Available online at www.sciencedirect.com
ScienceDirect
Plant Physiology and Biochemistry 46 (2008) 275–291
Plant Physiology and Biochemistry
www.elsevier.com/locate/plaphy

Research article

Structure and function of Rubisco

Inger Andersson [a,*], Anders Backlund [b]

[a] Department of Molecular Biology, Swedish University of Agricultural Sciences, Husargatan 3, BMC Box 590, S-751 24 Uppsala, Sweden
[b] Division of Pharmacognosy, Department of Medicinal Chemistry, Uppsala University, BMC Box 574, S-751 23 Uppsala, Sweden

Received 20 December 2007
Available online 12 January 2008

Abstract

Ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) is the major enzyme assimilating CO_2 into the biosphere. At the same time Rubisco is an extremely inefficient catalyst and its carboxylase activity is compromised by an opposing oxygenase activity involving atmospheric O_2. The shortcomings of Rubisco have implications for crop yield, nitrogen and water usage, and for the global carbon cycle. Numerous high-resolution crystal structures of different forms of Rubisco are now available, including structures of mutant enzymes. This review uses the information provided in these structures in a structure-based sequence alignment and discusses Rubisco function in the context of structural variations at all levels – amino acid sequence, fold, tertiary and quaternary structure – with an evolutionary perspective and an emphasis on the structural features of the enzyme that may determine its function as a carboxylase.

Keywords: Rubisco; Structure-function studies; CO_2/O_2 specificity; Evolution; Structure-based alignment

1. Introduction

Life on earth is almost exclusively dependent on the ability of photosynthetic organisms to sequester inorganic CO_2 of the atmosphere into organic carbon of the biosphere via the Calvin-Benson-Bassham pathway. The primary photosynthetic CO_2 reduction reaction, the binding of CO_2 to the acceptor-molecule ribulose-1,5-bisphosphate (RuBP) to form two molecules of 3-phosphoglycerate, is catalysed by the enzyme RuBP carboxylase/oxygenase (EC 4.1.1.39, Rubisco). Rubisco is found in most autotrophic organisms from prokaryotes (photosynthetic and chemoautotrophic bacteria, cyanobacteria and archaea) to eukaryotes (various algae and higher plants). It has been estimated that Rubisco can comprise up to 50% of the total soluble protein in the plant leaf or inside the microbe [21]. Perhaps less obvious, but as ubiquitous, is its presence in phytoplankton in the sea, which are estimated to provide more than 45% of global net primary production annually [25].

Due to the importance and abundance of Rubisco, aspects of the genetics, microbiology, molecular biology, biochemistry and evolution of the enzyme have been studied intensely. More recently, the realization of the importance of CO_2 as a greenhouse gas and the occurrence of rising levels of CO_2 in the Earth's atmosphere has drawn new attention to Rubisco's role in these processes. However, ultimately it is the unique and peculiar features of Rubisco as a catalyst that drives the continuous interest in the enzyme: Rubisco is an extremely slow catalyst and moreover its carboxylation activity is compromised by competing side-reactions, the most notable with another atmospheric gas, O_2, which attacks the same enediol-intermediate of RuBP. The opposing oxygenase activity of Rubisco results in the synthesis of phosphoglycolate, a molecule of limited use to most organisms. Phosphoglycolate is re-circulated by photorespiration, an energy-requiring salvage pathway. This causes a constant drain on the pool of the sugar substrate (RuBP) and results in a decrease of the efficiency of carbon fixation by up to 50%. The opposing

Abbreviations: RuBP, ribulose-1,5-bisphosphate; Rubisco, ribulose-1,5-bisphosphate carboxylase/oxygenase; RLP, Rubisco-like protein; PDB, Protein Data Bank; L, large subunit; S, small subunit; *rbc*L, Rubisco large subunit gene; *rbc*S, Rubisco small subunit gene; 2CABP, 2-carboxyarabinitol-1,5-bisphosphate.

* Corresponding author. Tel.: +46 18 471 4288; fax: +46 18 536 971.
E-mail address: inger@xray.bmc.uu.se (I. Andersson).

0981-9428/$ - see front matter © 2008 Elsevier Masson SAS. All rights reserved.
doi:10.1016/j.plaphy.2008.01.001

oxygenase activity is an intrinsic characteristic of Rubisco, the extent of which depends on the properties of the particular enzyme studied. Thus, the key to the efficiency of any particular Rubisco enzyme lies hidden in the fine details of its three-dimensional structure and this has motivated structural studies with the ultimate aim to suppress oxygenation by Rubisco as a means to improve crop yields. Detailed discussions of several aspects of Rubisco research are available [4,6,15,37,40,75,95–97,100]. The present article summarizes recent advances in the knowledge of the molecular basis for the function of Rubisco, and relates to the evolution of different photosynthesising organisms.

2. Differences in catalytic efficiency of enzymes from diverse origins

The efficiency with which CO_2 is able to compete with O_2 is quantified by the CO_2/O_2 specificity factor (often referred to as τ, or Ω) and is defined as V_cK_o/V_oK_c, where V_c and V_o are the maximal velocities of carboxylation and oxygenation, respectively, and K_c and K_o are the Michaelis constants for CO_2 and O_2, respectively [55]. Thus the relative rates for carboxylation and oxygenation are defined by the product of the specificity factor and the ratio of CO_2 to O_2 concentrations at the active site. The specificity values of Rubisco enzymes from different species and evolutionary lineages differ substantially [44]. Some photosynthesizing bacteria of the α-proteobacteria group have the lowest specificity values (5–40) whereas members of the Rhodophyta (red algae) have the highest (180–240). Chlorophyta, such as higher plants and green algae have intermediate specificity values in the range of 60–100. Taking this single parameter into consideration, it would appear that evolution in one aspect has worked by positive selection towards a more efficient enzyme. However, an inverse correlation between specificity and turnover rate (V_c or k_{cat} for carboxylation) has been observed [8,44] with e.g. bacteria displaying low specificity values and high turnover rates whereas higher plants have high specificity values coupled to low turnover rates. In addition, intracellular CO_2 and O_2 concentrations vary considerably among species because several organisms, including plants, have evolved mechanisms (carboxysomes, pyrenoids, C4- and CAM metabolisms) that concentrate CO_2 at the carboxylation site [32,80]. In order to compare net CO_2 fixation, it may thus be more appropriate to consider differences between the velocities of carboxylation and oxygenation in the intracellular gaseous environment [55,95]. It also appears that environmental factors, such as the temperature and the aridity of the environment from which the organism evolved are important factors that influence Rubisco's carboxylation capacity [28,84]. See also ref. [45].

3. Structure overview

To enable a discussion on similarity on a wider basis, an evolutionary diverse set of 109 Rubisco and Rubisco-like protein (RLP) amino acid sequences was compiled (Table 1). The treatment differs somewhat from most phylogenetic analyses, where a multiple sequence alignment is performed using a software tool such as ClustalX [42]. In these 'traditional' procedures, alignment of sequence data is achieved on grounds of an amino acid substitution matrix, defining the probability of exchanging one amino acid with another. By invoking penalties for sequence gap opening and extension etc., an alignment 'score' is obtained, which makes comparisons possible. Although useful in many instances, the limitations of this approach are evident when alignment of more complex sequences is attempted.

Table 1
List of accession numbers of *rbc*L and *rbc*S amino acid sequences used for the phylogenetic analysis. A complete and annotated alignment is available from the authors on request

AB004883, AB005248, AF012127, AB018007, AB018008, AB018555, AB020927, AB020942, AB020943, AB020945, AB042068, AB122071, AB233880, AE000782, AF034479, AF038430, AF041468, AF041820, AF046932, AF046933, AF047688, AF065474, AF101233, AF129925, AF139469, AF150665, AF195952, AF207527, AF211846, AF233069, AF267640, AF307091, AF479878, AY034235, AY177738, AY177748, AY450592, AY836181, AY836183, BA000001, BA000019, BA000040, CP000144, CP000254, CP000463, CP000477, CP000493, CP000505, D00206, D11140, D13539, D13539, D13971, D13971, D30764, D43621, D43622, D63675, D90204, DQ286015, EU196765, J01399, J01443, L02522, L10212, L13929, L14403, L15300, L32182, L42940, L76557, L77117, L82000, M15842, M17744, M24288, M26396, M34536, M35728, M59080, M64624, M76402, M98515, NC002932, NC_005296, NC006510, NC_000925, NC_001713, NC_002927, NC_007955, NZAAOT01000002, RCU23145, TFECBBMA, U05941, U28421, U29933, U38804, U39856, U46156, U46156, U55037, U91966, X00286, X00630, X00806, X03220, X03220, X03821, X03853, X04472, X04976, X05982, X07515, X13974, X15886, X15901, X16039, X17231, X17252, X17597, X51811, X51964, X52503, X53045, X53426, X54091, X54532, X55372, X55524, X55829, X57022, X57266, X57359, X57359, X62119, X66617, X69760, X70810, X75334, X83095, X86563, X89403, X98414, Y00322, Y00431, Z00044, Z95552

Rubisco is a fortunate case, because a sufficient number of three-dimensional structures of different evolutionary lineages are available in the Protein Data Bank (PDB), allowing for detailed evaluation of homologous amino acids in different species of the enzyme. From this information a structure based rather than substitution based alignment of Rubisco amino acid sequences can be attempted. The aim here is thus to enable discussion of function in the context of structural variations at all levels – amino acid sequence, fold, tertiary and quaternary structure – with an evolutionary perspective.

3.1. Molecular forms of Rubisco differ

All Rubisco enzymes are multimeric (Fig. 1). Two different types of subunits occur: catalytic large (L, 50–55 kDa), and small (S, 12–18 kDa) subunits. From a structural point of view an important distinction between different molecular forms of Rubisco depends on the presence or absence of the small subunit. The most common form (form I) of Rubisco is composed of large and small subunits in a hexadecameric structure, L8S8 (Fig. 1A–C). The molecule exhibits local 422 symmetry and consists of a core of four L2 dimers arranged around a four-fold axis, capped at each end by four small subunits [54]. The small subunit is not essential for

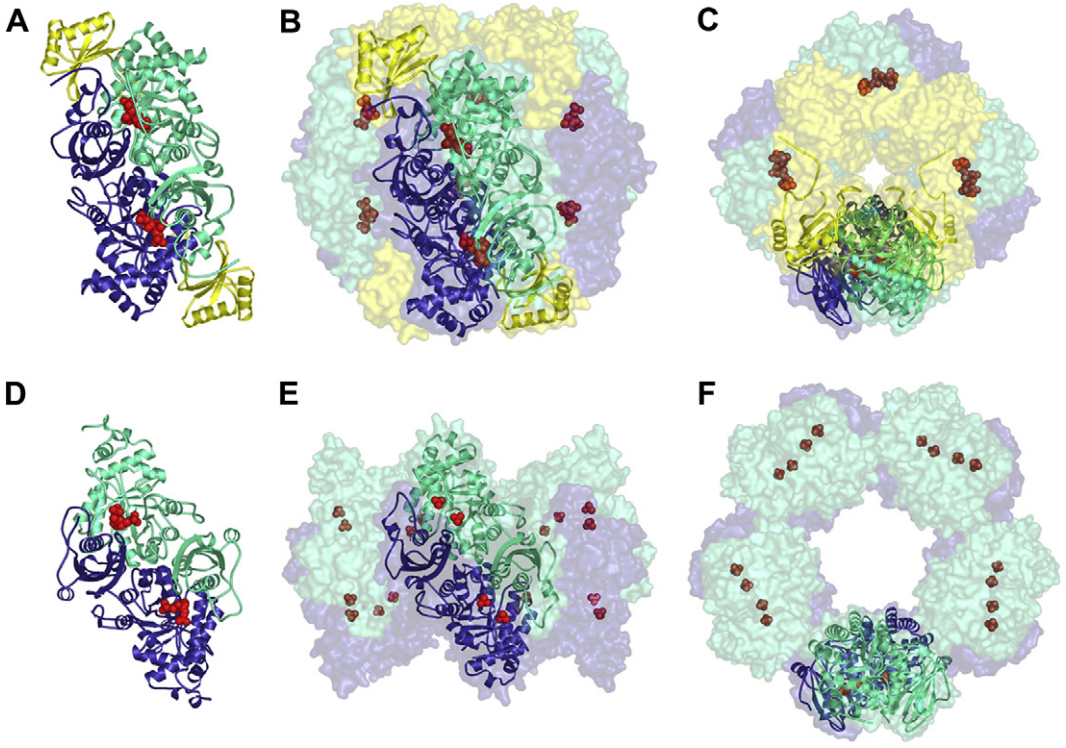


Fig. 1. The different arrangement of the quaternary structure of Rubisco enzymes showing the molecular symmetry. A) The L2S2 unit of form I Rubisco from spinach viewed along the 2-fold symmetry axis. Large subunits are blue and green, small subunits are yellow, and the substrate mimic (2CABP) is displayed as red spheres. B) The entire L8S8 hexadecamer is shown viewed along the same 2-fold axis and C) along the 4-fold axis. D) The dimeric form II Rubisco from *Rhodospirillum rubrum* showing the 2-fold symmetry. E) and F) The L10 Rubisco from *Thermococcus kodakaraensis* viewed along the 2-fold and 5-fold axes, respectively. Sulphate ions bound in the active site are displayed as red spheres.

catalysis, because several configurations of large subunits, including the *Synechococcus* large subunit octamer, still retain some carboxylase activity and have unperturbed specificity [5]. On the other hand, hybrids containing foreign small subunits and enzymes with mutations in the small subunit often display altered holoenzyme stability and/or specificity (reviewed in refs. [97,100], see also Section 4.1.4). Form I Rubisco is present in most chemoautotrophic bacteria, cyanobacteria, red and brown algae, and in all higher plants. The incorporation of a small subunit in Rubisco, giving rise to the form I enzymes, occurs at one specific stage in Rubisco evolution (as indicated in Fig. 2).

