SPECIAL ISSUE REVIEW PAPER

Multiple Rubisco forms in proteobacteria: their functional significance in relation to CO₂ acquisition by the CBB cycle

Murray Ronald Badger¹,²,* and Emily Jane Bek
¹ Molecular Plant Physiology Group, Research School of Biological Sciences, The Australian National University, Canberra, ACT, Australia
² ARC Centre of Excellence in Plant Energy Biology, Research School of Biological Sciences, The Australian National University, Canberra, ACT, Australia

Received 8 August 2007; Revised 19 September 2007; Accepted 22 October 2007

Abstract
Rubisco is the predominant enzymatic mechanism in the biosphere by which autotrophic bacteria, algae, and terrestrial plants fix CO₂ into organic biomass via the Calvin–Benson–Basham reductive pentose phosphate pathway. Rubisco is not a perfect catalyst, suffering from low turnover rates, a low affinity for its CO₂ substrate, and a competitive inhibition by O₂ as an alternative substrate. As a consequence of changing environmental conditions over the past 3.5 billion years, with decreasing CO₂ and increasing O₂ in the atmosphere, Rubisco has evolved into multiple enzymatic forms with a range of kinetic properties, as well as co-evolving with CO₂-concentrating mechanisms to cope with the different environmental contexts in which it must operate. The most dramatic evidence of this is the occurrence of multiple forms of Rubisco within autotrophic proteobacteria, where Forms II, IC, IBc, IAc, and IAq can be found either singly or in multiple combinations within a particular bacterial genome. Over the past few years there has been increasing availability of genomic sequence data for bacteria and this has allowed us to gain more extensive insights into the functional significance of this diversification. This paper is focused on summarizing what is known about the diversity of Rubisco forms, their kinetic properties, development of bacterial CO₂-concentrating mechanisms, and correlations with metabolic flexibility and inorganic carbon environments in which proteobacteria perform various types of obligate and facultative chemo- and photoautotrophic CO₂ fixation.

Key words: Autotrophic bacteria, chemotrophic bacteria, CO₂-concentrating mechanism, photosynthesis, proteobacteria, Rubisco.

Introduction
Rubisco is the predominant enzymatic mechanism in the biosphere by which autotrophic bacteria, algae, and terrestrial plants fix CO₂ into organic biomass (Tabita, 1999; Atomi, 2002). Its biochemical role is chiefly as the central part of the Calvin–Benson–Basham (CBB) reductive pentose phosphate pathway where it combines CO₂ with ribulose 1,5-bisphosphate to form two molecules of 3-phosphoglycerate. Hence, there has been tremendous research interest in its biochemistry, catalytic function, protein structure, and evolution, and a number of reviews have appeared over recent years that have covered various aspects of this research (Cleland et al., 1998; Spreitzer and Salvucci, 2002).

The imperfection of Rubisco
It is well recognized that Rubisco is not a perfect catalyst in terms of its enzymatic properties and the ecological and cellular environments in which it is situated (Tcherkez et al., 2006). These imperfections are to be found in three major areas. Firstly, its catalytic turnover rate is slow in relation to the rates of CO₂ fixation which occur in many autotrophic cells and this requires an investment of a significant amount of protein in this single catalytic step. This is especially so in higher plants, and it is telling that it is regularly cited as the most abundant protein on earth. A second impediment is that it has a low affinity for CO₂
as a substrate, although this has varied with evolution, which will be elaborated on below, and there has been a continual evolutionary struggle to improve Rubisco performance, by both increases in catalytic affinity for CO₂ and addition of active CO₂-acquisition mechanisms which help it to perform (Badger et al., 1998). Finally, it has suffered from the fact that molecular O₂ is an alternative substrate for the reaction (Rubisco oxygenase) and produces P-glycolate as a product. As a result, O₂ inhibits the catalytic affinity for CO₂ and cells must deal with the toxic P-glycolate thus produced (see Cleland et al., 1998; Spreitzer and Salvucci, 2002).

The major evolutionary forces which have acted on the evolution of Rubisco must be seen in terms of its imperfections listed above. Over the past 3.5 billion years, the earth’s biosphere has seen a dramatic change in the concentration of its substrates, with oxygenic photosynthesis causing O₂ to rise from an anaerobic starting point to reach as high as 35% O₂ in the late Paleozoic (Berner, 2001; Berner et al., 2003), and is currently 21%. There have also been significant concurrent declines in CO₂, from an initial concentration in the percentage range, to fall below 0.02% 280 Mya (Mora et al., 1999), and is currently 0.037% and rising (Berner and Kothavala, 2001). In addition, nitrogen as a limiting nutrient resource has varied tremendously in terrestrial and aquatic environments. Evolution has thus had to address how Rubisco is optimized with regard to CO₂ limitation, inhibition by O₂, and by the scarcity of nitrogen. Evolutionary changes in the enzyme which reflect responses to these pressures are evident in a significant decline in Kₘ(CO₂), increases in Vₘₐₓ for the carboxylase reaction (Tcherkez et al., 2006). From an organismal perspective, there has also been the evolution of a diverse array of active CO₂-concentrating mechanisms (CCMs) which elevate CO₂ at the active site of Rubisco thus largely minimizing its imperfections (Badger and Price, 1992; Badger et al., 1998).

The evolutionary diversification of Rubisco types and co-evolution between Rubisco properties and cellular CCMs is most evident in CO₂-fixing bacteria. Over the past few years there has been increasing availability of genomic sequence data for bacteria and this has allowed us to gain more extensive insights into this co-evolution. This review is focused on summarizing what is known about the diversity of Rubisco forms, their kinetic properties, development of CCMs, and correlations with the environment. In particular, the co-existence of multiple Rubisco types in many proteobacteria is explored.

