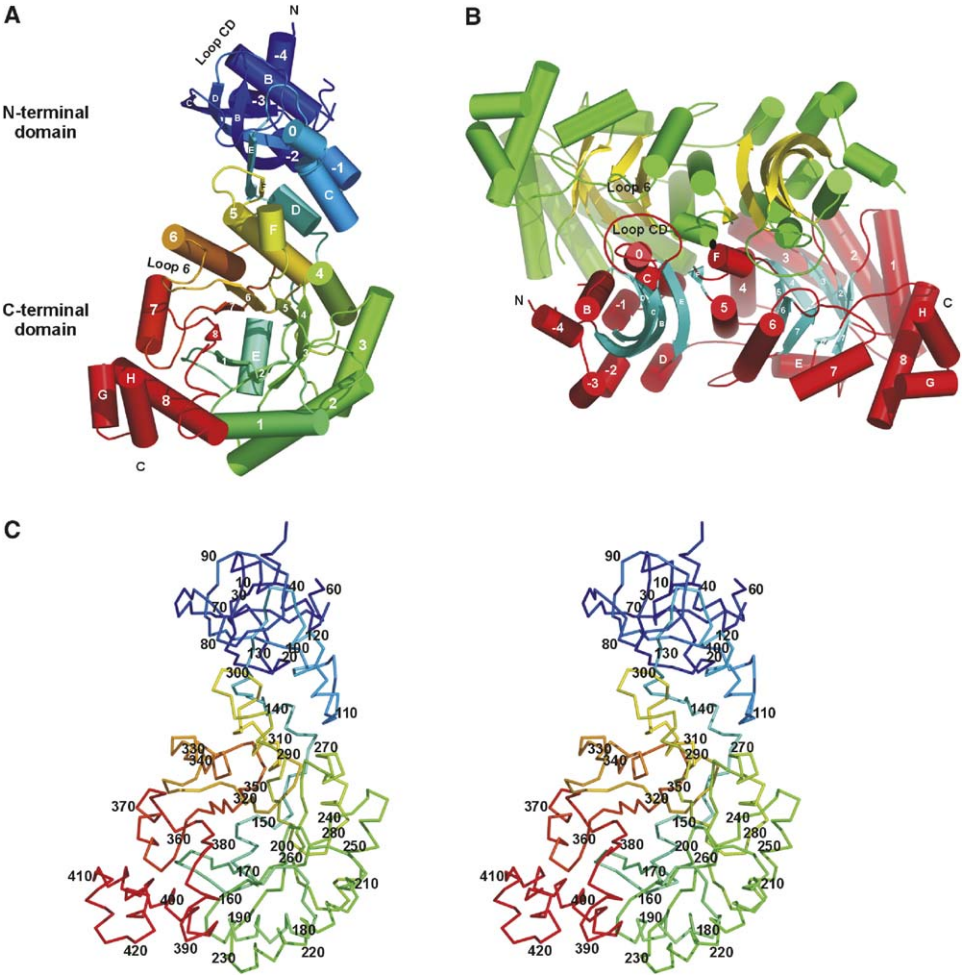

Figure 1. Monomer and Dimer Structure of *C. tepidum* RLP

(A) The monomer structure of the RLP. The monomer can be divided into two domains: a smaller N-terminal β sheet domain and a C-terminal α/β barrel domain. The structure is color-ramped from dark blue (N terminus) to red (C terminus) and the secondary structure is labeled. Loop CD and loop 6 are also indicated.

(B) The dimer structure of the RLP. The asymmetric unit contains a dimer. The active sites are formed between the N-terminal domain from one subunit and the C-terminal domain from the other subunit. The two monomers are colored in green and red. The black ellipse indicates the two-fold noncrystallographic symmetry axis.

(C) Stereoview of the Cα trace of the RLP monomer structure.

with sulfur oxidation, and inactivation of the *rlp* gene leads to an oxidative stress response (Hanson and Tabita, 2001, 2003). Currently it is unclear which reaction is affected in sulfur metabolism, but RLP appears to be somehow involved with the oxidation of thiosulfate, but not sulfide. Interestingly, the *R. palustris* genome encodes two RLPs as well as both form I and form II RuBisCO (Larimer et al., 2004). Sequence alignment analysis of RLPs from different organisms suggests that the two RLPs from *R. palustris* belong to different groups, perhaps indicating that these two RLPs may have different functions (Hanson and Tabita, 2003).

As there are more than two hundred complete genome sequences available, bioinformatic approaches can assist us in understanding potential functions of uncharacterized proteins. Several genome-context-based methods have been developed to infer protein functions based on comparisons of tens and hundreds of genome sequences: the Phylogenetic Profile method infers protein functional linkages between two proteins based on their correlated evolution in multiple genomes (Pellegrini et al., 1999); the Rosetta Stone method infers the linkages based on the fusion of two protein-encoded genes in another genome (Enright et al., 1999; Marcotte et al., 1999); the Gene Neighbor method assigns protein functional linkages based on the close proximity of two genes on the chromosomes in many genomes (Dandekar et al., 1998; Overbeek et al., 1999); and the Gene Cluster method infers the linkages between two genes based on the operon structure (Bowers et al., 2004; Pellegrini et al., 2001).

In this study the structure of the RLP from *C. tepidum* was determined using X-ray crystallography and the active site was analyzed and compared to other forms

Table 1. X-Ray Data Collection and Refinement Statistics of the RLP Structure

Data Collection	Native
Wavelength (Å)	1.127
Temperature (K)	100
Space group	$P2_1$
Cell parameters	
a (Å)	67.35
b (Å)	78.45
c (Å)	90.37
β (°)	99.95
Resolution (Å)[a]	87.7–2.0 (2.07–2.00)
Reflections	
Total	186,455
Unique	60,951
Completeness[a] (%)	97.5 (84.7)
$\langle I/\sigma \rangle$[a]	7.7 (2.0)
R_{sym}[a,b] (%)	12.9 (39.7)
Model Refinement	
R (%)/R_{free}[a] (%)	20.1/24.5 (25/31)
Number of atoms	
Protein	6419
Solvent	317
Rms bond length (Å)	0.013
Rms bond angles (°)	1.626
Ramachandran plot	
Most favored	641 residues, 92.2%
Additional allowed	48 residues, 6.9%
Generously allowed	3 residues, 0.4%
Disallowed	3 residues, 0.4%

[a] Statistics for the outer resolution shell are given in parentheses.
[b] $R_{sym} = \Sigma(I - \langle I \rangle)^2/\Sigma I^2$

of RuBisCO. The RLPs from *C. tepidum*, *R. palustris*, and *B. subtilis* were identified by sequence similarity to RubisCO and their functional linkages were calculated by comparing them with the genome sequences of 168 organisms in the Prolinks database (Bowers et al., 2004).

Results

Overall Structure of the RLP from *C. tepidum*

The monomer structure of the RLP from *C. tepidum* is similar to those of other RuBisCOs and may be divided into two domains (Figures 1A and 1C). The smaller N-terminal domain (residues 1–145) consists of a four-stranded β sheet with helices on one side of the sheet. The larger C-terminal domain (residues 146–435) consists of an eight-stranded α/β barrel with two additional small α helices forming a cap at the C terminus. The monomer of the RLP is composed of 435 residues, and the model contains most of the residues except N-terminal residues 1–3, C-terminal residues 429–435 and two disordered regions on the surface (residues 47–58 for chains A and B and residues 173–174 for chain B only) (Table 1). A RuBisCO active site residue, Q49, is within one of the disordered regions.

The RLP from *C. tepidum* forms homodimers (Figure 1B). In the asymmetric unit, two monomers form a closely packed dimer with two-fold symmetry. It is similar to the dimeric functional units of other forms of RuBisCO. When the dimer structures of the RLP and the RuBisCO from *R. rubrum* (5RUB, form II) are superimposed, the root mean square deviation (rmsd) of Cα atoms is 1.8 Å with 570 residues aligned, showing that they share a similar structure even though their amino acid sequence identity is low (30%).