Based on amino acid sequences of the form I enzymes, a distinction has been made [17,103] between green-type enzymes (forms I A and B from cyanobacteria, eukaryotic algae and higher plants) and red-type enzymes (forms I C and D from non-green algae and phototropic bacteria). To date, crystal structures of the form I B hexadecameric (L8S8) enzyme have been determined from *Spinacia oleracea* (spinach), *Nicotiana tabacum* (tobacco), *Oryza sativa* (rice), *Synechococcus* PCC6301, and *Chlamydomonas reinhardtii* (Table 2). No structure information is yet available for the form I A, however amino acid sequence data indicate a very close relationship with no readily observable differences. Crystal structures of form I C/D Rubisco have been determined from the enzymes from *Ralstonia eutropha* and *Galdieria partita* [35,101]. From an evolutionary perspective, however, it should be noted that there is no evidence for a multiple origin of Rubisco, and the endosymbiontic event resulting in the chloroplast organelle has been suggested to be a singularity. Hence, the different forms (including those of rhodophytes and phaeophytes) would be derived and not plesiomorphic. Evaluating the results from the structure-based alignment of *rbc*L/*rbc*S, with sequences as far as possible from the same genome, a somewhat different pattern emerges. As shown in Fig. 2, two large and very strongly supported groups (both from combined and separate analyses) emerges. One with mainly α-proteobacteria at the base, leading to several groups of algae, including red and brown algae, bacillariophytes, and dinoflagellates. A second group, with its base in α, β, and γ-bacteria, lead to cyanobacteria (and eventually green plants), with the Glaucocystophyta represented by *Cyanophora paradoxa* as the basal-most taxon among the latter.

The form II enzyme is a dimer of large subunits $(L2)_n$ and lacks small subunits (Fig. 1D). The form II enzyme was initially discovered in purple non-sulphur bacteria, but has also been found in several chemoautotropic bacteria [92] and in eukaryotic dinoflagellates [68,83]. This has interesting implications for evolutionary relationships. Several non-sulphur phototropic bacteria, i.e. *Rhodobacter sphaeroides*, *R. capsulatus*, several *Thiobacillus* sp., and *Hydrogenovibrio marinus* contain both form I and form II enzymes [31,38]. The first crystal structure of Rubisco was from the dimeric form II enzyme from *Rhodospirillum rubrum* [87,88]. In our analysis, the included genes of form II enzymes congregate in a strongly supported group at the base of the more advanced form I enzymes.

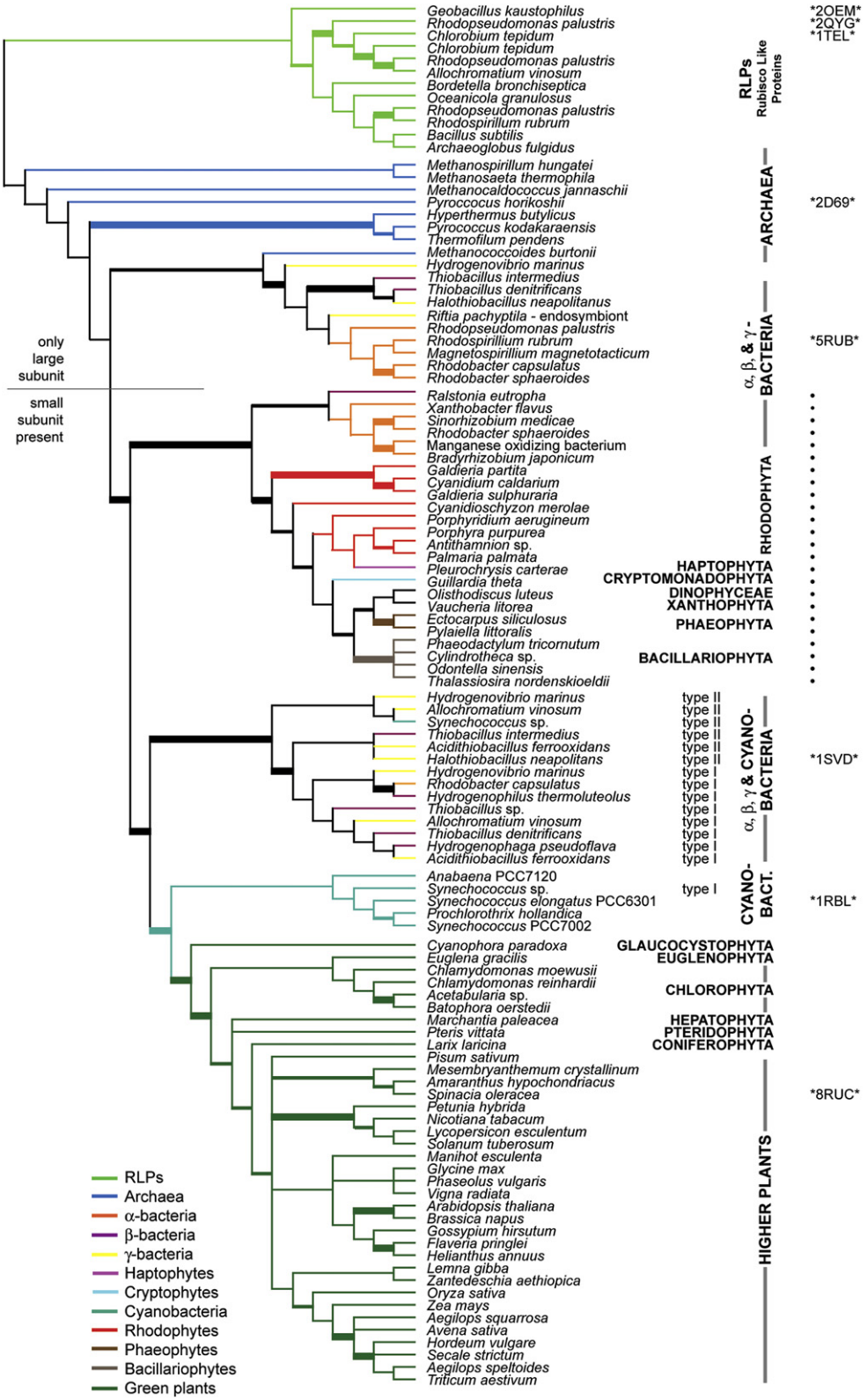


Fig. 2. Strict consensus of 58 equally parsimonious trees, resulting from analyses of combined alignment of *rbc*S and *rbc*L. The alignment matrix consisted of 621 informative characters from 109 taxa, with most parsimonious tree length 9069 steps, CI = 0.474, and RI = 0.362. Branches are colour coded according to systematic position, taxa labelled with • all feature the *rbc*S carboxy-terminal β-hairpin structure (see also Fig. 5), type I/II indicates *rbc*S gene subtype with differences in the amino-terminal region, and *ZXXX* indicates PDB code for models used in the sequence alignment. Branches with double or quadruple width indicate bootstrap support values >75% and >90%, respectively.

Table 2
Coordinates for Rubisco structures deposited in the database

PDB code	Organism		Dep. date	Complex	Resolution (Å)
5RUB	*Rhodospirillum rubrum*	[87]	May 1990	–	1.70
3RUB	Tobacco	[16]	May 1990	$SO_4^{2-} \times 2$	2.00
4RUB	Tobacco	[90]	May 1990	Mg^{2+}-CO_2/2CABP	2.70
9RUB	*Rhodospirillum rubrum*	[59]	Nov 1990	Mg^{2+}-CO_2/RuBP	2.60
1RUS	*Rhodospirillum rubrum*	[58]	Oct 1991	3PGA	2.90
2RUS	*Rhodospirillum rubrum*	[60]	Oct 1991	Mg^{2+}-CO_2	2.30
1RBA	*Rhodospirillum rubrum*	[94]	Nov 1991	D193N mutant	2.60
1RBL	*Synechococcus*	[70]	May 1993	Mg^{2+}-CO_2/2CABP	2.20
1RLC	Tobacco	[113]	Aug 1993	2CABP	2.70
1RLD	Tobacco	[114]	Dec 1993	–	2.50
1RSC	*Synechococcus*	[72]	Mar 1994	XuBP	2.30
1AUS	Spinach	[106]	Jun 1995	Mg^{2+}-CO_2	2.20
8RUC	Spinach	[2,54]	Feb 1996	Mg^{2+}-CO_2/2CABP	1.60
1RBO	Spinach	[110]	Oct 1996	4CABP	2.30
1RCO	Spinach	[110]	Oct 1996	XuBP	2.30
1RCX	Spinach	[107]	Dec 1996	RuBP	2.40
1RXO	Spinach	[107]	Dec 1996	Ca^{2+}-CO_2/RuBP	2.20
1AA1	Spinach	[108]	Jan 1997	Mg^{2+}-CO_2/3PGA	2.20
1BWV	*Galdieria partita*	[101]	Sep 1998	Mg^{2+}-CO_2/2CABP	2.40
1BXN	*Ralstonia eutropha*	[35]	Oct 1998	$PO_4^{3-} \times 2$	2.70
1EJ7	Tobacco	[20]	Mar 2000	$PO_4^{3-} \times 2$	2.45
1GK8	*Chlamydomonas reinhardtii*	[109]	Aug 2001	Mg^{2+}-CO_2/2CABP	1.40
1IR1	Spinach	[65]	Aug 2001	Mg^{2+}-CO_2/2CABP	1.80
1IR2	*Chlamydomonas reinhardtii*	[65]	Sep 2001	Mg^{2+}-CO_2/2CABP	1.84
1GEH	*Thermococcus kodakaraensis*	[52]	Nov 2001	$SO_4^{2-} \times 2$	2.80
1IWA	*Galdieria partita*	[73]	Apr 2002	SO_4^{2-}	2.60
1UPP	Spinach	[48]	Oct 2003	Ca^{2+}-CO_2/2CABP	2.30
1UPM	Spinach	[48]	Oct 2003	Ca^{2+}-CO_2/2CABP	2.30
1UW9	*Chlamydomonas reinhardtii*	[49]	Feb 2004	Mg^{2+}-CO_2/2CABP	2.05
1UWA	*Chlamydomonas reinhardtii*	[49]	Feb 2004	Mg^{2+}-CO_2/2CABP	2.30
1TEL	*Chlorobium tepidum*		May 2004	–	2.70
1WDD	Rice		May 2004	Mg^{2+}-CO_2/2CABP	1.35
1YKW	*Chlorobium tepidum*	[56]	Jan 2005	–	2.00
1SVD	*Halothiobacillus neapolitanus*		Mar 2005	SO_4^{2-}	1.80
1UZH	*Chlamydomonas reinhardtii*	[46]	May 2005	Mg^{2+}-CO_2/2CABP	2.20
1UZD	*Chlamydomonas reinhardtii*	[46]	May 2005	Mg^{2+}-CO_2/2CABP	2.40
2CXE	*Pyrococcus horikoshii*		June 2005	–	3.00
2CWX	*Pyrococcus horikoshii*		June 2005	–	2.00
2D69	*Pyrococcus horikoshii*		Nov 2005	SO_4^{2-}	1.90
2OEJ	*Geobacillus kaustophilus*	[41]	Dec 2006	PO_4^{3-}	2.55
2OEK	*Geobacillus kaustophilus*	[41]	Dec 2006	Mg^{2+}	1.80
2OEL	*Geobacillus kaustophilus*	[41]	Dec 2006	Mg^{2+}-HCO_3^-	1.80
2OEM	*Geobacillus kaustophilus*	[41]	Dec 2006	Mg^{2+}-2,3-diketohexane 1-P	1.70
2V63	*Chlamydomonas reinhardtii*	[47]	July 2007	Mg^{2+}-CO_2/2CABP V331A mutant	1.80
2V67	*Chlamydomonas reinhardtii*	[47]	July 2007	Mg^{2+}-CO_2/2CABP T342I	2.00
2V68	*Chlamydomonas reinhardtii*	[47]	July 2007	Mg^{2+}-CO_2/2CABP V331A/T342I mutant	2.30
2V69	*Chlamydomonas reinhardtii*	[47]	July 2007	Mg^{2+}-CO_2/2CABP D473E mutant	2.80
2V6A	*Chlamydomonas reinhardtii*	[47]	July 2007	Mg^{2+}-CO_2/2CABP V331A/G344S mutant	1.50
2QYG	*Rhodopseudomonas palustris*	[104]	Aug 2007	–	3.30
2VDH	*Chlamydomonas reinhardtii*	[29]	Oct 2007	Mg^{2+}-CO_2/2CABP C172S mutant	2.30
2VDI	*Chlamydomonas reinhardtii*	[29]	Oct 2007	Mg^{2+}-CO_2/2CABP C192S mutant	2.65

Apart from this group, there are no additional form II sequences associated with either of the other major groups.