Multiple Rubisco forms

Table 1 summarizes the recognized presence, structure, and function of four related Rubisco enzyme forms in archaea, bacteria, and eukaryotes. Of these, only Forms I, II, and III have been shown to have RuBP-dependent CO₂-fixing ability (for review, see Ashida et al., 2005). Form IV enzymes have been termed ‘Rubisco-like’ proteins and have been shown to be functional in the methionine salvage pathway in many bacteria (Ashida et al., 2003). There are suggestions that Form IV Rubisco may have been the evolutionary progenitor of CO₂-fixing Rubisco as it is known today (Ashida et al., 2005). Form III Rubiscos have to date only been found in archaea, and their CO₂-fixing role has been somewhat uncertain due to the lack of an identifiable phosphoribulokinase in archaea (Watson et al., 1999; Finn and Tabita, 2004), although some archaea have recently been identified with potential phosphoribulokinase genes (Mueller-Cajar and Badger, 2007). Most recently, it has been shown to participate in a pathway for AMP metabolism, with RuBP being supplied from AMP and phosphate (Sato et al., 2007). Rubisco Forms I and II are the enzymes which are directly involved in carbon metabolism associated with the CBB cycle and have a well-recognized autotrophic CO₂-fixing function which supports growth. These enzymes are found in a wide range of chemo-, organo-, and phototrophs,

<table>
<thead>
<tr>
<th>Rubisco form</th>
<th>Macromolecular structure</th>
<th>Phylogenetic occurrence</th>
<th>Enzymatic function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Green</td>
<td>Form IA L8S8</td>
<td>Cyanobacteria (Form IA), proteobacteria (Forms LAc and IAq)</td>
<td>CBB cycle</td>
</tr>
<tr>
<td></td>
<td>Form IB L8S8</td>
<td>Cyanobacteria (Form IBe), chlorophyte (green) algae, higher plants (Form IB)</td>
<td>CBB cycle</td>
</tr>
<tr>
<td>Red</td>
<td>Form IC L8S8</td>
<td>Proteobacteria</td>
<td>CBB cycle</td>
</tr>
<tr>
<td>Form II</td>
<td>Form ID L8S8</td>
<td>Heterokont and haptophyte algae (non-green)</td>
<td>CBB cycle</td>
</tr>
<tr>
<td>Form III</td>
<td>L2</td>
<td>Proteobacteria, archaea, and dinoflagellate algae</td>
<td>CBB cycle</td>
</tr>
<tr>
<td>Form IV: ‘Rubisco-like’</td>
<td>L2?</td>
<td>Bacteria, archaea, including both photosynthetic and non-photosynthetic bacteria</td>
<td>RuPP pathway</td>
</tr>
</tbody>
</table>

Table 1. Rubisco enzyme forms and their phylogenetic distribution

Information presented in this table was obtained from various review sources including Tabita (1999, 2004) and Ashida et al. (2005). The nature of Form LAc, LAc, and IBe Rubiscos is described in the text. Table S1 available at JXB online presents a detailed view of this phylogenetic distribution showing the presence of Forms I, II, III, and IV Rubisco genes in archaea, bacteria, algae, and plants.
occurring in bacteria, algae, and plants. Form I enzymes show considerable diversification and have been divided into four distinct clades. Forms IA and B belong to a ‘Green’ grouping and are found in proteobacteria, cyanobacteria, green algae, and higher plants. Forms IC and D are members of a ‘Red’ group, occurring in proteobacteria and non-green algae.

Table 2 shows a more complete view of the existence of different Form I and II Rubisco in proteobacteria and cyanobacteria. The data in the table are based on identification of Rubisco genes in publicly available genome sequence information, together with other information gathered from papers and protein database sequences. Table S1 in Supplementary data available at JXB online extends this table to include a summary of the presence of Forms I, II, III, and IV Rubisco genes in archaea, bacteria, and eukaryotes.

The nomenclature used for Rubisco forms in Table 2 is somewhat different from previous terminology, particularly in relation to Form I enzymes. As will be discussed below, Form IA enzymes are divided into two distinct types, IA<sub>C</sub> and IA<sub>q</sub>, based on two distinct types of small subunits and gene arrangements (see Figs 1, 3). Form IB enzymes have been subclassified into IB and IB<sub>c</sub> to indicate the Form IB<sub>c</sub> in cyanobacteria which is associated with carboxysomes (Badger et al., 2002). References to Form IC and D, II, III, and IV enzymes are in accordance with others (Tabita, 2004; Ashida et al., 2005).

From Table 2 and Table S1 in Supplementary data available at JXB online, it is evident that all obligate photoautotrophic organisms employing the CBB cycle possess a single form of Rubisco which has evolved to serve the primary role of CO<sub>2</sub> fixation during autotrophic capture of CO<sub>2</sub> by the CBB cycle from the external environment. This includes cyanobacteria with either Form IA<sub>C</sub> or IB<sub>c</sub>, chlorophyte algae with Form IB, chromophyte and rhodophyte algae (including haptophytes and heterokonts) with Form ID, dinoflagellates with Form II, and higher plants with Form IB (see Table S1 in Supplementary data available at JXB online). The general conclusion from this might be that when you have one primary carbon source for CO<sub>2</sub> fixation (CO<sub>2</sub> or HCO<sub>3</sub>−) then specialization in Rubisco becomes a primary driving force, particularly when this is associated with a chloroplast structure as it is in algae and higher plants.

For archaea, the predominant specialization in a Form III enzyme is most likely because of its proposed special role in AMP metabolism rather than autotrophic CO<sub>2</sub> fixation (Sato et al., 2007). However, the apparent presence of a Rubisco which may be more related to Form II enzymes in Methanococcoides burtonii (see Fig. S1 in Supplementary data available at JXB online) raises an interesting question about the role of this enzyme in this species. However, this species appears to have the genes associated with AMP metabolism (Sato et al., 2007) and it would seem most likely that it was also associated with this pathway.

The proteobacteria in general stand apart from these other organisms in having genes for up to three different forms of Rubisco including Form IA<sub>C</sub>, Form IA<sub>q</sub>, one or more Form IC, and Form II (Table 2). The absence of Form IB Rubisco from proteobacteria is interesting, given that cyanobacteria have evolved to contain either Form IA<sub>C</sub> or Form IB<sub>c</sub>. This situation may indicate that Form IB evolved more recently than the other Form I Rubiscos, being derived from Form IA. This has been previously speculated on and is linked to questions about the evolution of two types of carboxysomes in proteobacteria and cyanobacteria (Badger et al., 2002; Badger and Price, 2003).