Active Site of *C. tepidum* RLP

Like the other three forms of RuBisCO (Kitano et al., 2001; Schreuder et al., 1993a), the active site of the RLP is formed between two subunits, the C-terminal α/β barrel domain from one subunit and the N-terminal β sheet domain from another subunit. However, 10 of the 19 mechanistically significant residues of the RuBisCO active site differ in the RLP of *C. tepidum* (Figure 2). These replacements in the amino acid sequence alter the shape and chemical properties of the active site, making it evident that this RLP may not bind RuBP, but may perhaps bind a structurally related molecule. The differences between the active sites of the RLP and other RuBisCOs may be illustrated by visualizing the ability of the RLP to interact with the 6-carbon transition state analog, 2-carboxyarabinitol 1,5-bisphosphate (CABP), of RuBisCOs. From studies with form I RuBisCO, the residues involved in CABP binding can be divided into four groups: those involved in forming hydrogen bonds with (1) P1 phosphate, (2) P2 phosphate, and (3) the carboxyarabinitol backbone; and those involved in (4) coordinating the metal ion (Knight et al., 1990).

The RLP active site appears to be compatible with binding the P1 phosphate group. Some of the active site residues involved in forming hydrogen bonds are structurally conserved in the RLP compared to form I RuBisCO. The backbones of residues G382 and R383 can still make hydrogen bonds with the P1 phosphate group, as do the residues G403 and G404 in 8RUC (the structure of activated spinach RuBisCO [form I]) (Figure 3). The backbone of residue R327 in loop 6 superimposes well with that of K334 in 8RUC, and its side chain is quite close, so that R327 can possibly make hydrogen bonds with atoms O3P and O7 of CABP. Residue G381 in 8RUC is changed to S359 in the RLP, but S359 is capable of forming a hydrogen bond with atom O3P of CABP. Residue Q49 of the RLP is structurally equivalent to T65 of 8RUC and, in principle, is capable of donating a hydrogen bond to the P1 phosphate, although it is disordered in the RLP structure.

The largest change in the active site is at the P2 phosphate binding site. When the active sites of spinach RuBisCO (8RUC, form I) and *C. tepidum* RLP are superimposed, the P2 phosphate group of CABP does not fit into the active site of the RLP. In 8RUC, residue R295 donates two hydrogen bonds to the P2 phosphate group of CABP (Andersson, 1996), but in the RLP, the arginine is nonconservatively substituted by a phenylalanine (F288). Thus, residue F288's side chain would block the P2 phosphate group of RuBP from binding to the RLP (Figure 3). Similarly, in 8RUC, residue H327 donates a hydrogen bond to this same P2 phosphate group, but in the RLP the histidine is replaced by an isoleucine (I320). These two amino acid substitutions in the RLP contribute to the increase of the hydrophobicity of the substrate binding pocket in the RLP compared to forms I, II, and III RuBisCO.

The RLP active site appears compatible with binding

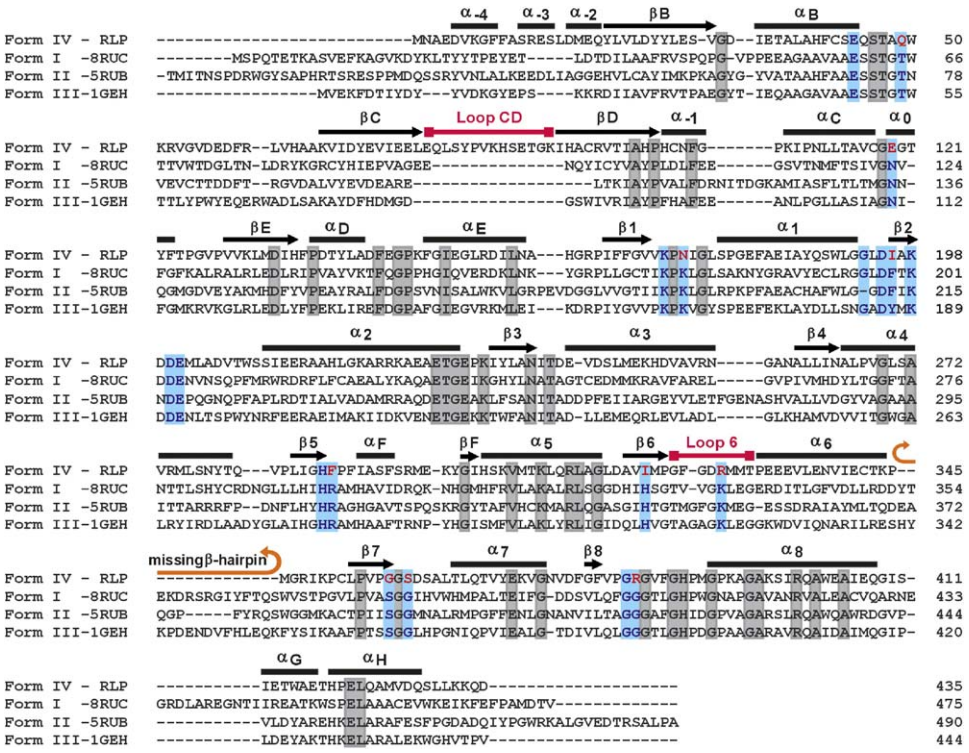

Figure 2. Structure-Based Sequence Alignment of Four Forms of RuBisCO

Sequences of the RLP from *C. tepidum* (form IV), spinach RuBisCO large subunit (8RUC, form I), RuBisCO from *R. rubrum* (5RUB, form II), and RuBisCO from *T. kodakaraensis* (1GEH, form III) were used in the alignment. Secondary structural elements (helices as bars and β strands as arrows) of the RLP are shown at the top of the sequences. The residue numbers of each sequence are shown on the right. Identical nonactive site residues in all four forms of RuBisCO are boxed in gray. The residues involved in substrate binding and catalytic activity are highlighted in blue boxes. In the RLP, ten (colored in red) out of 19 known mechanistically significant RuBisCO active site residues (colored in blue) are replaced. Loop 6, loop CD, and the missing β-hairpin in the RLP are also indicated.

the CABP backbone. The residues involved in making hydrogen bonds with the carboxyarabinitol moiety are partially conserved (Figure 3). Residue K172 in *C. tepidum* RLP, the equivalent to residue K175 in 8RUC, can still form a hydrogen bond with O2. Residue T173 in 8RUC, which contributes another hydrogen bond to O2, is replaced with V170 in the RLP. However, this residue is not conserved in forms II and III RuBisCO either, indicating that T173 probably does not play a key role in binding the substrate. Residue H294, which forms a hydrogen bond with O3, is conserved in the RLP (residue H287). Residue K177 forms a hydrogen bond with O6 in 8RUC, but is replaced by N174, which is partially disordered in the RLP structure. Residue N123 in 8RUC also contributes one hydrogen bond to O6. It is replaced by E119 in the RLP, but this residue can still possibly make a hydrogen bond with O6 of CABP. Residue S379's side chain in 8RUC contributes a hydrogen bond to O4, but this residue is changed to G357 in the RLP.

As with other RuBisCO proteins, the RLP active site is compatible with binding a metal ion. The residues involved in metal ion binding are structurally conserved in the RLP, including residues K198, D200, and E201. During RuBisCO catalysis, carbamylation at a particular lysine residue, K201 in 8RUC, must occur in order for the enzyme to be activated prior to carboxylation or oxygenation. The ϵ-amino group of the lysine reacts with a CO_2 molecule to form a carbamate, stabilized by the binding of Mg^{2+} or other divalent cations. The carbamate removes the C-3 proton and transfers it to residue K175 (Cleland et al., 1998; Spreitzer and Salvucci, 2002). Residue K201 in the activated spinach RuBisCO is carbamylated (Andersson, 1996). In our RLP crystal, the equivalent lysine residue, K198, was not carbamylated, as indicated by the electron density around the residue. There is also no electron density found for a metal ion near residue K198, indicating that the enzyme was not carbamylated, although the protein was incubated under conditions where bona fide RuBisCOs become activated (i.e., incubation with $NaHCO_3$ and Mg^{2+} in Tris·HCl [pH 8.0] prior to crystallization). This is consistent with the inability of the RLP to form an exchange-inert ternary complex with Mg^{2+} and CO_2 observed in a previous study (Hanson and Tabita, 2001).