The completion of genomic sequencing projects has led to the identification of *rbc*L sequences also in archaea, many of which are anoxygenic. Based on phylogenetic analyses archeal enzymes have been suggested to constitute a different group. This is also congruent with the present analysis, but partly with low support as estimated from bootstrap analyses. This

is obviously an anticipated result because the archaea form a separate kingdom, which is expected to manifest itself as a monophyletic entity in a phylogenetic analysis. This, however, does not necessarily imply that the archaeal Rubisco should be regarded as a separate form (form III) in the same way as e.g. the Solanaceae Rubisco forms a monophyletic entity without meriting the assignment to a separate form. In terms of quaternary structure the archaea are diverse and comprise L2, L8 and L10 enzymes. The crystal structure from *Thermococcus kodakaraensis* [52] features a pentameric structure composed of five L2-units, L10 (Fig. 1E,F). The general fold of the *T. kodakaraensis* Rubisco L-subunit is very similar to the L-subunit of *Rhodospirillum rubrum* Rubisco as is the arrangement of the large subunits in the dimer. The crystal structure of Rubisco from *Pyrococcus horikoshii* consists of an octamer of large subunits, L8 (Table 2). In our phylogenetic analysis, the archaea form a loosely knit group, with *Methanococcoides burtonii* at the base of the form II enzymes. In the former entity we also encounter *Geobacillus kaustophilus*, a bacterium of the Bacillae, which was first isolated from deep-sea sediments of the Mariana Trench. It is a thermophilic species, with an upper temperature limit above +70 °C. From amino acid sequences the similarities between *Geobacillus* and the archaea are difficult to observe, but with the aid of structure data [41] a firm conclusion on alignment can be made. Based on their similar ecological preferences, it is tempting to attribute the affinities to evolutionary convergence due to ecological adaptations.

Geobacillus kaustophilus belongs to a group of organisms that do not use CO_2 as the major source for carbon, but harbour putative Rubisco sequences differing from form I and II sequences [103]. The homologues were termed Rubisco-like proteins, RLPs, (reviewed in ref. [104]) and were assigned as a new form, IV. Members of the form IV subfamily do not catalyse the carboxylation reaction, but have been shown to play a role in sulphur metabolism [36,69]. RLPs from species of Bacilli (*Bacillus subtilis* and *Geobacillus kaustophilus*) have been found to catalyse enolisation of 2,3-diketo-5-methyl-thiopentyl-1-phosphate, a compound with structural similarity to RuBP [7,41]. This is true also for the RLP from *Microcystis aeruginosa* PCC 7806 [10]. A unique feature of *M. aeruginosa* is that it also harbours a gene encoding the form I Rubisco large subunit [10]. Crystal structures (Table 2) have been reported for the RLPs from the green sulphur bacterium *Chlorobium tepidum* [56], from *Geobacillus kaustophilus* [41], and from *Rhodopseudomonas palustris* [104]. From the phylogenetic analysis, it appears that the RLPs form an entity, and are not dispersed among the true Rubiscos, thus indicating a common origin for either Rubisco or RLPs. The order of appearance, however, is not entirely clear on basis of the sequences presently available.

3.2. The fold of Rubisco is conserved

Crystal structures representing all forms of Rubisco, including RLPs, have been determined (Table 2, Fig. 1). Despite apparent differences in amino acid sequence, the secondary structure is extremely well conserved throughout.

3.2.1. The large subunit

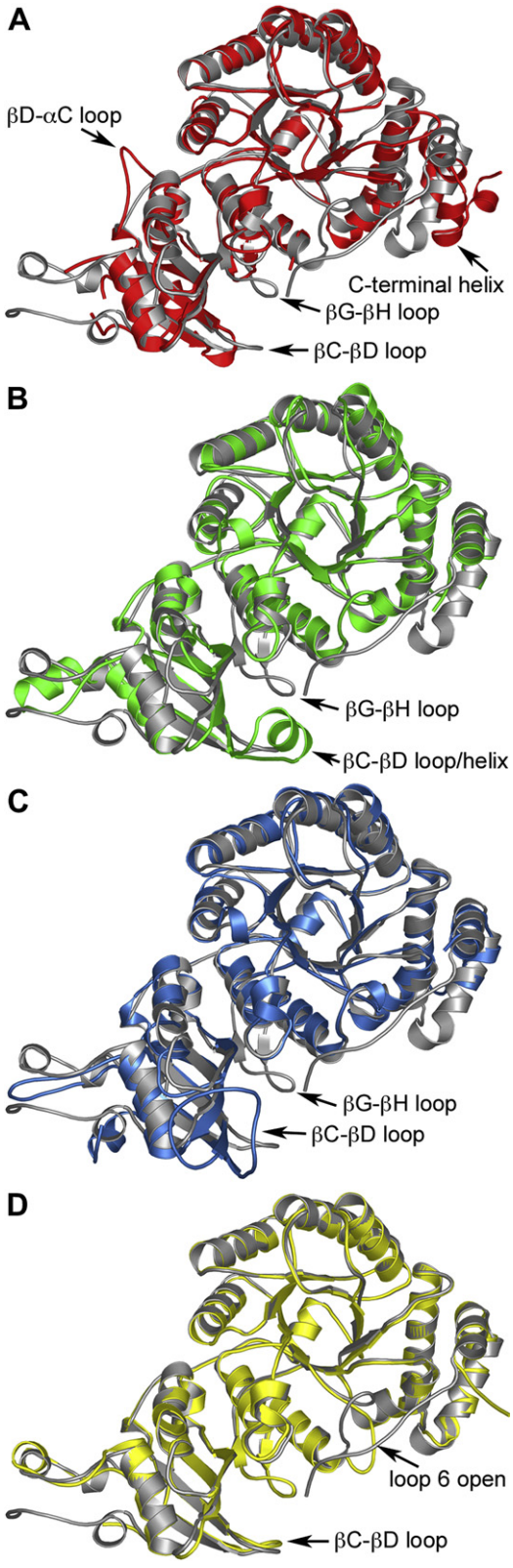
The overall fold of the large (catalytic) subunit is similar in all forms of Rubisco (Fig. 3): a smaller amino-terminal domain consisting of a four-to-five-stranded mixed β sheet with helices on one side of the sheet and a larger carboxy-terminal domain (reviewed in ref. [4]). The carboxy-terminal domain consists of eight consecutive βα-units arranged as an eight-stranded parallel α/β barrel structure. The active site is located at the carboxy-terminal end of the β-strands, with the loops connecting the βα-units contributing several residues involved in catalysis and substrate binding. Residues from the amino-terminal domain of the adjacent large subunit in the dimer complete the active site. Thus, the functional unit structure of Rubisco is an L2 dimer of large subunits harbouring two active sites (Fig. 1).

The secondary structure of the large subunit is well conserved (Fig. 3). One difference is the presence of an extra α-helix in the amino-terminal domain in the structure of Rubisco from *Rhodospirillum rubrum* (helix αA, residues 14–19, [86,88]). In addition, certain loop regions differ in length (Fig. 3, [86]).

Large subunits of RLPs display ca 30% sequence identity with form I–III Rubisco enzymes. The most significant difference in the structure of the RLPs from *Chlorobium tepidum* [56] and *Rhodopseudomonas palustris* [104] is the insertion of a 14-residue loop between β-strand C and D in the amino-terminal domain (βC–βD loop, Fig. 3C). In the RLP from *Geobacillus kaustophilus* this loop is also longer than in form I enzymes, but assumes a helical conformation (Fig. 3B) [41]. The shorter loop in plant and algal form I Rubisco enzymes interacts with Rubisco activase [74]. A second major difference is the absence in RLPs of a β-hairpin loop (βG–βH loop) between α-helix 6 and β-strand 7 at the bottom of the α/β-barrel (Fig. 3). This is also the case in the form III Rubisco from the archaeon *Pyrococcus horikoshii* (Table 2), and according to sequences also in *Methanospirillum hungatei*, *Methanosaeta thermophila*, *Chlorobium tepidum*, and *Geobacillus kaustophilus*.

3.2.2. The small subunit

Whereas the large subunits display relatively small variations in the different forms, the small subunit is more diverse. The common core structure consists of a four-stranded anti-parallel β-sheet covered on one side by two helices [54]. The most striking variations occur in two distinct locations, the loop between β strands A and B of the small subunit, the so-called βA–βB-loop, and the carboxy-terminus (reviewed in ref. [97]). The βA–βB loops of four small subunits line the openings of the solvent channel that traverses the holoenzyme (Fig. 4). Rubisco from prokaryotes and non-green algae have only ten residues in the loop (Fig. 4A) as illustrated by the structure of the cyanobacterial enzyme [71], but Rubisco from higher plants have 22 (Fig. 4B, [54]) and green algal Rubisco have 28 (Fig. 4C, [109]). Non-green algae and



some prokaryotes, which have ten residues in their βA−βB loops, display carboxy-terminal extensions that form β hairpin structures (βE−βF loop) in the spaces that are normally occupied by the longer βA−βB loops of the green algal and plant enzymes [35,101]. Four β-hairpin structures form a central β-barrel at the entrance to the central solvent channel (Fig. 4D). These enzymes also display a slightly longer loop between β-strands C and D. The small subunits of green algae also have longer carboxy-termini than those of higher plants, but these do not form β-hairpins [65,109] and do not appear to be essential to the function of the enzyme [97].

Judging by the arrangement of the small subunits covering a substantial area at two opposite ends of the L-subunit octamer, it is reasonable to assume a structural function of the small subunit, namely to assemble and concentrate the large catalytic subunits. However, considering that some Rubisco enzymes lack small subunits and have the lowest specificity values [44], it is tempting to speculate that the small subunits also contribute substantially to the differences in kinetic properties observed among different Rubisco enzymes.