The reasons for multiple Rubiscos in proteobacteria must presumably lie in the flexibility of the CO<sub>2</sub>-fixing lifestyles associated with their generally plastic metabolism which includes CO<sub>2</sub> derived from organic and inorganic substrates and environments where the levels of CO<sub>2</sub> and O<sub>2</sub> can fluctuate from high to low levels, and they have opted for a strategy where different forms of Rubisco are used under different environmental situations. The following part of this review will focus on how the presence and function of multiple forms of Form I and II Rubisco in proteobacteria can be interpreted from knowledge of the properties of the different forms of Rubisco and the metabolic behaviour and ecological environments of particular organisms.

A distinction between two different Form IA Rubiscos

Previous studies and reviews have classed all Form IA Rubiscos under one grouping. However, with multiple bacterial genome sequences becoming available, examination of Form IA Rubisco gene structure arrangements within the genomes clearly indicates that there are two different types of Form IA structures. This is illustrated in Fig. 1 where Rubisco gene region structures from a number of bacteria are shown. Gene region structures for two Form IA enzymes from Thiomicrospira crunogena are shown. One of the Form IA<sub>L</sub> and S gene pairs is associated with the genes coding for cbbQ1 and cbbO1. This Rubisco has been termed Form IA<sub>C</sub>. The other Form IA<sub>L</sub> and S gene pair is associated with a standard set of α-carboxysome genes (Cannon et al., 2001, 2003) and is presumably associated with the formation of a carboxysome structure. This has been termed Form IA<sub>q</sub>. A third Rubisco Form II is found near to the Form IA<sub>C</sub> gene set, associated with a second cbbQ/O gene pair. Such an arrangement of genes has been previously seen in Hydrogenovibrio marinus (Hayashi et al., 1999; Yoshizawa et al., 2004), where the same three sets of Rubisco genes are found.
Table 2. The presence of various Form I and II Rubiscos in proteobacteria and cyanobacteria together with information about their metabolism and the ecological habitats with which they are associated.

The documentation of the presence of these Rubisco forms was performed largely using the JGI Integrated Microbial Genomes browser (http://img.jgi.doe.gov/cgi-bin/pub/main.cgi) together with other information extracted from papers where the presence of multiple or single Rubisco forms are distinguished. The sequence nomenclature differs somewhat from previous studies in the subdivision of Forms 1A and 1B. Form IAq and Form IAc are distinguished in Fig. 2A and B by the amino acid sequences of the small subunits and their respective genome associations with cbbQ (for 1Aq) or α-carboxysome (Form 1Ac) genes. Form 1Bc genes are associated with β-carboxysome genes. Information on metabolism and ecological habitats was compiled from the DOE microbial genome project descriptions (http://genome.jgi-psf.org/mic_home.html), NCBI genome project database (http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=genomes), and Prescott et al. (2005). The potential presence of carboxysomes is indicated in the ‘Carb’ column (Y/N).