Loop 6

Loop 6 in the C-terminal α/β barrel domain has been observed to have two different conformations in form I RuBisCO structures (Duff et al., 2000; Schneider et al., 1992; Schreuder et al., 1993a). In the structure of unac-

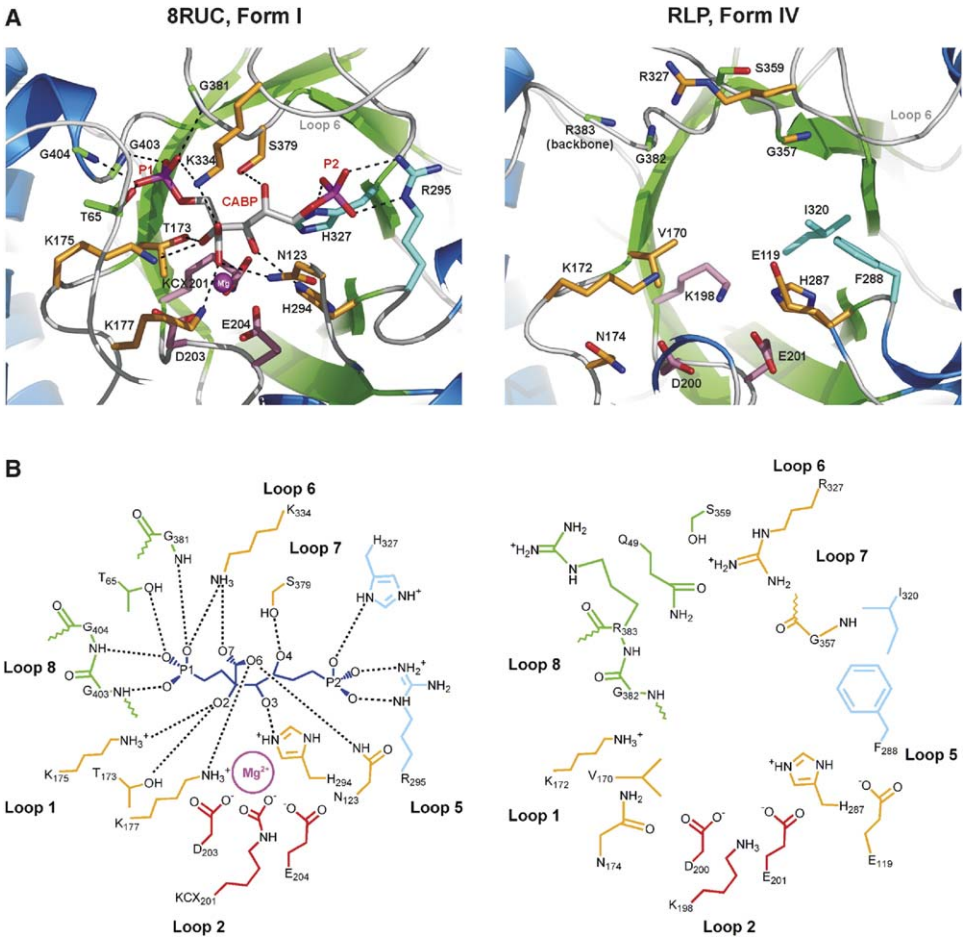

Figure 3. The Active Sites of Spinach RuBisCO (8RUC, Form I) with CABP Bound and *C. tepidum* RLP (Form IV)

(A) Cartoon representation of the active site residues and the hydrogen bonds formed between the 8RUC residues and CABP. The side chains of active site residues are shown in sticks, except for residue R383 in the RLP. Only the backbone carbon and nitrogen atoms of R383 in the RLP are shown. CABP is shown in white and the P1 and P2 phosphate groups of CABP are labeled in red and magenta. The residues involved in contributing hydrogen bonds with the P1 phosphate group are colored in green, the residues involved in making hydrogen bonds with the backbone of CABP are colored in orange, the residues coordinating the Mg^{2+} atom (shown in magenta) are colored in light red, and the residues involved in binding P2 phosphate group are colored in cyan. Not all parts of the structures are shown for the purpose of clarity.

(B) Schematic representation of (A) showing the hydrogen bonds between the active site residues and CABP. The color scheme is the same as in (A), except that CABP is shown in dark blue. Residue Q49 is disordered in the RLP structure, but is drawn in the schematic for the purpose of comparison.

tivated tobacco RuBisCO (1EJ7) (Duff et al., 2000), loop 6 is in an open conformation. In the structure of activated spinach RuBisCO (8RUC), loop 6 is in a closed conformation (Andersson, 1996). Upon RuBP or transition state analog binding, loop 6 folds over the active site and shields the substance from the solvent. The movement of loop 6 brings residue K334 close to the substrate and to donate a hydrogen bond to the 2-carboxy group of RuBP or CABP (Schneider et al., 1992). K334 is a conserved catalytic residue in bona fide RuBisCOs. It also forms hydrogen bonds with the P1 phosphate group and E60 in the N-terminal domain of the other subunit of the dimer. The available structure of *R. rubrum* RuBisCO (5RUB) is of unactivated enzyme and loop 6 is disordered in this structure (Schneider et al., 1990) (Figure 4).

Loop 6 in the RLP is ordered and adopts a closed conformation similar to that found in the structure of 8RUC (Figure 4). It also folds over and closes the active site. In place of K334 on loop 6, *C. tepidum* RLP has the structurally equivalent residue R327, which potentially donates a hydrogen bond to a carboxy group (Figures 2 and 3). This arginine residue is conserved in the RLPs from *C. tepidum, Chlorobium limicola*, and *R. palustris*, suggesting that it may be important for the catalytic activity of RLPs. In *R. rubrum* RuBisCO, it was observed that mutating K334 to arginine led to complete loss of catalytic activity (Soper et al., 1988). By analogy, it is not surprising that the RLP from *C. tepidum* (Hanson and Tabita, 2001), and possibly RLPs from other organisms, does not catalyze RuBP-dependent carboxylation.

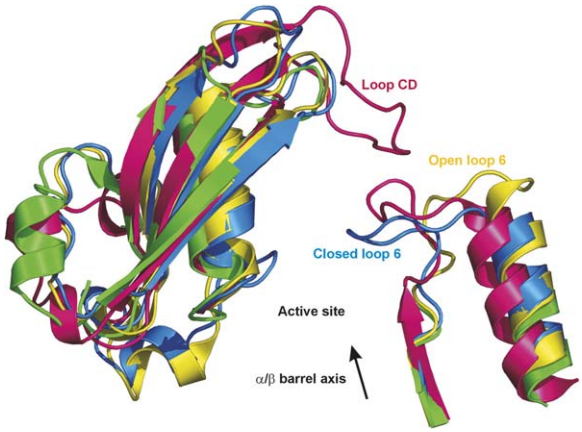

Figure 4. Loop 6 and Loop CD in the RuBisCO Structures

C. tepidum RLP is colored in dark pink, activated spinach RuBisCO (8RUC) is colored in blue, unactivated tobacco RuBisCO (1EJ7) is colored in yellow, and unactivated *R. rubrum* RuBisCO (5RUB) is colored in green. Loop 6 is in an open conformation in the tobacco RuBisCO as shown by the yellow loop in 1EJ7 structure. It is disordered in 5RUB structure (not modeled) as shown by the broken green structure. Loop 6 is in a closed conformation in the spinach RuBisCO structure, as shown by the blue loop in the 8RUC structure. It folds over the active site and shields the substrate from the solvent. In the RLP structure, loop 6 has a similar closed conformation as in 8RUC. Loop CD in the RLP structure is an extended insertion between β strands C and D. This loop folds over the side of the β sheet in the N-terminal domain and closes the active site further outside of loop 6 from an opposite direction. In the other RuBisCO structures, there is only a short loop between β strands C and D. The tip of the loop CD in the RLP is partially aligned with the C-terminal tail β strand in the activated form I RuBisCO structure (8RUC), which becomes ordered and extends toward the active site packing against loop 6 upon RuBP or transition-state analog binding. The axis of the α/β barrel is indicated with an arrow. Only parts of the structures are shown for the purpose of clarity.