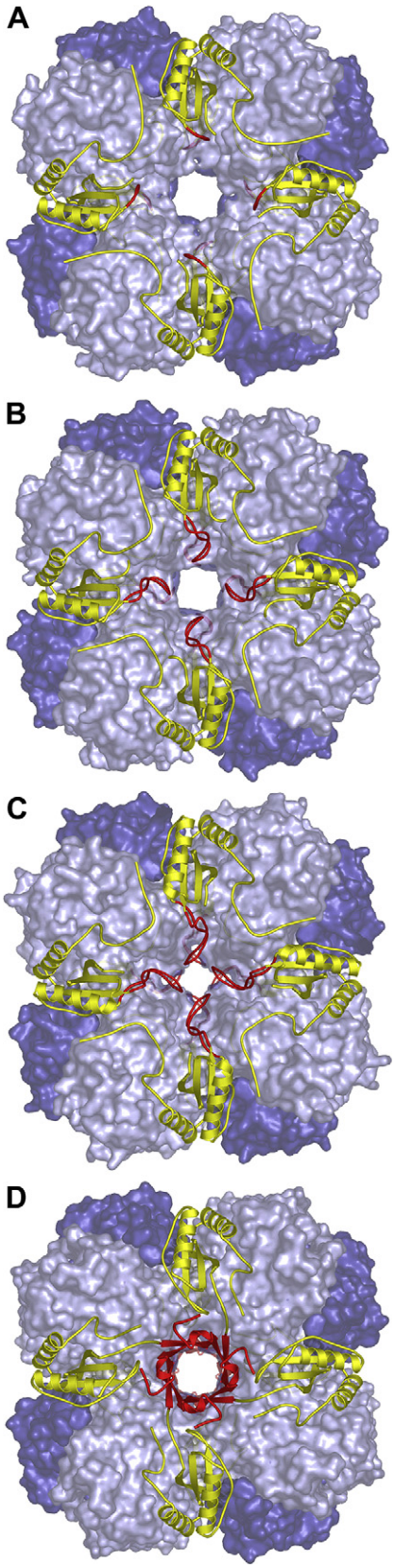
3.3. Sequence variation and phylogenetic relationships

Based on structure comparison the amino acid sequences of Rubisco large and small subunits have been aligned and analysed to provide a phylogenetic tree (Fig. 2), which can serve as a basis for discussions on evolution and distribution of traits, forms, and types among photosynthesizing organisms. Details on the actual analyses are presented below.

Although partly very well supported, it is obvious that the analysis of only two molecular markers, as these presented here, has its limitations. Due to restrictions of both the alignment itself, the nature of data with a restricted number of states, substitution asymmetries and dependencies, and the methods of analysis, it is only possible to provide a hypothesis that is valid within a restricted time window (e.g. refs. [1,9]). The width of this window depends on the substitution rates of the sequence fragments under analysis, which are by no means constant.

Early on, it was recognised that even though the *rbc*L sequence is highly conserved in plants, there were obvious differences between different parts of the gene − both with regard to codon position and sequence regions. In 1994 Albert and co-workers ([1] and references therein) studied the functional constraints on *rbc*L evolution, and possible effects of these on phylogenetic reconstruction. A similar study was published in 1997 [50], with largely congruent results. Recently Kapralov and Filatov [45] presented evidence from maximum likelihood and Bayesian modelling, which point

Fig. 3. Comparison of the large subunit from spinach, in grey, with the large subunits of (A) *Rhodospirillum rubrum* in red (PDB code 5RUB), (B) *Geobacillus kaustophilus* in green (2OEM), (C) *Rhodopseudomonas palustris* in blue (2QYG), and (D) *Halothiobacillus neapolitanus* in yellow (1SVD). Features that differ are indicated by arrows. See ref. [86] for assignment of secondary structure elements. Note that helices 7 and 8 of the α/β-barrel are displaced [86] in the enzyme from *Rhodospirillum rubrum* (A), but in the RLPs this region superimposes with the same region in the spinach enzyme (B, C).



to a positive selection on the *rbc*L gene. Although these and other separate studies point in a similar direction, a detailed analysis and interpretation of the patterns eludes us.

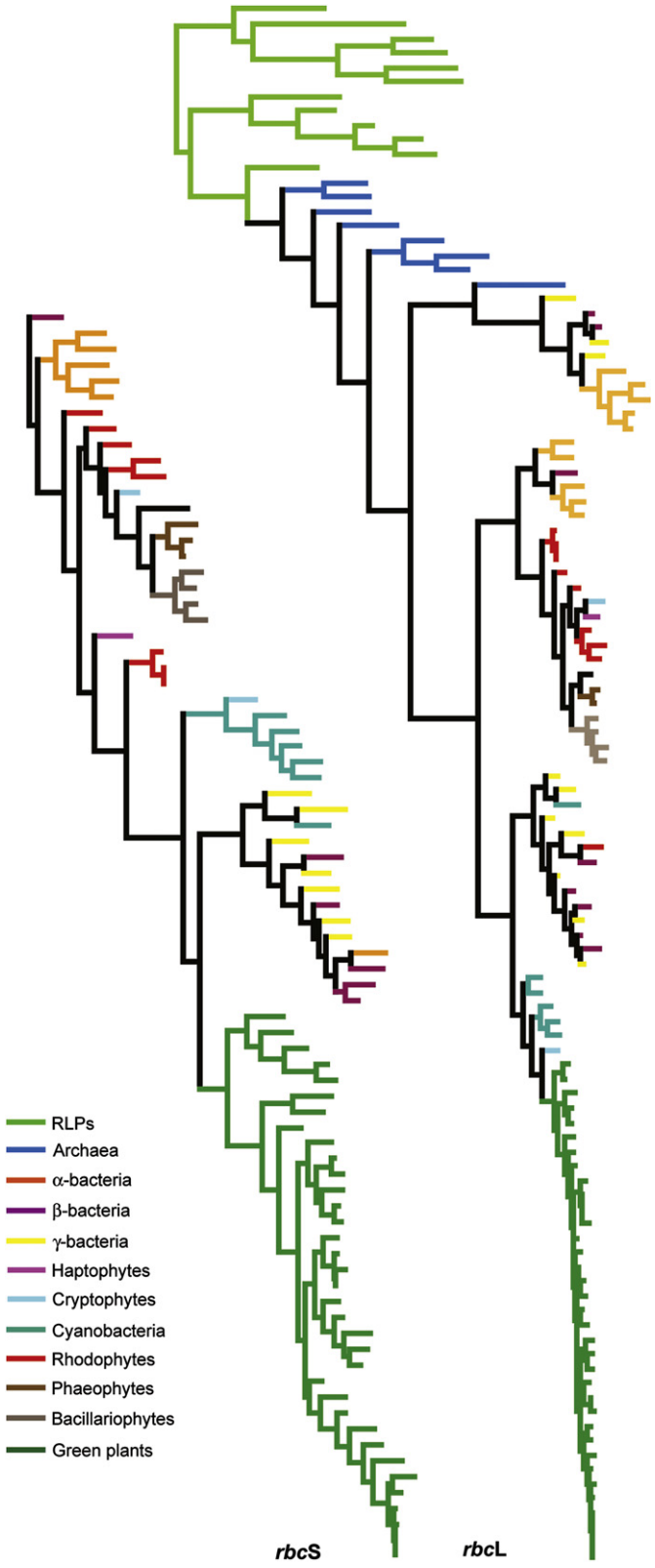
In Fig. 5, a summary of the evolutionary relationships of the *rbc*L and *rbc*S genes is presented. These tree diagrams are based on separate analyses of both genes, but are generally congruent with results from the combined analysis in Fig. 2, with the exception for minor, poorly supported, differences. Several unexpected and interesting features can be deduced from the tree, and it can also be used as a basis for understanding the distribution of the different enzyme forms, and for interpreting anomalous features of the sequences or structures.

The tree diagram is rooted using the RLPs, which are assumed to form a separate entity from the functional Rubiscos. Basal to the RLPs, forming a link with affinities both to these and the form II Rubiscos, we encounter the sequences of archaeal origin. With the tree rooted between Rubiscos and RLPs, the most plesiomorphic group among the Rubiscos are the form II enzymes. Alignment similarities, and as a consequence support, in this group are overall low.

The remainder of the tree comprises four large and well supported groups: i) α, β, and γ-bacteria with only large subunits; ii) α-proteobacteria and red algae; iii) α, β, γ, and cyanobacteria with several Rubisco operons; and iv) cyanobacteria and green plants. One interpretation of this division is that the chloroplasts of green plants and those of red algae and their relatives have different evolutionary origins. In both cases, the ancestry would be bacterial, but from different groups of bacteria or at least forming different lineages. This has earlier been suggested by e.g. Glöcker and co-workers after the complete sequencing of *Cyanidium caldarium* [33]. In the lineage with red algae, we also find e.g. haptophytes, xanthophytes, phaeophytes, and bacillariophytes, in congruence with most recent classifications. All of these taxa, including the α-proteobacteria, possess a unique sequential and structural feature, an extended carboxy-terminus of the small subunit (see also Fig. 4). The strong grouping here was, however, achieved also from an analysis excluding that portion of the sequences.

In the group of bacteria of mixed origins, we find all bacteria harbouring two different *rbc*L and *rbc*S genes (sometimes annotated as *cbb*L and *cbb*S or *cfx*L and *cfx*S) – with the sole exception of a *Synechococcus* type I gene placed at the base of the cyanobacteria and green plants group. The evolutionary significance of this arrangement is unclear, but it can be noted that all sequences of type I (resulting in enzymes of form I with an L8S8 arrangement) form one entity, while those of type II (also resulting in enzymes of form I with only a L8S8 arrangement) form a grade basal to the type I sequences.

Fig. 4. Variation of the βA–βB loop and the carboxy-terminus of the small subunit in various Rubisco enzymes. Large subunits are coloured blue, small subunits are yellow, and the βA–βB loops are red. A) *Synechococcus* PCC 6301 enzyme, PDB code 1RBL. B) Spinach enzyme, PDB code 8RUC. C) *Chlamydomonas reinhardtii* enzyme, PDB code 1GK8. D) *Galdieria partita* enzyme, PDB code 1BWV.



In the third, largest, group with cyanobacteria and green plants, there are no unexpected surprises. The euglenophytes, represented by the sequences from *Euglena gracilis*, are placed in a lineage together with the green algae. The close relationship between the genes of their photosynthesizing enzymes has been reported previously, and is not assumed to reflect the true phylogenetic relationships between these taxa.

The relative substitution rates in *rbc*S/L (Fig. 5) can be estimated in different lineages. Large groups of organisms have been colour coded, and both subtrees have been scaled to provide comparable branch lengths. The top section of the *rbc*L subtree lacks a corresponding portion in the *rbc*S tree, because small subunits are missing for the RLPs and other form II taxa such as the archaea. One notable difference is seen in the green plant branch, where the *rbc*-operon is lost when the *rbc*S gene is transferred to the nucleus. From this point the substitution rate of *rbc*L decreases significantly, indicating a strongly conserving force. The *rbc*S gene, however, is duplicated at several instances after the transfer. Presumably, this offers much larger possibilities to cope with deleterious substitutions, and hence this group is able to retain a comparably high substitution rate.

3.4. A rigid and conserved framework for the active site

The active site is located at the intra-dimer interface between the carboxy-terminal domain of one large subunit and the amino-terminal domain of the second large subunit in the L2 dimer (Fig. 1). In the L8S8 hexadecamer, the four L2 dimers are arranged in such a way that the entrances to the eight active sites face the outside of the hexadecameric molecule [3,54]).

Productive binding of the substrate requires prior activation of Rubisco. This process involves the carbamylation of an active-site lysine residue and the subsequent stabilisation of the lysyl-carbamate by a magnesium ion [57]. Binding of the substrate or inhibitors to the non-activated enzyme locks the enzyme in a closed unproductive form. Plants require the assistance of Rubisco activase to facilitate the release of tightly bound sugar phosphates (including the binding of the natural inhibitor 2-carboxyarabinitol 1-phosphate to the activated form of the enzyme) in order to restore activity (reviewed in ref. [79]).

The carboxy-terminal 8-stranded α/β-barrel of the large subunit forms a rigid framework for the active site. The loops connecting the β-strands of the barrel with the following α-helix contribute invariant residues that interact with the substrate [3,4,54]. Residues in loops 1, 2 and 5–8 (connecting

Fig. 5. Phylograms from separate analyses of *rbc*S (left) and *rbc*L (right) sections of alignment, with sample trees of 63 and 432 equally parsimonious trees, respectively. Branch lengths proportional to substitution rates are calculated as number of substitutions per length of each section. Both sections show comparable rates, with the exception of the green plants, where the *rbc*S-L operon is lost, and *rbc*S is transferred from the chloroplast to the nucleus followed by subsequent duplications. For RLPs and archaeal taxa, as well as some bacteria, no corresponding *rbc*S sequences are available, hence resulting in a phylogram with fewer terminal taxa.

β-strand 1 with α-helix 1, etc.) interact directly with the substrate or the metal ion. Two loop regions in the amino-terminal domain of the second large subunit in the dimer contribute additional residues to the active site. The majority of the residues lining the active site are charged or polar, reflecting the charged nature of the substrate [3]. The substrate binds in an extended conformation across the opening of the α/β-barrel and is anchored at two distinct phosphate-binding sites at opposite sides of the α/β-barrel and in the middle at the magnesium-binding site [3,54,107].