<table>
<thead>
<tr>
<th>Group</th>
<th>Rubisco form</th>
<th>Carb</th>
<th>Species</th>
<th>Oxygen</th>
<th>Metabolism</th>
<th>Ecological habitat</th>
<th>Taxonomy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alpha IAc</td>
<td>II N</td>
<td>Rhodospirillum rubrum ATCC 11170</td>
<td>Facultative anaerobe</td>
<td>Facultative photomixotroph</td>
<td>Multiple; freshwater</td>
<td>Rhodospirillales,</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>II N</td>
<td>Acidiphilum cryptum JF-5</td>
<td>Facultative anaerobe</td>
<td>Facultative chemolithomixotroph</td>
<td>Mine site, low pH, heavy metal-contaminated habitats</td>
<td>Rhodospirillales,</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC II N</td>
<td>Bradyrhizobium japonicum USDA 110</td>
<td>Aerobic</td>
<td>Organotroph</td>
<td>Root/soil</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Bradyrhizobium sp. BTA1</td>
<td>Aerobic</td>
<td>Facultative photo-organomixotroph</td>
<td>Stem/root/soil</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Nitrobacter hamburgensis</td>
<td>Aerobic</td>
<td>Facultative chemolithomixotroph</td>
<td>Soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Nitrobacter winogradskii Nb-255</td>
<td>Facultative anaerobe</td>
<td>Facultative chemolithomixotroph</td>
<td>Soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodopseudomonas palustris CGA009</td>
<td>Aerobic</td>
<td>Photoorganomixotroph</td>
<td>Temperate; soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodopseudomonas palustris BisA53</td>
<td>Facultative anaerobe</td>
<td>Facultative photo-organomixotroph</td>
<td>Soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodopseudomonas palustris BisB18</td>
<td>Facultative anaerobe</td>
<td>Facultative photo-organomixotroph</td>
<td>Soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodopseudomonas palustris BisB5</td>
<td>Facultative anaerobe</td>
<td>Facultative photo-organomixotroph</td>
<td>Soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodopseudomonas palustris HaA2</td>
<td>Facultative anaerobe</td>
<td>Facultative photo-organomixotroph</td>
<td>Soil/water</td>
<td>Rhizobiales, Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC N</td>
<td>Sinorhizobium melloti 1021</td>
<td>Aerobic</td>
<td>Organotroph</td>
<td>Soil/root nodules</td>
<td>Rhizobiales; Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC N</td>
<td>Xanthobacter autotrophicus Py2</td>
<td>Aerobic</td>
<td>Organotroph</td>
<td>Soil/sludge</td>
<td>Rhizobiales; Xanthobacteriaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Paracoccus denitrificans 1222</td>
<td>Aerobic</td>
<td>Facultative photoorganomixotroph</td>
<td>Soil and sewage sludge</td>
<td>Rhodobacters; Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodobacter sphaeroides 2.4.1</td>
<td>Aerobic</td>
<td>Facultative photomixotroph</td>
<td>Multiple</td>
<td>Rhodobacters; Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodobacter sphaeroides ATCC17025</td>
<td>Aerobic</td>
<td>Photomixotroph</td>
<td>Multiple</td>
<td>Rhodobacters; Bradyrhizobiales</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Dechloromoronas aromatica RCB</td>
<td>Facultative anaerobe</td>
<td>Organomixotroph — oxidizes benzene aerobically</td>
<td>Multiple; anoxic water sediments</td>
<td>Rhodocyclaces; Rhodocyclaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Rhodofex ferrireducens</td>
<td>Facultative anaerobe</td>
<td>Organomixotroph — reduces Fe(III)</td>
<td>Crater Lake, Oregon, depth 18 feet — anoxic sediments</td>
<td>Burkholderiales; Comamonadaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Hydrogenophaga pseudoflava</td>
<td>Facultative anaerobe</td>
<td>Facultative photomixotroph</td>
<td>Soils, hydrogen oxidizing, organic substrates</td>
<td>Burkholderiales; Comamonadaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC N</td>
<td>Burkholderia xenovorans LB-400</td>
<td>Aerobic</td>
<td>Facultative chemooorganomixotroph</td>
<td>PCB-containing landfill/soils</td>
<td>Burkholderiales; Burkholderiaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC N</td>
<td>Ralstoniaeutropha H16</td>
<td>Aerobic</td>
<td>Chemoorganotroph</td>
<td>PCB-containing soils</td>
<td>Burkholderiales; Burkholderiaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC N</td>
<td>Rubrivivax gelatinosus PM1</td>
<td>Aerobic</td>
<td>Photomixotroph</td>
<td>Freshwater, sewage, and activated sludge</td>
<td>Burkholderiales; Burkholderiaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Cupriavidus metallidurans CH34</td>
<td>Facultative anaerobe</td>
<td>Obligate chemolithotroph</td>
<td>Decantation tank of zinc factory</td>
<td>Burkholderiales; Burkholderiaceae</td>
<td></td>
</tr>
<tr>
<td>Alpha IAc</td>
<td>IC Y</td>
<td>Thiomonas intermedia K12</td>
<td>Aerobic</td>
<td>Obligate chemolithoautotroph</td>
<td>Soil and aquatic, both marine and freshwater</td>
<td>Burkholderiales; Burkholderiaceae</td>
<td></td>
</tr>
</tbody>
</table>
The amino acid alignments for all bacterial Form IA and Form IB small subunits are shown in Fig. 2. A major difference between the proteins coded for by the Form IA genes is to be found in the small subunits of the L8S8 enzymes. Figure S2 in Supplementary data available at JXB online also shows an alignment and phylogenetic tree for the corresponding large subunits, to demonstrate the much smaller differences in this protein subunit. All the Form IAq small subunit sequences are characterized by a six-amino-acid insertion at the N-terminal end of the protein. This holds for all the genes which are present in the database and where known gene arrangements are available. Interestingly, the alignment in Fig. 2 showing an equivalent insertion is also missing from Form IBc small subunits in cyanobacteria. Where pairs of Form IAc and IAq Rubiscos are present in the same organism (Hyrogenovibrio marinus, Thiomicrospira crunogena, Chromatium vinosum, and Acidithiobacillus ferrooxidans), the six-amino-acid insertion is always present in the Form IAq small subunit. There are two main groupings of Form IA small subunit sequences, one containing the α-cyanobacteria, and the other the chemoautotrophic proteobacteria. However, the chemoautotrophs Hyrogenovibrio marinus, Thiomicrospira crunogena, and Chromatium vinosum group phylogenetically with the cyanobacteria as far as the Form IAc sequences are concerned. Figure S3 in Supplementary data available at JXB online is included for completeness, showing amino acid sequence alignments of bacterial Form I small subunits together with a phylogenetic tree. Again the distinction between Form IAc and Form IAq is readily apparent.
The major difference between the two Form IA Rubisco small subunit types is obviously that the Form IAc is associated with carboxysomes while the Form IAq probably is not. Does this mean that this insertion is in some way associated with the ability or inability to form carboxysomes? From the crystal structure of the Form IAc Rubisco from *Halothiobacillus neapolitanus* (1SVD; CA Kerfeld et al., unpublished data, 2005) and the crystal structure of the cyanobacteria Form IBc enzyme (Newman et al., 1993; Newman and Gutteridge, 1994) it can be deduced that this region of the small subunit sequence occurs at a position which would be on the top exterior face of each small subunit, with potential interactions between adjoining small subunits and perhaps external proteins. Given that there are predicted interactions between the Rubisco holoenzyme and β-carboxysome shell proteins (Ludwig et al., 2000; Long et al., 2005) an insertion at this position could disrupt these interactions if they were occurring with the small subunits. Interestingly, in the Form IBc enzyme, this insertion point would be immediately adjacent to the conserved KERRY amino acid domain region of the small subunit, which has been shown to be present in the three or four small subunit repeat regions present in the β-carboxysomal shell protein CcmM (Ludwig et al., 2000). This region has been implicated in interactions between CcmM carboxysome protein and the large subunit of Rubisco, but the reality is not clear. However, the large carboxysomal shell proteins of α-carboxysomes, CsoS2 and CsoS3, show no analogous Form IA small subunit repeats (data not shown) so the analogy between these types of interactions is not clear at present.

The presence of two Form IA enzymes in *Chromatium vinosum* (*Allochromatium vinosum*) (Table 2) was noted some 20 years ago (Viale et al., 1990; Kobayashi et al., 1991), but its significance was previously unrecognized. Unfortunately, there is no genome sequence for this organism nor extension of sequence around the two gene sets. However, the analysis of small subunits here would clearly predict that *Chromatium* may have a set of α-carboxysome genes and even functional carboxysomes. No carboxysomes have been reported for *Chromatium*; however, sulphur granules similar in structure to carboxysomes with a surrounding protein membrane have been described (Jordan and Chollet, 1985). The reality of carboxysomes in this species remains to be established.

Another interesting observation concerns the Rubisco gene arrangement and type in *Thiobacillus denitrificans* (Fig. 3). This species has a Form IAq Rubisco and a set of α-carboxysome genes that are somewhat removed from the Form IAq gene set and closer to the Form II gene. The above analysis would suggest that *Thiobacillus denitrificans* should not possess functional carboxysomes. The fact that carboxysomes have not been observed in this species (Cannon et al., 2003) has been noted and is consistent with these predictions. Perhaps *Thiobacillus denitrificans* originally contained two different Form IA Rubisco gene sets as *Acidithiobacillus ferrooxidans* does, but they have been lost as this species has modified its ecological niche range (see discussion below).