Electrostatic Surface Potential

The electrostatic surface potential of the RLP is distinct from those of the other forms of RuBisCO. The front side of the dimeric RuBisCO from *R. rubrum* (5RUB, form II), which contains the active site openings, has a positive surface potential, whereas the RLP has a negative surface potential. The back side of the RLP is also more negatively charged than 5RUB. This is consistent with the finding that *C. tepidum* RLP possesses a function different from forms I, II, or III RuBisCO.

Structural Comparison of the Four Different Forms of RuBisCO

The monomer and dimer structures of the four forms of RuBisCO are well conserved, although the sequence identities among these proteins are relatively low (~30%). Comparisons of the monomer structural alignments of the four forms of RuBisCO based on four representative structures, 8RUC, 5RUB, 1GEH (the structure of RuBisCO from *Thermococcus kodakaraensis* [form III]), and the RLP, suggested that the RLP structure is distinct from the three other forms of RuBisCO and is more similar to forms I and III than to form II (Figure 5A). The structure diversity calculated based on the rmsd of Cα atoms and the number of the aligned residues (Lu, 2000) between the RLP and 1GEH, 8RUC, and 5RUB, are 1.84, 2.24, and 3.07, respectively. To avoid any bias that might result from comparing structures with particular conformations (open versus closed), we compiled and analyzed all 27 available RuBisCO protein structures from the Protein Data Bank (PDB). A structural alignment of the monomers from the RLP and all the 27 available RuBisCO protein structures gave a similar conclusion (Figure 5B).

There are two main differences between the RLP structure and the other three RuBisCO proteins. In the N-terminal domain of the RLP, there is an insertion of a ~14-residue loop, which we call loop CD (Figure 2). The loop extends the length of β strands C and D and approaches the active site opening from the direction opposite from loop 6 and from further away (Figure 4). In form I RuBisCO, the loop between β strands C and D is short and interacts with RuBisCO activase (Ott et al., 2000). The tip of the loop CD in the RLP structure superimposes well with the C-terminal tail β strand of activated form I RuBisCO. This C-terminal tail is disordered in unactivated RuBisCO and extends toward the active site and packs against loop 6 upon RuBP or transition-state analog binding. The conformational change of the C-terminal tail is thought to couple with the movement of loop 6 and to lock it in position (Schneider et al., 1992). Loop CD in the RLP structure may have a similar role in positioning loop 6.

The second major difference is the absence of a β-hairpin turn between helix 6 and β strand 7 of the C-terminal domain in the RLP (Figure 2). The β-hairpin turn is sandwiched between the C-terminal and N-terminal domains in all other RuBisCOs, perhaps to maintain the relative orientation of the two domains.

Protein Functional Linkages of RLPs

The RLP from *C. tepidum*, the structure of which is studied here, is the only RuBisCO protein encoded in its genome. The *rlp* gene resides in a cluster of seven genes on the chromosome, including two oxidoreductases, one bacteriochlorophyll (Bchl) c3(1) hydratase, and three hypothetical proteins (Eisen et al., 2002). In our functional linkage analysis of the *C. tepidum* genome, the *rlp* gene is tightly linked to its neighboring genes, an oxidoreductase (CT1771) and a conserved hypothetical protein (CT1773), by the Gene Cluster method (Figure 6). The *rlp* gene also has many coevolution linkages identified by the Phylogenetic Profile method (Pellegrini et al., 1999) to the enzymes in the bacteriochlorophyll biosynthesis pathway, such as bacteriochlorophyll synthases, protochlorophyllide reductases, and geranylgeranyl hydrogenase. These proteins form a cluster on the functional linkage network. The confidences of these Phylogenetic Profile linkages are relatively low (0.1–0.2); however, they are similar to the confidences of the linkages among the bacteriochlorophyll biosynthesis enzymes themselves. This suggests that the RLP of *C. tepidum* might be involved in processes that are evolutionarily related to bacteriochlorophyll biosynthesis.

R. palustris is the only organism known to date that has two RLPs: RLP1 and RLP2 (Hanson and Tabita, 2003; Larimer et al., 2004). Consistent with the se-

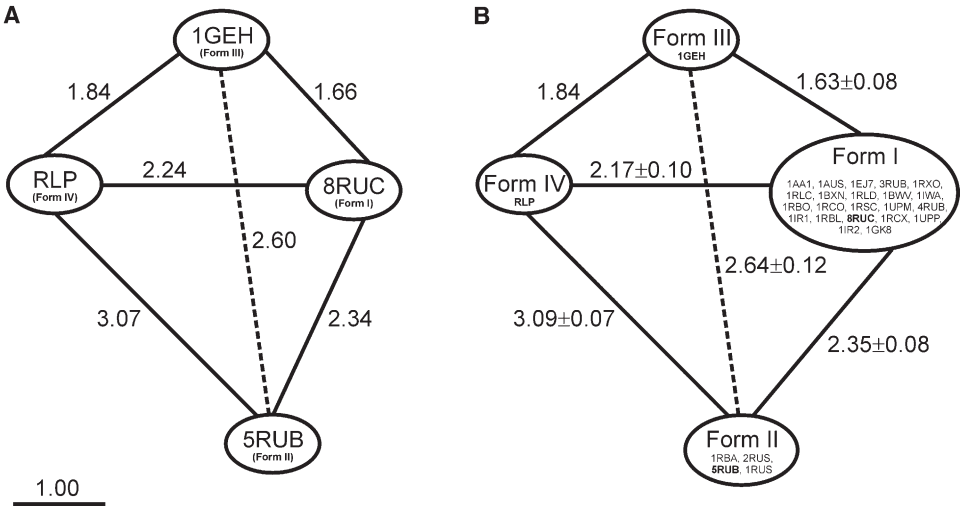

Figure 5. Comparison of the Four Forms of RuBisCO Based on Structural Alignments

(A) Structural comparison among 8RUC, 5RUB, 1GEH, and the RLP. Through the measure of structure diversity (labeled by the edges), the RLP is distinct from the three other forms of RuBisCO, and is more similar to 8RUC and 1GEH than to 5RUB.

(B) Structural comparison among the RLP and 27 available RuBisCO structures. Alignments of 28 RuBisCO structures gave a similar conclusion as in (A). The node "Form I" represents 22 structures of form I RuBisCO from various species with different conformations, and the node "Form II" represents 4 structures of *R. rubrum* RuBisCO. The PDB codes of the structures, the mean, and the standard deviation of structure diversities are given in the figure. The structure diversities are similar to the ones in (A) and the standard deviations are quite small. The distance between each pair of structures is drawn to scale by the structure diversity, except those between 1GEH and 5RUB in (A) and between form III and form II in (B) (dotted lines).

quence alignment analysis of RLPs from different organisms (Hanson and Tabita, 2003), our functional linkage analysis suggests that the two RLPs in the *R. palustris* genome may have different functions (Figure 6). The RLP2 is closely linked to both forms I and II RuBisCO with confidence level above 0.6, which is comparable to the linkage between form I and form II RuBisCOs. This suggests that the RLP2 of *R. palustris* may be involved in a process related to carbon fixation and, possibly, photosynthesis. The RLP2 is the ortholog of the RLP from *C. tepidum* with sequence identity of 66%. Similar to the RLP from *C. tepidum*, the RLP2 is linked to many enzymes in the bacteriochlorophyll biosynthesis pathway by the Phylogenetic Profile method. Again, these results suggest that the RLP2 may possibly have coevolved with those enzymes. In contrast, the other RLP in the genome, RLP1, is only weakly linked to forms I and II RuBisCO, suggesting that the RLP1 may have functionally diverged from the photosynthetic process, and may play a role in some other metabolic process.