3.5. Conformational changes during catalysis

The loop connecting β-strand 6 with α-helix 6 in the carboxy-terminal α/β-barrel is important for catalysis and specificity [13]. The sequence of loop 6 is well conserved among higher plants and green algae. Apart from Gly333 and Gly337, which are likely conserved to maintain flexibility in the hinge of the loop, one residue, Lys334, is strictly conserved within this group. Lys334 is located at the tip of the loop and interacts with the incoming gaseous substrate during catalysis. The Lys334 side chain extends into the active site and hydrogen bonds to the two oxygen atoms of the inhibitor 2-carboxyarabinitol-1,5-bisphosphate (2CABP) that are equivalent to those of substrate CO_2 [2,54]. It also makes hydrogen bonds with the γ-carboxylate of Glu60 and the hydroxyl group of Thr65 in the amino-terminal domain of the adjacent large subunit. Thr65 is conserved in all sequences studied except in *Chlorobium tepidum* and *Rhodopseudomonas palustris*, where it is replaced by Gln, and in ca 50% of the archaea where it is replaced by Trp.

Loop 6 is flexible, and as a part of the catalytic mechanism, movement of loop 6 partitions Rubisco structures into two states, open or closed [20]. The open state is associated with an unliganded active site or an active site occupied by loosely bound ligands (substrates or products). The closed state, in contrast, is associated with tightly bound substrates or inhibitors that are completely shielded from solvent and is a prerequisite for catalysis [20,106]. Apart from the movement of loop 6 (residues 331–338) to cover the opening of the α/β-barrel, the transition between open and closed active sites involves a rigid-body movement that brings the amino- and carboxy-terminal domains of adjacent subunits together, and the ordering of residues 63–69 of the amino-terminal domain. Packing of the carboxy-terminal strand (residues Trp463 to the carboxy-terminal end) against loop 6 completes the closure [89,106].

3.6. The active site

The prime reaction catalyzed by Rubisco (the addition of CO_2 and H_2O to RuBP to yield two molecules of 3-phosphoglycerate) involves multiple discrete steps and associated intermediates of variable stability (reviewed in refs. [15,37]). Crystal structures of complexes of the enzyme with substrate, product, or inhibitors have provided valuable information on the catalytic reaction (Table 2). The inhibitor 2CABP mimics the six-carbon carboxylated reaction intermediate 3-keto-2′-carboxyarabinitol 1,5-bisphosphate and forms an exchange-resistant complex ideal for crystallographic comparisons [2,54,65,71,90,101,109]. Complexes with ligands that bind less tightly, such as the natural substrate RuBP or the product 3-phosphoglycerate, or mimics thereof, have turned out to be more difficult to crystallise. Only in a few cases has it been possible to capture these complexes in the crystal [58,59,72,107,108,110]. To obtain a stable complex with RuBP, Ca^{2+} was used instead of Mg^{2+} [107], because the former ion does not support turnover to an appreciable extent [77]. The resulting structure, obtained by soaking of RuBP into crystals of the carbamylated spinach enzyme in the presence of Ca^{2+} is open, with loop 6 partially retracted. It appears that the presence of RuBP only is not enough to trigger the closing of the active site, because the complex of 2CABP with Ca^{2+} is of the closed type [48], indicating that it is not the calcium ion per se that prevents the conformational switch. It may be that the interaction of the substrate-CO_2 with Lys334 is required for proper loop closure. Lys334 is conserved in all organisms with the exception of the archaea, and those bacterial sequences associated with the archaea or RLP groups of the tree diagram in Fig. 2. Note however that a K334M mutant of *Synechococcus* PCC6301 still has all loops closed over 2CABP [71]. The question arises whether the striking similarity of the active sites in diverse complexes with 2CABP is at least to some extent induced as a result of the extraordinary stability of these complexes. Structural studies of a complex with a stable analogue of the endiol intermediate may provide some answers to this and other questions concerning the reactions catalysed by Rubisco, but such a mimic is yet to be found.

Computational methods have been used to probe the energetics of the reaction. Based on the 1.6 Å crystal structure of the activated complex of *Spinacia* Rubisco with 2CABP [2,54] and to a lesser extent on the activated complex with the substrate RuBP [107], quantum chemical calculations were performed in vacuo or using an abbreviated model of the substrate and the active site (see e.g. refs. [51,64,105]). While these calculations differed in certain details, they consistently identified the carbamylated Lys201 as a key player in catalysis. This lysine residue is conserved in all Rubiscos and RLPs studied here. Other residues are implicated in catalysis (e.g. Lys175, Lys334 and His294), although the conclusions concerning the role of these residues in catalysis are less straightforward. It is clear that further studies are needed before a unified and consistent mechanism can be derived.

Similar to form I–III Rubisco enzymes, the active site in RLPs is located at the interface of the two subunits sharing some (about half) of the mechanistically significant active-site residues found in authentic Rubisco enzymes [41,56,104]. RLPs contain the same Mg^{2+}-binding motif KDDE, with the flavours KDDH in *Rhodopseudomonas palustris* RLP1 and *Oceanicola granulosus* RLP1, KEDH in *Oceanicola granulosus* RLP2 and KNDE in all form II bacterial sequences. In the structure of the 'enolase' from

G. kaustophilus with the stable alternate substrate, 2,3-diketohexane 1-phosphate, Lys173, the homologue of K201 in the form I enzymes, is carbamylated [41]. Consistent with the different regiochemistry of the tautomerisation reaction catalysed by this enzyme, a different general base is identified in the *G. kaustophilus* 'enolase' [41]. The 1-phosphate binding site (P1 site) appears to be conserved in RLPs, but the P2 binding site is altered and seems most suited to bind a small hydrophobic group [41,56], in line with the fact that these enzymes do not catalyse RuBP carboxylation. The functionality of loop 6, participating in the transition between open and closed states appears to be conserved, whereas the residue at the tip of the loop, Lys334 in form I enzymes, is not [41,56,104]. In ca 50% of the RLPs studied this residue is replaced by Gly, Arg or Ser.

4. Structure and function

The key role played by Rubisco in carbon fixation makes it an obvious target for genetic manipulation with the aim to improve its catalytic properties (reviewed in ref. [75]). However, because of the size and complexity of Rubisco, it is not yet known which residues may improve CO_2/O_2 specificity or carboxylation catalytic efficiency, or account for differences in specificity among Rubisco enzyme from divergent species. Genetic engineering is an excellent method to examine the role of individual residues in catalysis, but the engineering of Rubisco enzymes from higher plants has been difficult due to several complications. Whereas the large subunit of the plant enzyme is coded by a single gene in the chloroplast genome, the small subunit is coded by a family of closely related nuclear genes (reviewed in ref. [97]). Advances in the genetic transformation of the genome of tobacco chloroplasts and more recently several other plant plastids (reviewed in ref. [62]) have opened up for molecular manipulation of higher-plant Rubisco [6,112]. However, hitherto unknown requirements of holoenzyme folding and assembly still preclude genetic manipulation of plant Rubisco in heterologous hosts (reviewed in ref. [6]), resulting in aggregated and misfolded protein. In contrast, bacterial Rubisco enzymes are readily expressed and assembled in *Escherichia coli* ([30] reviewed in refs. [37,100]).

Genetic screening and selection in *Chlamydomonas* has enabled the identification and recovery of mutants based on a photosynthesis-deficient phenotype because photosynthesis-deficient mutants can survive on a supply of acetate as a source of carbon [95]. Mutants can then be used to select for second-site mutations that restore Rubisco function and photosynthesis, allowing the identification of residues that influence catalysis in both large and small subunits (reviewed in ref. [100]). A system has also been designed for selection of randomly mutagenised Rubisco genes of the cyanobacterium *Synechococcus* PCC6301 with the photosynthetic bacterium *Rhodobacter capsulatus* as a host [93].

While this has allowed the identification of residues that influence catalysis in both large and small subunits we are far from a general understanding of what structural features influence catalysis and specificity of Rubisco. Most substitutions have a negative effect on specificity and even if a positive effect is achieved, this often influences the overall catalytic rate in a negative way. Also, the results obtained from one organism can often not be generalised to another organism, because differences at one particular site are likely to be complemented by substitutions at another site (e.g. refs. [81,99]). Nevertheless, a number of residues or groups of residues have been identified that influence catalysis and specificity. From the discussion in Section 2 it is obvious that the specificity factor is but one parameter that determines the net efficiency of Rubisco enzymes, but it can serve as an important first diagnostic parameter to indicate changes in efficiency of engineered Rubisco enzymes. It has been argued [111] that the majority of Rubiscos have been optimised in terms of specificity and catalytic turnover rates to the environment where the host organism resides. Considering the relatively modest number of kinetic data available from Rubiscos from diverse organisms to date, such a conclusion may turn out to be premature and warrants further studies.

4.1. Structural determinants of catalysis and stability

Numerous large- and small-subunit mutants have been described in the literature and several reviews have been dedicated to the analysis of site-directed mutants [37,75,95,97,100]. Reiteration of these data is not the purpose of the present article. However, detailed analysis of X-ray crystal structures of mutant Rubisco enzymes can provide additional information on interactions that are important for e.g. holoenzyme stability and catalysis and these are reviewed below. A remaining issue concerns the generality of these observations. In some cases the mutated residues are conserved in several organisms. In other cases, the residues in question are only conserved in closely related taxa. In these latter cases, alternative solutions may have evolved.

4.1.1. Large subunit, loop 6

Residue Val331 forms part of the hinge on which loop 6 moves (see Section 3.5) and is highly, but not strictly, conserved [71]. Replacement of Val331 by Ala in the green alga *Chlamydomonas reinhardtii* [13] caused considerable reduction in specificity and carboxylation turnover. Genetic selection for restored photosynthesis identified second-site T342I and G344S substitutions with increased specificity and modest improvement in V_c/K_c [13,14]. Similar results were obtained in *Synechococcus* Rubisco [34,76].

The crystal structure of the mutant V331A was analysed [47] along with the structures of the V331A/T342I- and V331A/G344S-revertant enzymes and the T342I- and G344S-suppressor enzymes (which contain wild-type Val331). In general, the mutations caused local, but significant deviations close to the substituted residues, but left the Cα backbone of the mutant enzymes unperturbed. The side chain of Val331 is situated in a hydrophobic pocket and makes van der Waals interactions with residues Thr342, Ile393, and the main-chain Cα atom of Arg339. Substitution of the valine

side chain by the smaller alanine in the V331A-mutant enzyme weakened these interactions considerably and created a small cavity that was partly filled by solvent. In the V331A/T342I-revertant enzyme, van der Waals contact was restored between residues 331 and 342. Whereas the main-chain conformation of the loop was not visibly altered by the mutation, the presence of the additional Ile342 Cγ atom caused changes in local conformations and interactions of several residues, e.g. Arg339, Glu338, and Ser328 close to I342. Because residues 331 and 342 flank Lys334 located at the apex of loop 6, it seems likely that the mutation may destabilise the loop or alter its flexibility, thereby influencing the catalytic properties of the mutant enzymes. Similar conclusions were drawn from the observation of altered catalysis in the tobacco enzyme induced by a L335V mutation, which caused a reduction in specificity and altered sensitivity to inhibitors [78,112]. In the V331A/G344S-revertant enzyme, the substitution of Gly344 by serine caused a displacement of the entire large-subunit α-helix 6 (residues 338–347), and resulted in a shift in backbone coordinates for these residues of approximately 1 Å [47]. The effect of this displacement is to bring Thr342 closer to Ala331. It thus appears that specificity is restored by two fundamentally different mechanisms in the two different revertants. The presence of Val331 in the G344S single mutant may limit the displacement of α-helix 6, and may account for the different catalytic properties of the V331A/G344S-revertant and G344S-suppressor enzymes.

Interpreting sequence variation in the 104 taxa *rbc*L + *rbc*S matrix, it appears that Val331 and Thr342 are introduced in the form I-genes along the branch leading from archaea to the more apomorphic parts of the tree. Val331 and the entire motif VVGKLEG are conserved in all eukaryotes, as well as in the majority of α, β, and γ-bacteria included in our alignments, with the notable exception of *Xanthobacter*, *Ralstonia*, *Sinorhizobium* and *Rhodobacter* where the motif reads AVGKLEG. In the genes giving rise to form II enzymes, however, the corresponding motif reads MGY/FGKME. With regard to Thr342, this position is further modified to Ile or Val in two groups. One consists of the same four genera which show V331A, described above. The other group comprises a set of taxa belonging to the Bacillariophyta-Phaeophyta clade, including some, but not all, members of Rhodophyta. This evolutionary coinciding, and putatively related, establishment of Val331 and Thr342, and their simultaneous change to Ala331 and Ile/Val342, support the observations from mutant studies of a dependence between these two positions. In genes giving rise to form II enzymes, this position is also conserved, but as a Gly.