Previous assessment of Rubisco forms in bacteria has aggregated all Form IA Rubiscos together without making any distinction between those which are part of carboxysome structures and those that are not. Unfortunately, this distinction is not easily made on large subunit sequences which have generally been the basis of enzyme form classification. However, there is a very important relevance to recognizing the present or absence of carboxysomes with respect to recognizing if certain bacteria are undertaking an active CCM to enhance the effectiveness of CO2 capture. CCMs have been recognized in gamma- proteobacteria with Form IAc Rubiscos and carboxysomes (Baker et al., 1998; Dobrinski et al., 2005) and it can be reasonably assumed that they will occur regularly when a Form IAc Rubisco is found in the genome.
Gene arrangements associated with Form I and II Rubiscos

Previous reviews have spent some time focusing on the flanking genes which are associated with different Rubisco genes in bacteria (Kusian and Bowien, 1997; Shively et al., 1998; Tabita, 1999; Dubbs and Tabita, 2004). The possible variation is illustrated by the arrangements shown in Figs 1 and 3, and a summary of these gene arrangement patterns is given in Table 3.

Form II Rubiscos are found in two arrangements. One arrangement is in association with the \( \text{cbb} \) genes which code for enzymatic components of the CBB cycle (Cbb metabolic). In the second arrangement, the Form II Rubisco large subunit gene is associated with the downstream presence of the \( \text{cbbQ} \) and \( \text{cbbO} \) genes.

Previous analyses of these genes and the function of their protein products in \( \text{Hydrogenovibrio marinus} \) (Hayashi et al., 1999) have indicated that they may be involved in some form of post-translational regulation of both Form II and Form IAq enzymes (see below) in these species, with the \( \text{CbbQ} \) protein possessing two nucleotide-binding motifs showing interactions with ATP and \( \text{Mg}^{2+} \) (Hayashi and Igarashi, 2002).

These LII-cbb(metabolic) and LII-cbbQ-CbbO gene arrangements are well correlated with the metabolic functioning of the organisms in which they occur (Kusian and Bowien, 1997; Hayashi et al., 1999; Tabita, 1999; Scott et al., 2006). Facultative autotrophs which can switch between CBB cycle \( \text{CO}_2 \) fixation and external organic substrates as sources of carbon have the L-Cbb(metabolic) arrangement, showing clustering with Cbb metabolic genes. This presumably allows the expression of the CBB cycle to be tightly regulated by the availability of organic and inorganic substrates. On the other hand, obligate autotrophs must continually express the CBB cycle and close regulation with Rubisco is not necessary. In these cases the Cbb metabolic genes are located elsewhere in the genome.

It appears as though all Rubisco Form IC LS genes are found in association with \( \text{cbbX} \) located downstream. This gene code for a protein of unknown function but is thought to be involved with some form of post-translational activation of the Rubisco (Li and Tabita, 1997). This is generally, but not always, correlated with the presence of a number of \( \text{cbb} \) metabolic genes in the same cluster and the fact that Form IC enzymes appear to be associated with facultative autotrophs (mixotrophs) (see Table 2). These species include the beta-proteobacteria \( \text{Rubrivivax gelatinosus} \), \( \text{Cupriavidus necator} \), and \( \text{Burkholderia xenovorans} \), and the alpha-proteobacteria \( \text{Bradyrhizobium japonicum} \), \( \text{Bradyrhizobium BTA1} \), \( \text{Rhodopseudomonas palustris} \) (various strains), \( \text{Nitrobacter hamburgensis} \), \( \text{Nitrobrochoinigradsy} \), and \( \text{Xanthobacter agilis} \).
flavus, Paracoccus denitrificans, Rhodobacter sphaeroides, Rhodobacter capsulatus, and Sinorhizobium meliloti (deduced from gene arrangements accessed at JGI Integrated Microbial Genome Browser, http://img.jgi.doe.gov/cgi-bin/pub/main.cgi). Two exceptions to this correlation appear to be the beta-proteobacterium Nitrosospira multiformis and the gamma-proteobacterium Nitrosococcus oceanii, which are both obligate chemolithotrophs and have a single Form IC Rubisco gene set.

The distinguishing feature of the Form IAq gene arrangement is the universal presence of two genes downstream of the LS gene pair (Kusian and Bowien, 1997;
These two genes code for CbbQ and CbbO and the distribution of the cbb metabolic genes elsewhere in the genome. Form IAq Rubisco appears predominantly in obligate chemolithotrophs such as the beta-proteobacteria *Thiobacillus denitrificans*, *Cupravidus metallidurans*, and *Nitrosomonas europaea*, and the gamma-proteobacteria *Hydrogenovibrio marinus*, *Thiomicrospira crunogenea*, and *Acidithiobacillus ferrooxidans*. However, it is also found in the alpha-proteobacterium *Rhodopseudomonas BisB5*, the beta-proteobacterium *Hydrogenophaga pseudoalva*, and the gamma-proteobacterium *Methyllococcus capsulatus*, which are facultative in their use of organic and inorganic electron donors for CO2 fixation.

Form IAc Rubiscos are found in both obligate and facultative autotrophs (photo- and chemolithotrophs) and are always associated together with a cluster of a-carboxysome genes. The presence of Rubisco in a-carboxysomes is presumably part of the CCM syndrome found in these organisms (Badger and Price, 2003; Dobrinski *et al.*, 2005). The presence of Form IAc Rubisco and a-carboxysomes has been attributed to lateral gene transfer events (Badger *et al.*, 2002). The cbb metabolic genes are found elsewhere in the genome.

Form IBc Rubisco is found only in beta-cyanobacteria and the LS genes may be associated with the beta-carboxysome gene cluster. In those species where they are not adjacent, a beta-carboxysome gene cluster is located elsewhere in the genome. Beta-cyanobacteria are the only bacteria with the Form IB lineage of Rubisco.