The RLP of *B. subtilis* is also called YkrW, and it has been experimentally shown to be involved in a methionine salvage pathway (Ashida et al., 2003), in support of its proposed function based on genetic studies (Murphy et al., 2002; Sekowska and Danchin, 2002). Our functional linkage analysis shows that the RLP from *B. subtilis* is in a network of proteins involved in the methionine salvage pathway, including YkrS, YkrT, YkrV, YkrX, YkrY, and YkrZ (Figure 6). This is consistent with the biochemical studies (Ashida et al., 2003) and demonstrates that functional linkages can provide clues about the functions of proteins.

Discussion

The differences in the active site structures from the other three forms of RuBisCO suggest that *C. tepidum* RLP catalyzes reactions other than carboxylation, possibly enolization. The substrate for the RLP might resemble RuBP in that it could have a P1 phosphate group and a small carbohydrate backbone, but in place of the P2 phosphate, this substrate should have a small hydrophobic group. When we modeled the structure of *B. subtilis* RLP, which catalyzes enolization, we observed that the residues coordinating the P1 phosphate group and the metal ion, which are important for enolization, are relatively conserved between *C. tepidum* RLP and *B. subtilis* RLP, although the P2 site is different.

The catalytic mechanism of *C. tepidum* RLP may also be different from bona fide RuBisCOs. Two major states coupled with the catalytic cycle have been observed in RuBisCO structures: open and closed. In the open state, substrates bind and products leave the active site, and catalysis occurs in the closed state, where the active site is sequestered from the solvent (Duff et al., 2000). Differences between the open and closed structures of form I RuBisCOs are found mainly in four loops near the active site in the large subunit (Schreuder et al., 1993a): (1) the loop at the N terminus (residues 9–21), (2) residues 64–68, (3) loop 6, and (4) the C-terminal tail starting at residue 468. These loops are either disordered or open in the unactivated enzyme, but become ordered or closed upon RuBP or transition-state analog binding. The corresponding structural elements in the RLP structure have features similar to both open and

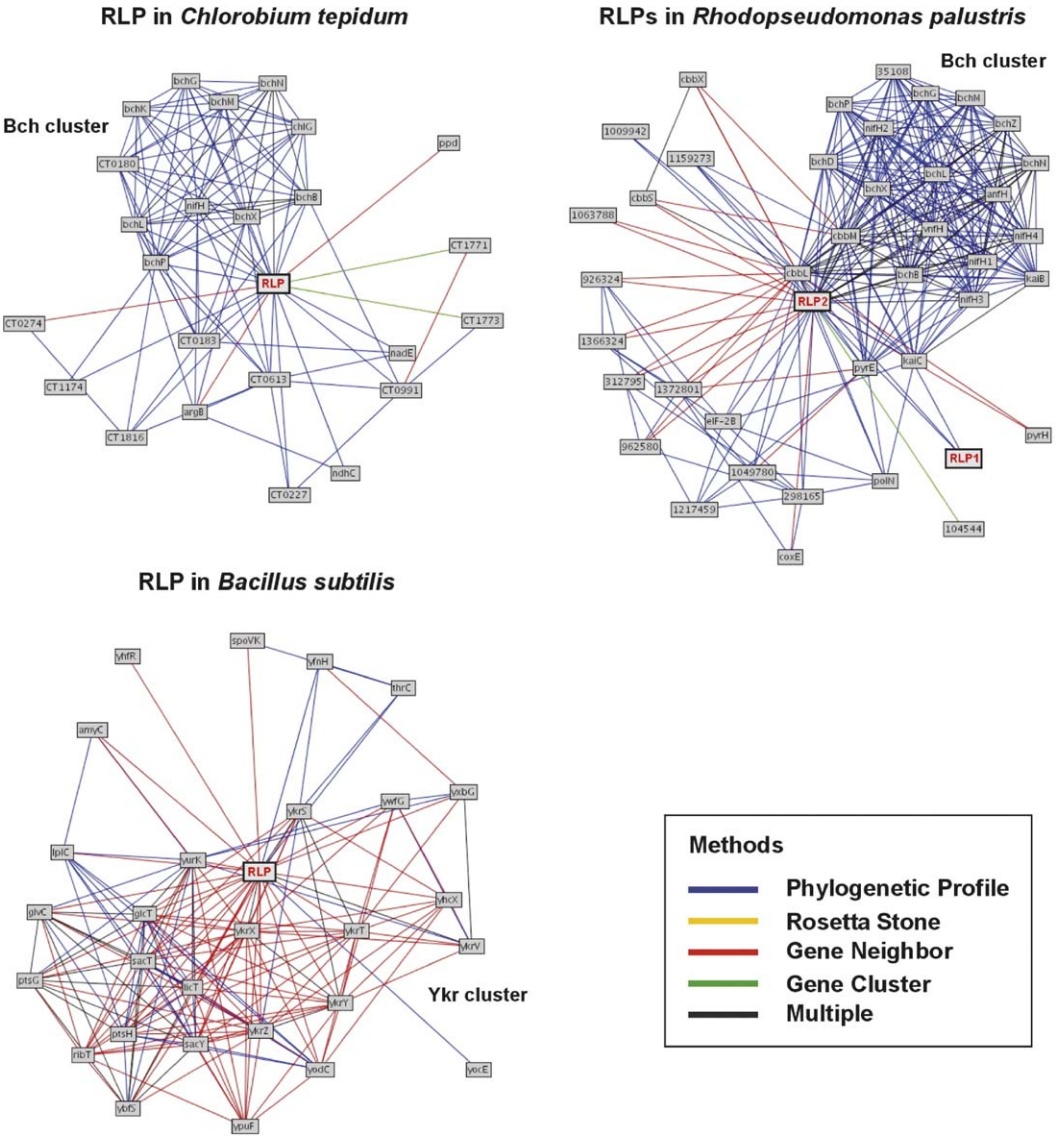

Figure 6. The Inferred Functional Linkages of the RLPs from *C. tepidum*, *R. palustris*, and *B. subtilis*

In *C. tepidum*, the RLP is linked to a large number of enzymes involved in the bacteriochlorophyll biosynthesis pathway, shown as the bch cluster. The linkages are weak, but comparable to the linkages among the bch proteins themselves. In *R. palustris*, the RLP2 is tightly linked to both form I large subunit (CbbL) and form II (CbbM) RuBisCO encoded in the genome. The confidences of the two linkages are high, 0.70 and 0.61, and are comparable to the confidence of the linkage between form I and form II RuBisCOs (0.62). In addition, similar to the RLP from *C. tepidum*, the RLP2 of *R. palustris* is also linked to a cluster of bacteriochlorophyll biosynthesis enzymes (bch cluster). In contrast, the RLP1 of *R. palustris* is only weakly linked to forms I and II RuBisCOs. In *B. subtilis*, the RLP is linked to a cluster of methionine salvage pathway proteins, such as YkrS, YkrT, YkrV, YkrX, YkrY, and YkrZ. The boxes represent the proteins, and the edges represent the linkages between pairs of proteins.

closed form I RuBisCOs. The loop composing residues 47–60 (equivalent to residues of 64–77 in 8RUC) is disordered in the RLP. This is similar to the open state of form I RuBisCO. Loop 6 in the RLP was found in the closed conformation. Although the C-terminal residues of the RLP are disordered, loop CD partially superimposes with the C-terminal tail β strand of form I RuBisCO that becomes ordered and folds against loop 6 upon RuBP or transition-state analog binding. This conformation at the active site is more similar to a closed form I RuBisCO. The mixed conformation states of the loops observed in the RLP, discussed above, suggest that the RLP may have a different catalytic mechanism coupled with the structure from bona fide RuBisCOs.