4.1.2. Large subunit, carboxy-terminus

As mentioned above, apart from the closure of loop 6, movement of the carboxy-terminus against loop 6 also seems to be intimately involved in the transition from the open to the closed state of the Rubisco active site [20,89,106]. Results from site-directed mutagenesis indicate that although the carboxy-terminus is not absolutely required for catalysis, it is important for maximal activity and stability [22,34,66,82]. In particular, Asp473 was proposed to serve as a latch responsible for placing the large-subunit carboxy-terminus over loop 6, and stabilizing the closed conformation required for catalysis [20]. Directed mutagenesis and chloroplast transformation in *Chlamydomonas reinhardtii* showed that although Asp473 is not essential for catalysis, mutations D473A and D473E caused substantial decreases in catalytic efficiency and specificity [85]. Analysis of the crystal structure of D473E [47] revealed that the relatively modest substitution of Asp473 by Glu causes disorder of the last six carboxy-terminal residues. The most likely conclusion is that the disruption of the contacts of residue 473 (with Arg134 and His310) in the mutant enzymes causes destabilisation of the under-lying loop 6.

4.1.3. Interface between large and small subunits and the bottom of the α/β-barrel

Mutant screening and selection in *Chlamydomonas reinhardtii* identified a mutant Rubisco with alterations in stability and kinetic parameters that were caused by a Leu290 to Phe (L290F) substitution in the chloroplast encoded large subunit [11]. The L290F enzyme had a 13% decrease in CO_2/O_2 specificity and reduced thermal stability *in vivo* and *in vitro* [11,12]. Second-site suppressor substitutions A222T and V262L restores photosynthesis of the L290F mutant at the restrictive temperature [39] and the L290F/A222T and L290F/V262L double-mutant enzymes have improved thermal stability properties and specificity values close to the levels of the wild-type enzyme [19,39]. Analysis of the crystal structures of wild-type [109] and mutant enzymes [49] showed that these substitutions do not perturb the structure significantly besides introducing a different amino acid side-chain. Residue 290 is located at the beginning of β-strand 5 at the bottom of the α/β-barrel (i.e. away from the active site) with its Cα 18 Å from the active-site magnesium ion, suggesting that the L290F substitution must affect the relative rates of carboxylation and oxygenation in an indirect way. Residue 222 is positioned roughly in the middle of α-helix 2 of the α/β barrel. It is located almost midway between residue 290 in the same large subunit and residue 290 in the nearest neighbouring large subunit (at 15 and 19 Å distances, respectively) and 20 Å away from the active site magnesium ion. Thus the effect of the A222T suppressor substitution must also be indirect. Although no structure was obtained for the V262L mutant enzyme, analysis of the wild-type enzyme structure [109] indicated that the same must hold for V262L.

These long-range interactions indicated that protein dynamics in the mutants may be affected and this conclusion is underscored by the temperature-sensitive phenotype of the L290F mutant. Close scrutiny of the structures [49] revealed that the mutant residues interact via van der Waals contacts within the same large subunit and also via a path involving a neighbouring small subunit. In addition, distinct regions were identified at the subunit interface and around the central solvent channel that showed significant and systematically increased atomic temperature factors in the L290F mutant enzyme compared to wild-type. These regions coincide with residues on the interaction paths between mutant and suppressor

sites and may thus explain the temperature-conditional phenotype of the L290F mutant strain. This suggests alterations in subunit interactions may influence protein dynamics, and thereby affect catalysis [49]. Residues Ala222, Val262 and Leu290 are conserved in green algae and higher plants. A mutant of *Nicotiana tabacum* with a G322A substitution in the large subunit of Rubisco was found to lack holoenzyme [91]. Gly 322 is located in a loop at the amino-terminal end of the α/β-barrel and presumably the substitution prevents holoenzyme assembly. Substitution of residues on the interface between large and small subunits [18,99] influence catalysis and specificity.

4.1.4. Small subunit, βA–βB-loop

The βA–βB loop of the small subunit is the most divergent structural feature of Rubisco enzymes. Numerous experiments have shown that specific interactions of the βA–βB loop can influence holoenzyme stability and catalytic performance ([98], reviewed in ref. [97]). Directed mutagenesis and chloroplast import showed that an R53E substitution in the pea βA–βB loop was particularly detrimental to holoenzyme assembly [27]. Based on sequence alignment and directed mutagenesis, Arg59 of the *Chlamydomonas* enzyme was thought to be homologous with Arg53 in plant Rubisco. When the R59E mutant was created in *Chlamydomonas reinhardtii*, Rubisco also assembled, albeit at a lower level and with an associated decrease in thermal stability [98]. When the X-ray crystal structure of *Chlamydomonas reinhardtii* Rubisco was solved [109], it became clear that the backbone atoms of Asn54 in the *Chlamydomonas* enzyme align with those of Arg53 in the *Spinacia* enzyme and that instead the guanidino group of Arg59 from a neighboring small subunit of *Chlamydomonas* Rubisco is in perfect structural alignment with the guanidino group of Arg53 in the same small subunit of the spinach enzyme [109]. This illustrates the value of structural information for sequence alignment.

Transformation of the green alga *Chlamydomonas reinhardtii* with chimeric small subunits where the 28-residue *Chlamydomonas* loop was replaced with the shorter loop sequences of *Spinacia* and *Synechococcus* [46] resulted in the recovery of photosynthesis-competent colonies. However, when the growth phenotypes of the transformants were assessed, the short-loop transformants failed to grow on minimal medium in the light at 35 °C. Thus, although the nature of the βA–βB loop sequence is not essential for holoenzyme assembly, it can influence holoenzyme stability *in vivo* and *in vitro*. X-ray crystal structures of the engineered proteins [46] revealed remarkable similarity between the introduced βA–βB loops and the respective loops in the *Synechococcus* and spinach enzymes. Analysis of the kinetic constant showed that the *Synechococcus* loop caused decreases in carboxylation V_{max}, $K_m(O_2)$, and CO_2/O_2 specificity whereas the construct with the spinach loop caused complementary decreases in carboxylation V_{max}, $K_m(O_2)$, and $K_m(CO_2)$ without a change in specificity. Such changes may arise as a result of differences in the extent and character of the interaction-area between the small-subunit βA–βB loop and the adjacent large subunit in the chimeric enzymes [46]. Also loop dynamics is likely to change in the chimera.

4.1.5. Residues close to the active site

When probing the catalytic reaction by mutagenesis, substitution of residues directly involved in the binding of substrates or activator metal ion is of limited value, because the substitution abolishes overall carboxylase activity [37]. In contrast, residue Cys172 is located in β-strand 1 of the α/β barrel and is close to the active site but does not interact directly with the substrate. Substitution of either of the flanking residues, Gly171 or Thr173, in *Chlamydomonas reinhardtii* Rubisco eliminates carboxylase activity [95], whereas the substitution of Cys172 to Ser has several interesting consequences. It increases resistance to stress-induced degradation in vivo and in vitro [67], this was also observed with a C172A mutant enzyme [63]. However, the C172S substitution also results in a modest but significant (11%) increase in the specificity factor compared to the wild-type enzyme but leaves V_{max} for carboxylation unchanged [29]. It is interesting to note that the predicted coupling of an increased specificity and a decreased carboxylation turnover rate [111] does not seem to hold here.

Analysis of the crystal structure of the C172S mutant Rubisco showed that the substitution to Ser at position 172 causes a shift of the main chain backbone atoms of β-strand 1 of the α/β-barrel. This is likely to result in shifts of the position of individual residues in the active site. In particular Thr173, which interacts with the C2-hydroxyl of 2CABP, or Lys175, which has been suggested play a dual role in catalysis as an acid, as a base, or both [15,51,64,107], see above, are likely to be influenced. Perhaps more importantly, the shift in main-chain residues is likely to influence dynamics of the active site, thereby affecting catalysis.

Considering Cys172 is conserved in green algae and higher plants, the question arises why it has not been substituted by Ser during evolution for the gain of substrate specificity and oxidative resistance? As pointed out by Garcia-Murria et al. [29] the conservation of Cys172 may be part of a general design of redox control of CO_2 fixation and Rubisco turnover.

5. Concluding remarks

Rubisco is not only the most abundant protein, it has also been thoroughly examined. More than 3500 *rbc*L sequences are presently available from public databases for the purpose of phylogenetic reconstruction of photosynthetic lineages. Numerous crystals structures of Rubisco from diverse origins, including site-directed mutants, have been determined and form a basis for attempts to understand functional details. In this study we have attempted to provide a link between these two fields by mapping the variation of the Rubisco large and small subunits onto the three-dimensional structure. The amino acid sequences of Rubisco large and small subunits have been aligned based on a comparison of currently available structures and analysed to provide a phylogenetic tree,

Table 3
Summary of statistics from phylogenetic analysis

Alignment	Characters	Taxa	No. of EPT[a]	Tree-length	CI[b]	RI[c]
*rbc*L + *rbc*S	621	109	58	9069	0.474	0.362
*rbc*L	500	108	432	7475	0.547	0.746
*rbc*S	139	79	63	1704	0.456	0.696

[a] EPT, Equally Parsimonious Trees.
[b] CI, Consistency Index.
[c] RI, Retention Index.

which can serve as a basis for discussions on molecular evolution. Recent results from structural studies are discussed in the light of this analysis.

6. Methods section

6.1. Structure based sequence comparison and phylogenetic analysis

Amino acid sequences of large (*rbc*L) and small (*rbc*S) subunits of Rubisco and Rubisco-like proteins (RLPs) for a biologically diverse set of organisms were compiled from NCBI public domain databases, summarized in Table 1. Amino acid sequence alignment was performed according to the principles outlined above, following a structure based principle rather than the common-place substitution based methods. The reference sequences were aligned on the basis of the corresponding PDB structures. The complete alignment consisted of 109 taxa and a total of 711 characters, of which as many as 621 were parsimony informative. The complete alignment as a NEXUS [61] file, is available from the authors on request.

Three separate phylogenetic analyses, with their corresponding bootstrap analyses, were performed under the assumption of Fitch parsimony [26] with the PAUP* software package [102] on a 8*3 GHz Apple PowerMac. Primary analyses were made on the *rbc*L, *rbc*S and combined *rbc*L + S sequence fragments, respectively. Each round with 1000 random sequence addition replicates, followed by TBR branch swapping, saving all most parsimonious trees. The phylogenetic analyses were followed by bootstrap analysis [24], with 1000 bootstrap replicates, each replicate with three random addition sequences followed by SPR branch swapping. A summary of analysis statistics including retention index (RI, [23]) and consistency index (CI, [53]), are presented in Table 3.

6.2. Structural alignment

The superposition of Rubisco models was performed with the least-squares comparison function in O [43].

Acknowledgements

This work was supported by the Swedish Research Council for Environment, Agricultural Sciences, and Spatial Planning (FORMAS), Helge Ax:sson Johnssons stiftelse, and the European Union.