Over the past 3.5 billion years, the earth’s atmosphere has evolved so that oxygen has been elevated from 0% to 21% and CO2 has declined from >1% to very limiting levels in the past 10 000 years (<0.025%) (Berner, 2001; Berner and Kothavala, 2001; Berner *et al.*, 2003). It is clear that these atmospheric and environmental changes have caused the evolution of the kinetic properties of Rubisco both between and within form types (Badger and Andrews, 1987; Badger *et al.*, 1998; Tabita, 1999). This evolution seems to be have been driven by the need to optimize kinetic properties to match the nature of CO2 substrate availability and the O2 levels of the catalytic environments in which they operate. However, the evolution of Rubisco has also been intimately linked with the development of CCMs which serve to capture decreasing levels of external inorganic carbon and elevate CO2 around the catalytic site of the enzyme (Badger *et al.*, 1998).

Table 4 shows a summary of some of the relevant kinetic properties of the various Form I and II Rubisco proteins identified in this review and the potential
presence of CCMs. Within each of the forms there is often a considerable spread of kinetic values for $K_c$. This spread may be associated with the exact nature of the CO$_2$ environment in which Rubisco operates. If the external CO$_2$ is generally high during the CO$_2$-fixation period in which the Rubisco is active or there is intervention of a CCM, then it is likely that higher $K_c$ values will have evolved. This is illustrated for the spread of values in non-green algae, green algae, and higher plants. In all cases the higher $K_c$ values are found in species with active CCMs while the low values are associated with what appears to be passive acquisition of CO$_2$ by photosynthesis (Badger et al., 1998).

Form II Rubiscos are characterized by a low $S_{rel}$ value (indicating a low discrimination against O$_2$ as an alternative substrate), a poor affinity for CO$_2$, and a relatively high $k_{cat}$. Based on these properties it has been generally concluded that Form II enzymes are adapted to functioning in low-O$_2$ and high-CO$_2$ environments. There is no clear evidence in proteobacteria that Form II Rubiscos operate in conjunction with an active CCM, but they are found in gamma-proteobacteria species, such as *Thiomicrora crunogenea* and *Hydrogenovibrio marinus* (Table 2), with multiple Rubisco forms where CCM mechanisms clearly operate at some time in the life cycle (Table 2), with multiple Rubisco forms where CCM mechanisms clearly operate at some time in the life cycle under some environmental conditions (Dobrinski et al., 2005) (see discussion below).

Form IC enzymes show improved $S_{rel}$ values and this is generally assumed to be the first measure of the evolution of enzymes that are better adapted to increased levels of O$_2$ (Tabita, 1999). They have medium to low affinities for CO$_2$, indicating an adaptation to environments with medium to high CO$_2$ but with O$_2$ present. Even though they are closely related to Form ID enzymes associated with red and non-green algae, they do not show the high $S_{rel}$ and low $K_c$ values found with some of these enzymes.

Form IAq enzymes show kinetic properties which overlap with those found for Form IC Rubiscos. Unlike their closely related Form IAc counterparts, they are not associated with carboxysomes and are most probably adapted to environments with medium to high CO$_2$ but with O$_2$ present. They are present together with Form IAc enzymes in some gamma-proteobacteria, including *Hydrogenovibrio marinus*, *Thiomicrora crunogenea*, and *Acidithiobacillus ferrooxidans* (Table 2), where clearly there is the potential for inorganic carbon accumulation at some times (Cannon et al., 2001; Dobrinski et al., 2005).

Form IAc Rubiscos are somewhat poorly studied with regard to their kinetic properties. They appear to have $S_{rel}$ values that are less than the values for the Form IBc Rubiscos associated with β-carboxysomes, and probably medium to low affinity for CO$_2$, although this is based on a very limited sample (see Table S3 in Supplementary data available at *JXB* online). Studies with the α-cyanobacteria *Synechococcus* and *Prochlorococcus* marine species and the gamma-proteobacterium *Thiomicrora crunogenea* (Hassidim et al., 1997; Dobrinski et al., 2005) clearly show they are part of a CO$_2$-concentrating mechanism and are able to function at low external CO$_2$ levels as a consequence of this.

Finally, the Form IBc Rubiscos from β-cyanobacteria are well studied, and have been shown to have medium $S_{rel}$ values and the lowest affinities for CO$_2$. They function as part of a well-characterized β-carboxysome-associated CCM, enabling efficient photosynthesis at very low levels of external CO$_2$.

**The carbon fixation strategies and growth habitats of proteobacteria**

Within the bacteria, there are a range of metabolic approaches to utilizing the CBB cycle and Rubisco for CO$_2$ fixation resulting in increased biomass. Two major types of autotrophic behaviour can be discerned among the various chemo- and phototrophs, namely obligate and facultative. Obligate autotrophs are generally solely dependent on CO$_2$ and the CBB cycle as the sole source of...
net carbon for growth, and these largely include the chemo- and photolithoautotrophs. Faculative autotrophs show metabolic flexibility allowing these organisms to grow on organic substrates as alternative carbon and energy sources (Kusian and Bowien, 1997; Shively et al., 1998). The organotrophs use only organic substrates but use the CBB cycle for CO₂ fixation and synthesis processes.

Obligate chemo- (e.g. many beta- and gamma-proteobacteria) and photoautotrophs in Table 2 can be seen as specialists in their growth strategies, and their metabolism is optimized for oxidation of a small number of reduced donors (Shively et al., 1998). In chemo-autotrophs, the electron donors include H₂, reduced nitrogen (e.g. NH₃, NO₂), sulphur (e.g. S₂O²⁻, H₂S), metals (e.g. Fe²⁺, Mn²⁺), and carbon compounds (e.g. CO, CH₄, CH₃OH). These reduced donors are produced from various sources, including anthropogenic, biological, and geological processes, which include industry and agriculture, anaerobic metabolism in sediments and animal guts, and volcanic activity (Shively et al., 1998, 2001). Photo-autotrophs can use either anoxygenic (PSI driven) photosynthesis with electrons derived from reduced inorganic or organic substrates; or oxygenic (PSI plus PS2) photosynthesis with electrons derived from H₂O. Anoxygenic photosynthesis is found within the proteobacteria, while oxygenic photosynthesis is confined to cyanobacteria.

By contrast, facultative autotrophic bacteria display a great metabolic diversity which allows them to grow on a wide range of substrates. A clear distinction with obligate autotrophs is that catabolic (including the CBB pathway) and anaerobic pathways of carbon metabolism are inducible (Kusian and Bowien, 1997; Shively et al., 1998; Dubbs and Tabita, 2004). Many of these bacteria have the ability to grow mixotrophically, enabling situations where autotrophic and heterotrophic growth metabolisms can be simultaneously engaged.