Bioinformatic analysis of the genome sequences of *C. tepidum* and *R. palustris* presented in the Results suggests that the RLP from *C. tepidum* and its ortholog, the RLP2 from *R. palustris*, have coevolved with the enzymes in the bacteriochlorophyll biosynthesis pathway. Hanson and Tabita (2001) reported that the loss of RLP expression in *C. tepidum* causes less BChl *c* synthesis in the cell. They observed that cultures of

the RLP-knockout strain of *C. tepidum* possessed a lighter green color and contained 20% less bacteriochlorophyll c than wild-type cultures at similar biomass densities. The absorbance of BChl *c* has a blue shift of 5–6 nm, which could be due to a lower degree of aggregation. Frigaard and Bryant (2004) have made several mutants deficient in BChl *c* biosynthesis in *C. tepidum*. Mutant strains lacking normal methylation of BChl *c* contain less BChl *c* or BChl *d* in the cell than the wild-type. The mutants also showed a blue shift in the absorption of BChl *c* or BChl *d* aggregates. The phenotype of the RLP knockout strain resembles those of the mutants lacking enzymes in BChl *c* biosynthesis. Thus, it is tempting to speculate that the function of the RLP is somehow related to or affects the bacteriochlorophyll biosynthesis process.

Despite the diversity in sequences of different forms of RuBisCO, the dimer, which is the functional unit, is structurally well conserved. Bona fide RuBisCO proteins have several unfavorable characteristics as enzymes: low turnover rate, low affinity for the substrate CO_2, CO_2 fixation activity diminished by side reactions, and, of course, oxygenation (Cleland et al., 1998). The disadvantages of RuBisCO as a catalyst may reflect the evolution and possible multifunctionality of the RuBisCO ancestor protein. As illustrated by Ashida et al. (2003), RuBisCO from *R. rubrum* can complement an RLP deletion strain and function as an enolase in the methionine salvage pathway in *B. subtilis*. This suggests that the active site of RuBisCO can accommodate a relatively wide range of substrates and, therefore, RuBisCO is potentially capable of evolving new catalytic activities. To date, no ancestral gene of RuBisCO has been identified. Investigating the function of RLPs from different organisms, from both structural and bioinformatic perspectives, as well as from a functional standpoint, may shed light on the evolution of RuBisCO proteins and their functions.

Experimental Procedures

Overexpression and Purification of the Recombinant RLP

The *rlp* gene from *C. tepidum* with a mutation at residue 328 (Met to Val) had previously been cloned in pET-11a (Hanson and Tabita, 2001) and was overexpressed in *E. coli* strain BL-21. Cells were grown at 37°C to an OD_{600} of 0.6. IPTG (0.1 mM) was added to induce expression overnight at 25°C. All purification steps were conducted either at 4°C or on ice. Cells were resuspended in 20 ml TEMMB (20 mM Tris·Cl, pH 8.0, 1.0 mM EDTA, 10 mM $MgSO_4$, 10 mM BME, and 50 mM $NaHCO_3$) containing 0.1 M NaCl. The cell lysate was clarified by low-speed centrifugation (10,000 × g, 10 min) and then incubated at 50°C for 15 min. The resulting heat-stable extract was clarified by ultracentrifugation (100,000 × g, 1 hr) followed by filtration. This was loaded onto a Q-Sepharose HP column and eluted using a linear gradient from 100 to 300 mM NaCl in TEMMB. The RLP-containing fractions were eluted between 150 and 180 mM NaCl. Concentrated protein was loaded onto a Superdex 200 gel filtration column and eluted with TEMMB containing 150 mM NaCl. This purification yielded 293 mg of homogeneous RLP as determined by SDS-PAGE.

Crystallization and X-Ray Data Collection

The purified protein was stored in a buffer containing 20 mM Tris·HCl (pH 8.0), 3 mM $MgCl_2$, and 50 mM $NaHCO_3$. For crystallization, 11 mM D-xylulose 5-phosphate was added to the protein. The protein concentrated to 30 mg/ml was mixed with the reservoir buffer at a 1:1 ratio. The best diffracting crystals were grown in 20% PEG3350 and 40 mM magnesium formate by hanging drop vapor diffusion method at 22°C. Thin plate crystals appeared within a week. Crystals were flash frozen in liquid nitrogen at –173°C. Cryogenic X-ray diffraction data were collected at the Beamline 8.2.2 at the Advanced Light Source (Lawrence Berkeley National Laboratory, Berkeley, CA). The crystal-to-detector distance was set to 260 mm. Exposure time was 10 seconds per frame, and the oscillation angle was set to 1°. Data were indexed and processed using DENZO and SCALEPACK (Otwinowski and Minor, 1997).

Phasing, Structure Determination, and Structure Analysis

The initial phases were calculated by molecular replacement using EPMR (Kissinger et al., 1999). The structure of the RuBisCO from *R. rubrum* (5RUB) at 5 Å resolution was used as the search model. The structural model was fitted to the electron density map using O (Jones et al., 1991). Refinement was performed by CNS (Brunger et al., 1998) initially for rigid-body refinement and simulated annealing, and by REFMAC5.2 (CCP4, 1994) in the later stage for TLS combined with restrained refinement. Noncrystallographic symmetry was applied in the refinement. Fragments of the two monomers (residues 7–40, 90–160, 225–320, and 333–375) were imposed with tight main-chain and loose side-chain noncrystallographic symmetry restraints. Surface hydrophobicity was calculated by atomic solvation parameters (Eisenberg and McLachlan, 1986). The electrostatic surface potential was calculated based on the PARSE model (Sitkoff et al., 1994). The stereochemistry of the structure was assessed by PROCHECK (CCP4, 1994; Morris et al., 1992), ERRAT (Colovos and Yeates, 1993), What_check (Hooft et al., 1996), and Prove (Pontius et al., 1996). PDB coordinates of RuBisCO structures 8RUC (Andersson, 1996), 5RUB (Schneider et al., 1990), 1GEH (Kitano et al., 2001), 1EJ7 (Duff et al., 2000), and other RuBisCO structures were used for structure comparison. The structural alignments of different RuBisCOs were performed using MAPS (Zhang et al., 2003). Default values of the three criteria of the alignment program were used: the length of the aligned segments should be equal or greater than three residues; the direction difference formed by the Cα-Cβ atoms between each pair of aligned residues should be less than 45°; and the distance between the aligned Cα atoms should be shorter than 3.8 Å. The structure diversity was calculated based on: Structure diversity = $RMS/(N_{match}/N_0)^a$, where RMS is the rmsd of the distances between the matched Cα atoms, N_{match} is the number of matching residues, N_0 is the average number of residues in the two compared structures, and a is the default number (1.5) to control the relative contributions between the rmsd and the number of matched residues (Lu, 2000). The substrate binding model in the active site was generated using LIGPLOT (Wallace et al., 1995). Figures 1, 3A, and 4 were prepared by PyMol (DeLano, 2002).

On the Ramachandran plot (Ramachandran and Sasisekharan, 1968), all the residues are in the most favorable and additional allowed regions, except for residues S46, A204, and F290 from both chains, which are in the generously allowed and disallowed regions. These residues are in well-defined electron density. They are all in the vicinity of the active site, near the α/β barrel C-terminal end. These residues are not directly involved in ligand binding and probably play a role in maintaining the architecture of the active site. Residue S46 is conserved in all RuBisCOs. The side chain of S46 forms a hydrogen bond with the carbonyl oxygen of C42, and it has similar Phi/Psi angles as the equivalent residue S50 from the *R. rubrum* RuBisCO structure (5RUB).

Protein Functional Linkage Analysis

The functional linkages of the RLPs from *C. tepidum* (GI: 21674586), *R. palustris* (RLP1, GI: 39935238; RLP2, GI: 39933339) and *B. subtilis* (GI: 16078423) were calculated by comparing the individual genome sequence of each organism to other 167 fully sequenced genome sequences. The methods were described in detail by Bowers et al. (2004). Default parameters were used (e.g., the sequence similarity threshold was set to e^{-4}). The proteins that are linked to the RLPs were plotted as the nodes in the network and the links were indicated by the edges. The functional linkages

of each protein are available from the Prolinks database at http://mysql5.mbi.ucla.edu/cgi-bin/functionator/pronav.