References

[1] V. Albert, A. Backlund, K. Bremer, DNA characters and cladistics: the optimization of functional history, in: R.W. Scotland, D.J. Siebert (Eds.), Systematics Association Special Volume, vol. 52, Oxford, Clarendon Press, London, 1994, pp. 249–272.
[2] I. Andersson, Large structures at high resolution: spinach ribulose-1,5-bisphosphate carboxylase/oxygenase at 1.6 Å resolution, J. Mol. Biol. 259 (1996) 160–174.
[3] I. Andersson, S. Knight, G. Schneider, Y. Lindqvist, T. Lundqvist, C.I. Brändén, G.H. Lorimer, Crystal structure of the active site of ribulose-bisphosphate carboxylase, Nature 337 (1989) 229–234.
[4] I. Andersson, T.C. Taylor, Structural framework for catalysis and regulation in ribulose-1,5-bisphosphate carboxylase/oxygenase, Arch. Biochem. Biophys. 414 (2003) 130–140.
[5] T.J. Andrews, Catalysis by cyanobacterial ribulose-bisphosphate carboxylase large subunits in the complete absence of small subunits, J. Biol. Chem. 263 (1988) 12213–12219.
[6] T.J. Andrews, S.M. Whitney, Manipulating ribulose bisphosphate carboxylase/oxygenase in the chloroplasts of higher plants, Arch. Biochem. Biophys. 414 (2003) 159–169.
[7] H. Ashida, Y. Saito, C. Kojima, K. Kobayashi, N. Ogasawara, A. Yokota, A functional link between RuBisCO-like protein of Bacillus and photosynthetic RuBisCO, Science 302 (2003) 286–290.
[8] G. Bainbridge, P. Madgwick, S. Parmar, R. Mitchell, M. Paul, J. Pitts, A.J. Keys, M.A.J. Parry, Engineering Rubisco to change its catalytic properties, J. Exp. Bot. 46 (1995) 1269–1276.
[9] K. Bremer, The limits of amino acid sequence data in angiosperm phylogenetic reconstruction, Evolution 42 (1988) 795–803.
[10] A. Carré-Mlouka, A. Méjean, P. Quillardet, H. Ashida, Y. Saito, A. Yokota, I. Callebaut, A. Sekowska, E. Dittmann, C. Bouchier, N.T. de Marsac, A new Rubisco-like protein coexists with a photosynthetic Rubisco in the planktonic cyanobacteria microcystis, J. Biol. Chem. 281 (2006) 24462–24471.
[11] Z. Chen, C.J. Chastain, S.R. Al-Abed, R. Chollet, R.J. Spreitzer, Reduced CO_2/O_2 specificity of ribulose-bisphosphate carboxylase oxygenase in a temperature-sensitive chloroplast mutant of *Chlamydomonas*, Proc. Natl. Acad. Sci. USA 85 (1988) 4696–4699.
[12] Z. Chen, S. Hong, R.J. Spreitzer, Thermal-instability of ribulose-1,5-bisphosphate carboxylase oxygenase from a temperature-conditional chloroplast mutant of *Chlamydomonas-reinhardtii*, Plant Physiol. 101 (1993) 1189–1194.
[13] Z. Chen, R.J. Spreitzer, Chloroplast intragenic suppression enhances the low specificity of mutant ribulose-1,5-bisphosphate carboxylase/oxygenase, J. Biol. Chem. 264 (1989) 3051–3053.
[14] Z. Chen, W. Yu, J.H. Lee, R. Diao, R.J. Spreitzer, Complementing amino-acid substitutions within loop-6 of the α/β-barrel active-site influence the CO_2/O_2 specificity of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase, Biochemistry 30 (1991) 8846–8850.
[15] W.W. Cleland, T.J. Andrews, S. Gutteridge, F.C. Hartman, G.H. Lorimer, Mechanism of Rubisco: the carbamate as a general base, Chem. Rev. 98 (1998) 549–561.
[16] P.M.G. Curmi, D. Cascio, R.M. Sweet, D. Eisenberg, H.A. Schreuder, 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 (1992) 16980–16989.
[17] C.F. Delwiche, J.D. Palmer, Rampant horizontal transfer and duplication of rubisco genes in eubacteria and plastids, Mol. Biol. Evol. 13 (1996) 873–882.
[18] Y.C. Du, S.R. Peddi, R.J. Spreitzer, Assessment of structural and functional divergence far from the large subunit active site of ribulose-1,5-bisphosphate carboxylase/oxygenase, J. Biol. Chem. 278 (2003) 49401–49405.

[19] Y.C. Du, R.J. Spreitzer, Suppressor mutations in the chloroplast-encoded large subunit improve the thermal stability of wild-type ribulose-1,5-bisphosphate carboxylase/oxygenase, J. Biol. Chem. 275 (2000) 19844–19847.
[20] A.P. Duff, T.J. Andrews, P.M.G. Curmi, 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 (2000) 903–916.
[21] R.J. Ellis, Most abundant protein in the world, Trends Biochem. Sci. 4 (1979) 241–244.
[22] M.G. Esquivel, M. Anwaruzzaman, R.J. Spreitzer, Deletion of nine carboxy-terminal residues of the Rubisco small subunit decreases thermal stability but does not eliminate function, FEBS Lett. 520 (2002) 73–76.
[23] J.S. Farris, The retention index and the rescaled consistency index, Cladistics 5 (1989) 417–419.
[24] J. Felsenstein, Confidence limits on phlyogenies: an approach using the bootstrap, Evolution 39 (1985) 783–791.
[25] C.B. Field, M.J. Behrenfeld, J.T. Randerson, P. Falkowski, Primary production of the biosphere: Integrating terrestrial and oceanic components, Science 281 (1998) 237–240.
[26] W.M. Fitch, Toward defining the course of evolution: minimum change for a specific tree topology, Syst. Zool 20 (1971) 406–416.
[27] R. Flachmann, H.J. Bohnert, Replacement of a conserved arginine in the assembly domain of ribulose-1,5-bisphosphate carboxylase/oxygenase small subunit interferes with holoenzyme formation, J. Biol. Chem. 267 (1992) 10576–10582.
[28] J. Galmes, J. Flexas, A.J. Keys, J. Cifre, R.A.C. Mitchell, P.J. Madgwick, R.P. Haslam, H. Medrano, M.A.J. Parry, Rubisco specificity factor tends to be larger in plant species from drier habitats and in species with persistent leaves, Plant Cell Environ. 28 (2005) 571–579.
[29] Garcia-Murria, M.G., Karkehabadi, S., Marín-Navarro, J., Satagopan, S., Spreitzer, R.J., Andersson, I., Moreno, J., Structural and functional consequences of the substitution of vicinal residues Cys172 and Cys192 in the large subunit of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Chlamydomonas reinhardtii*, Bioch. J., in press.
[30] A.A. Gatenby, S.M. van der Vies, D. Bradley, Assembly in *E. coli* of a functional multisubunit ribulose bisphosphate carboxylase from a blue-green alga, Nature 314 (1985) 617–620.
[31] J.L. Gibson, F.R. Tabita, Different molecular forms of D-ribulose-1,5-bisphosphate carboxylase from *Rhodopseudomonas sphaeroides*, J. Biol. Chem. 252 (1977) 943–949.
[32] M. Giordano, J. Beardall, J.A. Raven, CO_2 concentrating mechanisms in algae: mechanisms, environmental modulation, and evolution, Annu. Rev. Plant Biol. 56 (2005) 99–131.
[33] G. Glöckner, A. Rosenthal, K. Valentin, The structure and gene repertoire of an ancient red algal plastid genome, J. Mol Evol. 51 (2000) 382–390.
[34] S. Gutteridge, D.F. Rhoades, C. Herrmann, Site-specific mutations in a loop region of the C-terminal domain of the large subunit of ribulose bisphosphate carboxylase/oxygenase that influence substrate partitioning, J. Biol. Chem. 268 (1993) 7818–7824.
[35] S. Hansen, V.B. Vollan, E. Hough, K. Andersen, The crystal structure of Rubisco from *Alcaligenes eutrophus* reveals a novel central eight-stranded β-barrel formed by β-strands from four subunits, J. Mol. Biol. 288 (1999) 609–621.
[36] T.E. Hanson, F.R. Tabita, 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 (2001) 4397–4402.
[37] F.C. Hartman, M.R. Harpel, Structure, function, regulation, and assembly of D-ribulose-1,5-bisphosphate carboxylase oxygenase, Annu. Rev. Biochem. 63 (1994) 197–234.
[38] N.R. Hayashi, A. Oguni, T. Yaguchi, S.Y. Chung, H. Nishihara, T. Kodama, Y. Igarashi, Different properties of gene products of three sets ribulose 1,5-bisphosphate carboxylase/oxygenase from a marine obligately autotrophic hydrogen-oxidizing bacterium, *Hydrogenovibrio marinus* strain MH-110, J. Ferment. Bioeng. 85 (1998) 150–155.
[39] S. Hong, R.J. Spreitzer, Complementing substitutions at the bottom of the barrel influence catalysis and stability of ribulose-bisphosphate carboxylase/oxygenase, J. Biol. Chem. 272 (1997) 11114–11117.
[40] R.L. Houtz, A.R. Portis, The life of ribulose 1,5-bisphosphate carboxylase/oxygenase-posttranslational facts and mysteries, Arch. Biochem. Biophys. 414 (2003) 150–158.
[41] H.J. Imker, A.A. Fedorov, E.V. Fedorov, S.C. Almo, J.A. Gerlt, Mechanistic diversity in the RuBisCO superfamily: the "Enolase" in the methionine salvage pathway in *Geobacillus kaustophilus*, Biochemistry 46 (2007) 4077–4089.
[42] F. Jeanmougin, J.D. Thompson, M. Gouy, D.G. Higgins, T.J. Gibson, Multiple sequence alignment with Clustal X, Trends Biochem. Sci. 23 (1998) 403–405.
[43] T.A. Jones, J.Y. Zou, S.W. Cowan, M. Kjeldgaard, Improved methods for building protein models in electron density maps and the location of errors in these models, Acta Crystallogr. A 47 (1991) 110–119.
[44] D.B. Jordan, W.L. Ogren, Species variation in the specificity of ribulose biphosphate carboxylase/oxygenase, Nature 291 (1981) 513–515.
[45] M.V. Kapralov, D.A. Filatov, Widespread positive selection in the photosynthetic Rubisco enzyme, BMC Evol. Biol. 7 (2007) 73.
[46] S. Karkehabadi, S.R. Peddi, M. Anwaruzzaman, T.C. Taylor, A. Cederlund, T. Genkov, I. Andersson, R.J. Spreitzer, Chimeric small subunits influence catalysis without causing global conformational changes in the crystal structure of ribulose-1,5-bisphosphate carboxylase/oxygenase, Biochemistry 44 (2005) 9851–9861.
[47] S. Karkehabadi, S. Satagopan, T.C. Taylor, R.J. Spreitzer, I. Andersson, Structural analysis of altered large-subunit loop-6/carboxy-terminus interactions that influence catalytic efficiency and CO_2/O_2 specificity of ribulose 1,5-bisphosphate carboxylase/oxygenase, Biochemistry 46 (2007) 11080–11089.
[48] S. Karkehabadi, T.C. Taylor, I. Andersson, Calcium supports loop closure but not catalysis in Rubisco, J. Mol. Biol. 334 (2003) 65–73.
[49] S. Karkehabadi, T.C. Taylor, R.J. Spreitzer, I. Andersson, Altered intersubunit interactions in crystal structures of catalytically compromised ribulosebisphosphate carboxylase/oxygenase, Biochemistry 44 (2005) 113–120.
[50] E.A. Kelogg, N.D. Juliano, The structure and function of RuBisCO and their implications for systematic studies, Am. J. Bot. 84 (1997) 413–428.
[51] W.A. King, J.E. Gready, T.J. Andrews, Quantum chemical analysis of the enolization of ribulose bisphosphate: the first hurdle in the fixation of CO_2 by Rubisco, Biochemistry 37 (1998) 15414–15422.
[52] K. Kitano, N. Maeda, T. Fukui, H. Atomi, T. Imanaka, K. Miki, Crystal structure of a novel-type archaeal Rubisco with pentagonal structure, Structure 9 (2001) 473–481.
[53] A.G. Kluge, J.S. Farris, Quantitative phyletics and the evolution of anurans, Syst. Zool. 18 (1969) 1–32.
[54] S. Knight, I. Andersson, C.-I. Brändén, Crystallographic analysis of ribulose 1,5-bisphosphate carboxylase from spinach at 2.4 Å resolution: subunit interactions and the active site, J. Mol. Biol. 215 (1990) 113–160.
[55] W.A. Laing, W.L. Ogren, R.H. Hageman, Regulation of soybean net photosynthetic CO_2 fixation by the interaction of CO_2, O_2, and ribulose 1,5-diphosphate carboxylase, Plant Physiol. 54 (1974) 678–685.
[56] H. Li, M.R. Sawaya, F.R. Tabita, D. Eisenberg, Crystal structure of a RuBisCO-like protein from the green sulfur bacterium *Chlorobium tepidum*, Structure 13 (2005) 779–789.
[57] G.H. Lorimer, H.M. Miziorko, Carbamate formation on the epsilon-amino group of a lysyl residue as the basis for the activation of ribulosebisphosphate carboxylase by CO_2 and Mg^{2+}, Biochemistry 19 (1980) 5321–5328.
[58] T. Lundqvist, G. Schneider, Crystal-structure of the binary complex of ribulose-1,5-bisphosphate carboxylase and its product, 3-phospho-D-glycerate, J. Biol. Chem. 264 (1989) 3643–3646.
[59] T. Lundqvist, G. Schneider, Crystal-structure of activated ribulose-1,5-bisphosphate carboxylase complexed with its substrate, ribulose-1,5-bisphosphate, J. Biol. Chem. 266 (1991) 12604–12611.