Obligate and facultative autotrophs occupy considerably different ecological niches as a result of their different approaches to acquiring reduced substrate electron donors. Obligate photo- and chemoaotrophs grow where there is a continuous or fluctuating supply of reduced inorganic compounds and a low abundance and turnover of organic donors. Conversely, facultative chemoautotrophs grow best when both inorganic and organic reduced compounds are present.

One important consideration with chemoautotrophs is that the reduced energy substrates, such as H₂ and reduced sulphur compounds, are produced anaerobically, in sediments or hydrothermal vents, for example, and these subsequently accumulate and diffuse to aerobic zones. Many chemoautotrophic bacteria require O₂ as an electron acceptor and are therefore limited to the oxic layer where they compete with chemical oxidation of the reduced inorganic electron donors. This results in these bacteria positioning themselves at the oxic/anoxic interface of sediments, stratified lakes, or hydrothermal vents. This interface is usually highly dynamic and would be expected to present a variable environment where fluctuations in reduced compounds, CO₂, and O₂ would be expected (Shively et al., 1998; Overmann, 2001). An added limitation for photoautotrophs is that they must also gain access to adequate light and this further determines their potential distribution.

Correlations between Rubisco complement and ecological niche

The above considerations result in autotrophic proteobacteria growing in a range of environments which differ widely in the spatial and temporal variation in CO₂ and O₂ which influence the effectiveness of Rubisco in fixing CO₂. Table 2 shows combined information about Rubisco complement, metabolic type, and the nature of the environments in which included bacteria generally occur. Using the general conclusions about functional niches which may be occupied by the different Form I and II Rubiscos (Table 4), the following conclusions and speculations between Rubisco complement and ecological adaptation can be drawn for bacterial species shown in Table 2.

Alpha-proteobacteria

The Rubisco-containing alpha-proteobacteria encompass a number of species in the Rhodospirillales, Rhizobiales, and Rhodobacteriales. They are an extremely flexible metabolic group including members that can perform anoxygenic photosynthesis (the non-sulphur purple bacteria), aerobic or anaerobic respiration, and fermentation. During phototrophy and chemo-autotrophy they can use a number of electron donors including H₂ and organic carbon compounds, while acquiring new carbon through the CBB cycle. However, the CBB cycle performs two distinct functions in these organisms, acting to supply new carbon during autotrophic growth and as an electron sink, which participates in maintaining redox balance during heterotrophic growth on organic substrates (Dubbs and Tabita, 2004). This is particularly true for phototrophs where excess electrons are produced as a result of oxidation of reduced organic substrates.

The alpha-proteobacteria inhabit environments which vary considerably in their O₂ and CO₂ levels. Inhabiting soil, sludge, water, and plant roots and stems, some environments can be anaerobic and high in CO₂ (sludge, water-filled soils), high in CO₂ but aerobic (aerated soils, plant root nodules), or low in CO₂ and aerobic (plant stem nodules). It is not surprising then that, overall, this group of bacteria contains the four main proteobacterial
Rubiscos, Forms II, IC, Iaq, and IAc (Table 2), and the complement in each species can be interpreted in terms of an adaptation to the particular characteristics of the environment they inhabit.

The presence of Form II and Form IC enzymes has been extensively studied in a number of the non-sulphur purple photosynthetic bacteria shown in Table 2, including *Rhodospirillum rubrum*, *Rhodopseudomonas palustris*, *Rhodobacter sphaeroides*, and *Rhodobacter capsulatus* (Dubbs and Tabita, 2004). The response of the expression of both Rubisco forms has been well studied, being related to the metabolic state and redox state of the cells and the CO2 and O2 levels (see Kusian and Bowien, 1997; Dubbs and Tabita, 2004). The expression of Form II Rubisco varies somewhat between species, and can be highly expressed under photoheterotrophic, photoautotrophic, and chemoheterotrophic growth conditions (Dubbs and Tabita, 2004). However, in general it does appear that the engagement of Form II is favoured under conditions of high CO2 (>1.5%) and low O2. Form IC Rubisco has been shown to be induced during photo-autotrophic growth at medium levels of CO2 (<1.5% CO2; Dubbs and Tabita, 2004). Based on the kinetic properties of the Rubisco forms (Table 4), it would be reasonable to expect that Form II Rubiscos would be favoured at high CO2 and low O2, while Form IC would be utilized when environments become more aerobic with somewhat reduced levels of CO2.

While most of the non-purple sulphur phototrophic bacteria have Form IC and Form II Rubiscos, there is variation. *Rhodospirillum rubrum* has only one Form II gene and presumably is restricted to engaging the CBB cycle under conditions of high CO2 and low O2. By contrast, *Rhodopseudomonas palustris* BisB5 and *Rhodobacter sphaeroides* ATCC 17025 have Form II, Form IC, and Form Iaq Rubisco gene sets. A clear distinction between the kinetic advantages of Form IC versus Form Iaq enzymes is not immediately obvious (see Table 4). One could speculate that Form Iaq may be more associated with growth under photolithoautotrophic conditions as these genes are not associated with other *cbb* metabolic genes and may be engaged when inorganic substrates are oxidized. Under these conditions the CO2 may be reduced (0.1–1%) and there may be more O2.

The photoorganotroph *Bradyrhizobium* sp. BTA1 is an interesting variant in that it has a Form IAc Rubisco suggesting a CCM and operation of Rubisco under low CO2 conditions. *Bradyrhizobium* BTA1 is a stem-nodulating bacterium of legumes (Giraud et al., 2000) and as such may perform anoxygenic photosynthesis in a relatively aerobic and low CO2 (<0.1%) environment, being surrounded by the green photosynthetic tissues of the plant stem and having immediate access to the gaseous atmosphere surrounding the stem. Thus Form IAc Rubisco may be important in the light during photo-organotrophic growth, while Form IC Rubisco may be used in the dark when CO2 levels are presumably much higher.