Acknowledgments

We thank Cedric Bobst and Mary Johnston for preparing purified RLP, Inna Pashkov for assistance in crystallization of the RLP, Dr. Duilio Cascio and the staff at ALS Beamline 8.2.2 for their assistance in data collection, and Howard Hughes Medical Institute and DOE-BER for support. F.R.T. was supported by NIH grant GM24497 and DOE BER grant DE-FG02-91ER20033.

Received: November 8, 2004
Revised: February 17, 2005
Accepted: February 19, 2005
Published: May 10, 2005

References

Andersson, I. (1996). Large structures at high resolution: the 1.6 A crystal structure of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase complexed with 2-carboxyarabinitol bisphosphate. J. Mol. Biol. *259*, 160–174.

Andersson, I., and Taylor, T.C. (2003). Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase. Arch. Biochem. Biophys. *414*, 130–140.

Ashida, H., Saito, Y., Kojima, C., Kobayashi, K., Ogasawara, N., and Yokota, A. (2003). A functional link between RuBisCO-like protein of *Bacillus* and photosynthetic RuBisCO. Science *302*, 286–290.

Baker, T.S., Eisenberg, D., and Eiserling, F. (1977). Ribulose bisphosphate carboxylase: a two-layered, square-shaped molecule of symmetry 422. Science *196*, 293–295.

Bowers, P.M., Pellegrini, M., Thompson, M.J., Fierro, J., Yeates, T.O., and Eisenberg, D. (2004). Prolinks: a database of protein functional linkages derived from coevolution. Genome Biol. *5*, R35.

Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. (1998). Crystallography & NMR system: a new software suite for macromolecular structure determination. Acta Crystallogr. D Biol. Crystallogr. *54*, 905–921.

Chapman, M.S., Suh, S.W., Cascio, D., Smith, W.W., and Eisenberg, D. (1987). Sliding-layer conformational change limited by the quaternary structure of plant RuBisCO. Nature *329*, 354–356.

Cleland, W.W., Andrews, T.J., Gutteridge, S., Hartman, F.C., and Lorimer, G.H. (1998). Mechanism of RuBisCO: the carbamate as general base. Chem. Rev. *98*, 549–562.

CCP4 (Collaborative Computational Project, Number 4)(1994). The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. *50*, 760–763.

Colovos, C., and Yeates, T.O. (1993). Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. *2*, 1511–1519.

Curmi, P.M., Cascio, D., Sweet, R.M., Eisenberg, D., and Schreuder, H. (1992). Crystal structure of the unactivated form of ribulose-1,5-bisphosphate carboxylase/oxygenase from tobacco refined at 2.0-Å resolution. J. Biol. Chem. *267*, 16980–16989.

Dandekar, T., Snel, B., Huynen, M., and Bork, P. (1998). Conservation of gene order: A fingerprint of proteins that physically interact. Trends Biochem. Sci. *23*, 324–328.

DeLano, W.L. (2002). The PyMOL Molecular Graphics System (http://www.pymol.org).

Duff, A.P., Andrews, T.J., and Curmi, P.M. (2000). The transition between the open and closed states of RuBisCO is triggered by the inter-phosphate distance of the bound bisphosphate. J. Mol. Biol. *298*, 903–916.

Eisen, J.A., Nelson, K.E., Paulsen, I.T., Heidelberg, J.F., Wu, M., Dodson, R.J., Deboy, R., Gwinn, M.L., Nelson, W.C., Haft, D.H., et al. (2002). The complete genome sequence of *Chlorobium tepidum* TLS, a photosynthetic, anaerobic, green-sulfur bacterium. Proc. Natl. Acad. Sci. USA *99*, 9509–9514.

Eisenberg, D., and McLachlan, A.D. (1986). Solvation energy in protein folding and binding. Nature *319*, 199–203.

Ellis, R.J. (1979). The most abundant protein in the world. Trends Biochem. Sci. *4*, 241–244.

Enright, A.J., Iliopoulos, I., Kyrpides, N.C., and Ouzounis, C.A. (1999). Protein interaction maps for complete genomes based on gene fusion events. Nature *402*, 86–90.

Finn, M.W., and Tabita, F.R. (2003). Synthesis of catalytically active form III ribulose 1,5-bisphosphate carboxylase/oxygenase in archaea. J. Bacteriol. *185*, 3049–3059.

Frigaard, N.U., and Bryant, D.A. (2004). Seeing green bacteria in a new light: genomics-enabled studies of the photosynthetic apparatus in green sulfur bacteria and filamentous anoxygenic phototrophic bacteria. Arch. Microbiol. *182*, 265–276.

Hansen, S., Vollan, V.B., Hough, E., and Andersen, K. (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.

Hanson, T.E., and Tabita, F.R. (2001). A ribulose-1,5-bisphosphate carboxylase/oxygenase (RuBisCO)-like protein from *Chlorobium tepidum* that is involved with sulfur metabolism and the response to oxidative stress. Proc. Natl. Acad. Sci. USA *98*, 4397–4402.

Hanson, T.E., and Tabita, F.R. (2003). Insights into the stress response and sulfur metabolism revealed by proteome analysis of a *Chlorobium tepidum* mutant lacking the RuBisCO-like protein. Photosynth. Res. *78*, 231–248.

Hooft, R.W., Vriend, G., Sander, C., and Abola, E.E. (1996). Errors in protein structures. Nature *381*, 272.

Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, M. (1991). Improved methods for building protein models in electron density maps and the location of errors in these models. Acta Crystallogr. A *47*, 110–119.

Kissinger, C.R., Gehlhaar, D.K., and Fogel, D.B. (1999). Rapid automated molecular replacement by evolutionary search. Acta Crystallogr. D Biol. Crystallogr. *55*, 484–491.

Kitano, K., Maeda, N., Fukui, T., Atomi, H., Imanaka, T., and Miki, K. (2001). Crystal structure of a novel-type archaeal rubisco with pentagonal symmetry. Structure (Camb) *9*, 473–481.

Klenk, H.P., Clayton, R.A., Tomb, J.F., White, O., Nelson, K.E., Ketchum, K.A., Dodson, R.J., Gwinn, M., Hickey, E.K., Peterson, J.D., et al. (1997). The complete genome sequence of the hyperthermophilic, sulphate-reducing archaeon *Archaeoglobus fulgidus*. Nature *390*, 364–370.

Knight, S., Andersson, I., and Branden, C.I. (1990). Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution: subunit interactions and active site. J. Mol. Biol. *215*, 113–160.

Kunst, F., Ogasawara, N., Moszer, I., Albertini, A.M., Alloni, G., Azevedo, V., Bertero, M.G., Bessieres, P., Bolotin, A., Borchert, S., et al. (1997). The complete genome sequence of the gram-positive bacterium *Bacillus subtilis*. Nature *390*, 249–256.

Larimer, F.W., Chain, P., Hauser, L., Lamerdin, J., Malfatti, S., Do, L., Land, M.L., Pelletier, D.A., Beatty, J.T., Lang, A.S., et al. (2004). Complete genome sequence of the metabolically versatile photosynthetic bacterium *Rhodopseudomonas palustris*. Nat. Biotechnol. *22*, 55–61.

Lu, G. (2000). TOP: a new method for protein structure comparisons and similarity searches. J. Appl. Crystallogr. *33*, 176–183.

Lundqvist, T., and Schneider, G. (1989a). Crystal structure of the binary complex of ribulose-1,5-bisphosphate carboxylase and its product, 3-phospho-D-glycerate. J. Biol. Chem. *264*, 3643–3646.

Lundqvist, T., and Schneider, G. (1989b). Crystal structure of the complex of ribulose-1,5-bisphosphate carboxylase and a transition state analogue, 2-carboxy-D-arabinitol 1,5-bisphosphate. J. Biol. Chem. *264*, 7078–7083.