[60] T. Lundqvist, G. Schneider, Crystal-structure of the ternary complex of ribulose-1,5-bisphosphate carboxylase, Mg(II), and activator CO_2 at 2.3-Å resolution, Biochemistry 30 (1991) 904–908.
[61] D.R. Maddison, D.L. Swofford, W.P. Maddison, Nexus: an extensible file format for systematic information, Syst. Biol. 46 (1997) 590–621.
[62] P. Maliga, Plastid transformation in higher plants, Annu. Rev. Plant Biol. 55 (2004) 289–313.
[63] Y. Marcus, H. Altman-Gueta, A. Finkler, M. Gurevitz, Dual role of cysteine 172 in redox regulation of ribulose 1,5-bisphosphate carboxylase/oxygenase activity and degradation, J. Bacteriol. 185 (2003) 1509–1517.
[64] H. Mauser, W.A. King, J.E. Gready, T.J. Andrews, CO_2 fixation by Rubisco: computational dissection of the key steps of carboxylation, hydration, and C–C bond cleavage, J. Am. Chem. Soc. 123 (2001) 10821–10829.
[65] E. Mizohata, H. Matsumura, Y. Okano, M. Kumei, H. Takuma, J. Onodera, K. Kato, N. Shibata, T. Inoue, A. Yokota, Y. Kai, Crystal structure of activated ribulose-1,5-bisphosphate carboxylase/oxygenase from green alga *Chlamydomonas reinhardtii* complexed with 2-carboxyarabinitol-1,5-bisphosphate, J. Mol. Biol. 316 (2002) 679–691.
[66] M.K. Morell, H.J. Kane, T.J. Andrews, Carboxylterminal deletion mutants of ribulosebisphosphate carboxylase from rhodospirillum-rubrum, FEBS Lett. 265 (1990) 41–45.
[67] J. Moreno, R.J. Spreitzer, C172S substitution in the chloroplast-encoded large subunit affects stability and stress-induced turnover of ribulose-1,5-bisphosphate carboxylase/oxygenase, J. Biol. Chem. 274 (1999) 26789–26793.
[68] D. Morse, P. Salois, P. Markovic, J.W. Hastings, A nuclear-encoded form II RuBisCO in dinoflagellates, Science 268 (1995) 1622–1624.
[69] B.A. Murphy, F.J. Grundy, T.M. Henkin, Prediction of gene function in methylthioadenosine recycling from regulatory signals, J. Bacteriol. 184 (2002) 2314–2318.
[70] J. Newman, C.-I. Brändén, T.A. Jones, Structure determination and refinement of ribulose 1,5-bisphosphate carboxylase/oxygenase from *Synechococcus* PCC6301, Acta. Crystallogr. D 49 (1993) 548–560.
[71] J. Newman, S. Gutteridge, The X-ray structure of *Synechococcus* ribulose-bisphosphate carboxylase/oxygenase-activated quaternary complex at 2.2-Å resolution, J. Biol. Chem. 268 (1993) 25876–25886.
[72] J. Newman, S. Gutteridge, Structure of an effector-induced inactivated state of ribulose 1,5-bisphosphate carboxylase/oxygenase – the binary complex between enzyme and xylulose 1,5-bisphosphate, Structure 2 (1994) 495–502.
[73] Y. Okano, E. Mizohata, Y. Xie, H. Matsumura, H. Sugawara, T. Inoue, A. Yokota, Y. Kai, X-Ray structure of Galdieria Rubisco complexed with one sulfate ion per active site, FEBS Lett. 527 (2002) 33–36.
[74] C.M. Ott, B.D. Smith, A.R. Portis, R.J. Spreitzer, 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 (2000) 26241–26244.
[75] M.A.J. Parry, P.J. Andralojc, R.A.C. Mitchell, P. Madgwick, A.J. Keys, Manipulation of Rubisco: the amount, activity, function and regulation, J. Exp. Bot 54 (2003) 1321–1333.
[76] M.A.J. Parry, P. Madgwick, S. Parmar, M.J. Cornelius, A.J. Keys, Mutations in loop six of the large subunit of ribulose-1,5-bisphosphate carboxylase affect substrate specificity, Planta 187 (1992) 109–112.
[77] M.A.J. Parry, C.N.G. Schmidt, A.J. Keys, S. Gutteridge, Activation of ribulose 1,5-bisphosphate carboxylase by Ca^{2+}, FEBS Lett. 159 (1983) 107–111.
[78] F.G. Pearce, T.J. Andrews, The relationship between side reactions and slow inhibition of ribulose-bisphosphate carboxylase revealed by a loop 6 mutant of the tobacco enzyme, J. Biol. Chem. 278 (2003) 32526–32536.
[79] A. Portis, Rubisco activase – Rubisco's catalytic chaperone, Photosynth. Res. 75 (2003) 11–27.
[80] D.G. Price, M.R. Badger, F.J. Woodger, B.M. Long, Advances in understanding the cyanobacterial CO_2-concentrating-mechanism (CCM): functional components, Ci transporters, diversity, genetic regulation and prospects for engineering into plants Advance Access published on, J. Exp. Bot. (July 4, 2007). doi:10.1093/jxb/erm112.
[81] R.T. Ramage, B.A. Read, F.R. Tabita, Alteration of the alpha helix region of cyanobacterial ribulose 1,5-bisphosphate carboxylase/oxygenase to reflect sequences found in high substrate specificity enzymes, Arch. Biochem. Biophys. 349 (1998) 81–88.
[82] B. Ranty, T. Lundqvist, G. Schneider, M. Madden, R. Howard, G.H. Lorimer, Truncation of ribulose-1,5-bisphosphate carboxylase oxygenase (rubisco) from *Rhodospirillum rubrum* affects the holoenzyme assembly and activity, EMBO J. 9 (1990) 1365–1373.
[83] R. Rowan, S.M. Whitney, A. Fowler, D. Yellowlees, Rubisco in marine symbiotic dinoflagellates: form II enzymes in eukaryotic oxygenic phototrophs encoded by a nuclear multigene family, Plant Cell. 8 (1996) 539–553.
[84] R.F. Sage, Variation in the k_{cat} of Rubisco in C_3 and C_4 plants and some implications for photosynthetic performance at high and low temperature, J. Exp. Bot. (2002) 609–620.
[85] S. Satagopan, R.J. Spreitzer, Substitutions at the Asp-473 latch residue of *Chlamydomonas* ribulosebisphosphate carboxylase/oxygenase cause decreases in carboxylation and CO_2/O_2 specificity, J. Biol. Chem. 279 (2004) 14240–14244.
[86] G. Schneider, S. Knight, I. Andersson, C.I. Brändén, Y. Lindqvist, T. Lundqvist, Comparison of the crystal-structures of L2 and L8S8 rubisco suggests a functional-role for the small subunit, EMBO J. 9 (1990) 2045–2050.
[87] G. Schneider, Y. Lindqvist, C.-I. Brändén, G. Lorimer, Three-dimensional structure of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Rhodospirillum rubrum* at 2.9 Å resolution, EMBO J. 5 (1986) 3409–3415.
[88] G. Schneider, Y. Lindqvist, T. Lundqvist, Crystallographic refinement and structure of ribulose-1,5-bisphosphate carboxylase from *Rhodospirillum-rubrum* at 1.7-A resolution, J. Mol. Biol. 211 (1990) 989–1008.
[89] H.A. Schreuder, S. Knight, P.M.G. Curmi, I. Andersson, D. Cascio, C.-I. Brändén, D. Eisenberg, Formation of the Rubisco active site by a disorder-order transition from the unactivated to the activated form, Proc. Natl. Acad. Sci. USA 90 (1993) 9968–9972.
[90] H.A. Schreuder, S. Knight, P.M.G. Curmi, I. Andersson, D. Cascio, R.M. Sweet, C.I. Brändén, D. Eisenberg, Crystal-structure of activated tobacco rubisco complexed with the reaction-intermediate analog 2-carboxy-arabinitol 1,5-bisphosphate, Protein Sci. 2 (1993) 1136–1146.
[91] T. Shikanai, C.H. Foyer, H. Dulieu, M.A.J. Parry, A. Yokota, A point mutation in the gene encoding the Rubisco large subunit interferes with holoenzyme assembly, Plant Mol. Biol. 31 (1996) 399–403.
[92] J.M. Shively, G. van Keulen, W.G. Meijer, Something from almost nothing: carbon dioxide fixation in chemoautotrophs, Annu. Rev. Microbiol. 52 (1998) 191–230.
[93] S.A. Smith, F.R. Tabita, Positive and negative selection of mutant forms of prokaryotic (cyanobacterial) ribulose-1,5-bisphosphate carboxylase/oxygenase, J. Mol. Biol. 331 (2003) 557–569.
[94] E. Söderlind, G. Schneider, S. Gutteridge, Substitution of Asp193 to Asn at the active-site of ribulose-1,5-bisphosphate carboxylase results in conformational-changes, Eur. J. Biochem. 206 (1992) 729–735.
[95] R.J. Spreitzer, Genetic dissection of Rubisco structure and function, Annu. Rev. Plant Physiol, Plant Mol. Biol. 44 (1993) 411–434.
[96] R.J. Spreitzer, Questions about the complexity of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase, Photosynth. Res. 60 (1999) 29–42.
[97] R.J. Spreitzer, Role of the Rubisco small subunit, Arch. Biochem. Biophys. 414 (2003) 141–149.
[98] R.J. Spreitzer, M.G. Esquivel, Y.C. Du, P.D. McLaughlin, Alanine-scanning mutagenesis of the small-subunit βA–βB loop of chloroplast ribulose-1,5-bisphosphate carboxylase/oxygenase: substitution at Arg-71 affects thermal stability and CO_2/O_2 specificity, Biochemistry 40 (2001) 5615–5621.
[99] R.J. Spreitzer, S.R. Peddi, S. Satagopan, Phylogenetic engineering at an interface between large and small subunits imparts land-plant kinetic properties to algal Rubisco, Proc. Nat. Acad. Sci. USA 102 (2005) 17225–17230.
[100] R.J. Spreitzer, M.E. Salvucci, Rubisco: structure, regulatory interactions, and possibilities for a better enzyme, Annu. Rev. Plant Biol. 53 (2002) 449–475.

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

[102] D.L. Swofford, PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), Version 4.0, Sinauer Associates, Sunderland, Massachusetts, 2003.

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

[104] F.R. Tabita, T.E. Hanson, H. Li, S. Satagopan, J. Singh, S. Chan, Function, structure, and evolution of the Rubisco-like proteins and their Rubisco homologs, Microbiol. Mol. Biol. Rev. 71 (2007) 576–599.

[105] O. Tapia, H. Fidder, V.S. Safont, M. Oliva, J. Andres, Enzyme catalysis: transition structures and quantum dynamical aspects: modeling Rubisco's oxygenation and carboxylation mechanisms, Int. J. Quant. Chem. 88 (2002) 154–166.

[106] T.C. Taylor, I. Andersson, 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 (1996) 95–101.

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

[108] T.C. Taylor, I. Andersson, Structure of a product complex of activated spinach ribulose-bisphosphate carboxylase, Biochemistry 36 (1997) 4041–4046.

[109] T.C. Taylor, A. Backlund, R.J. Spreitzer, K. Björhall, A. Andersson, First crystal structure of Rubisco from a green alga – *Chlamydomonas reinhardtii*, J. Biol. Chem. 276 (2001) 48159–48164.

[110] T.C. Taylor, M.D. Fothergill, I. Andersson, 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 (1996) 32894–32899.

[111] G.G.B. Tcherkez, G.D. Farquhar, T.J. Andrews, Despite slow catalysis and confused substrate specificity, all ribulose bisphosphate carboxylases may be nearly perfectly optimized, Proc. Natl. Acad. Sci. USA 103 (2006) 7246–7251.

[112] S.M. Whitney, S. von Caemmerer, G.S. Hudson, T.J. Andrews, Directed mutation of the Rubisco large subunit of tobacco influences photorespiration and growth, Plant Physiol 121 (1999) 579–588.

[113] K.Y.J. Zhang, D. Eisenberg, 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 (1994) 64–69.

[114] K.Y.J. Zhang, D. Eisenberg, Solid state phase transition in the crystal structure of ribulose 1,5-bisphosphate carboxylase/oxygenase, Acta Crystallogr. (1994) 258–262 D 50.