Lundqvist, T., and Schneider, G. (1991a). Crystal structure of activated ribulose-1,5-bisphosphate carboxylase complexed with its substrate, ribulose-1,5-bisphosphate. J. Biol. Chem. *266*, 12604–12611.

Lundqvist, T., and Schneider, G. (1991b). Crystal structure of the ternary complex of ribulose-1,5-bisphosphate carboxylase, Mg(II), and activator CO2 at 2.3-Å resolution. Biochemistry *30*, 904–908.

Maeda, N., Kitano, K., Fukui, T., Ezaki, S., Atomi, H., Miki, K., and Imanaka, T. (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.

Marcotte, E.M., Pellegrini, M., Ng, H.L., Rice, D.W., Yeates, T.O., and Eisenberg, D. (1999). Detecting protein function and protein-protein interactions from genome sequences. Science *285*, 751–753.

Morris, A.L., MacArthur, M.W., Hutchinson, E.G., and Thornton, J.M. (1992). Stereochemical quality of protein structure coordinates. Proteins *12*, 345–364.

Murphy, B.A., Grundy, F.J., and Henkin, T.M. (2002). Prediction of gene function in methylthioadenosine recycling from regulatory signals. J. Bacteriol. *184*, 2314–2318.

Ott, C.M., Smith, B.D., Portis, A.R., Jr., and Spreitzer, R.J. (2000). Activase region on chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase: nonconservative substitution in the large subunit alters species specificity of protein interaction. J. Biol. Chem. *275*, 26241–26244.

Otwinowski, Z., and Minor, W. (1997). Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. *276*, 307–326.

Overbeek, R., Fonstein, M., D'Souza, M., Pusch, G.D., and Maltsev, N. (1999). The use of gene clusters to infer functional coupling. Proc. Natl. Acad. Sci. USA *96*, 2896–2901.

Pellegrini, M., Marcotte, E.M., Thompson, M.J., Eisenberg, D., and Yeates, T.O. (1999). Assigning protein functions by comparative genome analysis: protein phylogenetic profiles. Proc. Natl. Acad. Sci. USA *96*, 4285–4288.

Pellegrini, M., Thompson, M., Fierro, J., and Bowers, P. (2001). Computational method to assign microbial genes to pathways. J. Cell. Biochem. Suppl. *96*, 106–109.

Pontius, J., Richelle, J., and Wodak, S.J. (1996). Deviations from standard atomic volumes as a quality measure for protein crystal structures. J. Mol. Biol. *264*, 121–136.

Ramachandran, G.N., and Sasisekharan, V. (1968). Conformation of polypeptides and proteins. Adv. Protein Chem. *23*, 283–438.

Schneider, G., Lindqvist, Y., and Branden, C.I. (1992). RuBisCO: structure and mechanism. Annu. Rev. Biophys. Biomol. Struct. *21*, 119–143.

Schneider, G., Lindqvist, Y., and Lundqvist, T. (1990). Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase from *Rhodospirillum rubrum* at 1.7 Å resolution. J. Mol. Biol. *211*, 989–1008.

Schreuder, H.A., Knight, S., Curmi, P.M., Andersson, I., Cascio, D., Branden, C.I., and Eisenberg, D. (1993a). Formation of the active site of ribulose-1,5-bisphosphate carboxylase/oxygenase by a disorder-order transition from the unactivated to the activated form. Proc. Natl. Acad. Sci. USA *90*, 9968–9972.

Schreuder, H.A., Knight, S., Curmi, P.M., Andersson, I., Cascio, D., Sweet, R.M., Branden, C.I., and Eisenberg, D. (1993b). Crystal structure of activated tobacco RuBisCO complexed with the reaction-intermediate analogue 2-carboxy-arabinitol 1,5-bisphosphate. Protein Sci. *2*, 1136–1146.

Sekowska, A., and Danchin, A. (2002). The methionine salvage pathway in *Bacillus subtilis*. BMC Microbiol. *2*, 8.

Shibata, N., Inoue, T., Fukuhara, K., Nagara, Y., Kitagawa, R., Harada, S., Kasai, N., Uemura, K., Kato, K., Yokota, A., and Kai, Y. (1996). Orderly disposition of heterogeneous small subunits in D-ribulose-1,5-bisphosphate carboxylase/oxygenase from spinach. J. Biol. Chem. *271*, 26449–26452.

Sitkoff, D., Lockhart, D.J., Sharp, K.A., and Honig, B. (1994). Calculation of electrostatic effects at the amino terminus of an alpha helix. Biophys. J. *67*, 2251–2260.

Soderlind, E., Schneider, G., and Gutteridge, S. (1992). Substitution of ASP193 to ASN at the active site of ribulose-1,5-bisphosphate carboxylase results in conformational changes. Eur. J. Biochem. *206*, 729–735.

Soper, T.S., Mural, R.J., Larimer, F.W., Lee, E.H., Machanoff, R., and Hartman, F.C. (1988). Essentiality of Lys-329 of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* as demonstrated by site-directed mutagenesis. Protein Eng. *2*, 39–44.

Spreitzer, R.J., and Salvucci, M.E. (2002). RuBisCO: structure, regulatory interactions, and possibilities for a better enzyme. Annu. Rev. Plant Biol. *53*, 449–475.

Sugawara, H., Yamamoto, H., Shibata, N., Inoue, T., Okada, S., Miyake, C., Yokota, A., and Kai, Y. (1999). Crystal structure of carboxylase reaction-oriented ribulose 1,5-bisphosphate carboxylase/oxygenase from a thermophilic red alga, *Galdieria partita*. J. Biol. Chem. *274*, 15655–15661.

Tabita, F.R. (1999). Microbial ribulose 1,5-bisphosphate carboxylase/oxygenase: A different perspective. Photosynth. Res. *60*, 1–28.

Taylor, T.C., and Andersson, I. (1996). Structural transitions during activation and ligand binding in hexadecameric RuBisCO inferred from the crystal structure of the activated unliganded spinach enzyme. Nat. Struct. Biol. *3*, 95–101.

Taylor, T.C., and Andersson, I. (1997a). Structure of a product complex of spinach ribulose-1,5-bisphosphate carboxylase/oxygenase. Biochemistry *36*, 4041–4046.

Taylor, T.C., and Andersson, I. (1997b). The structure of the complex between RuBisCO and its natural substrate ribulose 1,5-bisphosphate. J. Mol. Biol. *265*, 432–444.

Taylor, T.C., Backlund, A., Bjorhall, K., Spreitzer, R.J., and Andersson, I. (2001). First crystal structure of RuBisCO from a green alga, *Chlamydomonas reinhardtii*. J. Biol. Chem. *276*, 48159–48164.

Taylor, T.C., Fothergill, M.D., and Andersson, I. (1996). A common structural basis for the inhibition of ribulose 1,5-bisphosphate carboxylase by 4-carboxyarabinitol 1,5-bisphosphate and xylulose 1,5-bisphosphate. J. Biol. Chem. *271*, 32894–32899.

Wallace, A.C., Laskowski, R.A., and Thornton, J.M. (1995). LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. *8*, 127–134.

Watson, G.M., Yu, J.P., and Tabita, F.R. (1999). Unusual ribulose 1,5-bisphosphate carboxylase/oxygenase of anoxic Archaea. J. Bacteriol. *181*, 1569–1575.

Zhang, K.Y., Cascio, D., and Eisenberg, D. (1994). Crystal structure of the unactivated ribulose 1,5-bisphosphate carboxylase/oxygenase complexed with a transition state analog, 2-carboxy-D-arabinitol 1,5-bisphosphate. Protein Sci. *3*, 64–69.

Zhang, Z., Lindstam, M., Unge, J., Peterson, C., and Lu, G. (2003). Potential for dramatic improvement in sequence alignment against structures of remote homologous proteins by extracting structural information from multiple structure alignment. J. Mol. Biol. *332*, 127–142.

Accession Numbers

The structure has been deposited in the Protein Data Bank (PDB code 1YKW).