The response of the CBB cycle and Rubisco expression is likely to be similar between the photo- and chemoheterotrophs in the alpha-proteobacteria, except that the need for the CBB cycle to dissipate excess electrons will be less critical (Kusian and Bowien, 1997; Dubbs and Tabita, 2004). The facultative chemotroph *Acidiphilum cryptum* has Form II and Form IC Rubiscos, suggesting similar roles to the phototrophs. By contrast, the organotrophs *Bradyrhizobium japonicum* and *Xanthobacter autotrophicus* only have Form IC, suggesting engagement of Rubisco under conditions of medium CO2 and suboxic O2 as may be found in soils and root nodules. However, it is significant that two of the facultative chemolithoorganotrophs, *Nitrobacter hamburgensis* and *Nitrobacter winogradski*, have Form IAc Rubisco and *α*-carboxysome genes as well as Form IC Rubisco. This suggests the potential engagement of CCMs under low CO2 conditions. They grow by oxidizing both NO2 and, to a lesser extent, organic substrates, and inhabit soil and water environments (Starkenburg et al., 2006). This would suggest that Form IAc Rubisco and a CCM may be engaged under nitrite growth, particularly in water where CO2 may be reduced, while Form IC may be more important during growth on organic substrates.

**Beta-proteobacteria**

The beta-proteobacteria encompass largely obligate and facultative chemotrophs but they also include some photomixotrophs. Members of the Burkholderales and Nitrosomonadales are the predominant representatives of this group. They are generally less flexible in their metabolic strategies than many of the alpha-proteobacteria and this is related to the greater number of obligate chemolithotrophs represented in this group. The four proteobacterial Rubisco forms (II, IC, Iaq, and IAc) are also found in various members of the group.

The regulation of Rubisco expression has been studied extensively in the chemo-organotroph *Ralstonia eutropha* (Bowien and Kusian, 2002). It has only Form IC Rubisco which is induced under conditions of autotrophic growth on organic substrates and is repressed during heterotrophic growth. *Burkholderia xenovorans* appears similar in oxidizing organic substrates and having Form IC Rubisco. Engagement of Rubisco under medium CO2 with some O2 present is suggested. The photomixotroph *Rubrivivax gelatinosus* has two Form IC Rubiscos and homologies with the non-purple sulphur alpha-proteobacteria. However, the absence of a Form II Rubisco suggests CBB cycle activity under moderate CO2 levels with suboxic levels of O2 rather than anaerobic conditions.
Several of the obligate chemolithotrophs have a single Form IAq Rubisco. This includes *Capriavidus metallicidurans*, *Thiobacillus denitrificans*, and *Nitrosomonas europaea*. This is consistent with the obligate and more specialized metabolic nature of these bacteria. However, as discussed above, the advantage of a Form IAq Rubisco is not immediately obvious but is correlated with chemolithotrophic-driven CO₂ fixation at moderate CO₂ and O₂ levels. *Thiobacillus denitrificans* is interesting (see Fig. 3) in that it has a Form IAq Rubisco and α-carboxysome genes nearby. This may have been a remnant of a previous metabolic state which was adapted to lower CO₂ environments or a transition to a future state. *Hydrogenophaga pseudoflava*, a CO and H₂ oxidizer, also has a single IAq Rubisco, and would suggest that autotrophic metabolism using the CBB cycle is primarily associated with chemolithotrophic metabolism oxidizing H₂. *Nitrospira multiformis* is an exception being an obligate chemolithotroph, deriving its energy from ammonia oxidation, but having a single Form IC Rubisco. This is also true of other *Nitrospira* species (Utaker et al., 2002). Whether this is associated with advantages of this particular Rubisco, such as higher CO₂ environments, or a result of stochastic phylogenetic events remains to be determined.

Form IAc Rubisco and α-carboxysome genes are present in the obligate chemolithotrophs *Nitrosomonas europa* and *Thiimonas intermedia* K12. *Nitrosomonas europa*, like *Nitrosomonas europa* is an ammonia-oxidizing chemotroph, and is found in eutrophic environments (Chain et al., 2003). The distinction between it and *N. europa* must be found in the nature of the levels of CO₂ commonly found during CO₂ fixation. Presumably *N. europa* occupies aerobic but eutrophic environments where other CO₂-fixing activity must deplete the inorganic carbon and thus a high affinity CO₂-fixation system is needed. *Thiimonas intermedia* is a sulphur-oxidizing bacterium and contains both a Form IA and a Form II Rubisco. This could be explained by variation in its environment between anaerobic soils with high CO₂ and low O₂ and more aerobic aquatic environments with low CO₂.

**Gamma-proteobacteria**

The gamma-proteobacteria encompass a range of obligate photo- and chemolithotrophs, and some organotrophs which encompass families that include Chromatiaceae, Thiotrichales, and Acidithiobacillales. Form IAq and Form IA Rubiscos are dominant in this group, but there are some occurrences of Form IC and Form II enzymes. The photosynthetic (anoxygenic photosynthesis) purple sulphur bacteria are contained within the Chromatiaceae, but of these only *Allochromatium vinosum* is represented in Table 2. Species in the subfamily Chromatiaceae generally inhabit freshwater lakes and intertidal sandflats, while species in the Ectothiorhodospiraceae are associated with hypersaline waters. Many of the species oxidize reduced sulphur with and without the aid of anoxygenic photosynthesis.

Evidence for the potential participation of Form IAq Rubisco and carboxysomes in a high-affinity CCM in chemolithotrophs comes most strongly for some species in this group. Carboxysomes were originally discovered in the Thiobacilli, most of which are included in this group (Cannon et al., 2001, 2003). Insertional mutagenesis of carboxysome genes in *Halothiobacillus neapolitanus* resulted in a high CO₂ requirement for growth (Baker et al., 1998); active accumulation of inorganic carbon has been demonstrated in *Thiomicrospira crunogena*, and the affinity for inorganic carbon is increased when the cells are grown at limiting CO₂ (Dobrinski et al., 2005), and Form IAq Rubisco is induced when *Hydrogenovibrio marinus* is grown at low CO₂ (Yoshizawa et al., 2004). This evidence strongly supports the notion that Form IAq Rubisco and carboxysomes are a part of an active inorganic carbon acquisition mechanism in many chemotrophs.

Single genes for Form IAq Rubisco are found in the two members of the Ectothiorhodospiraceae, *Halorhodospira halophila* and *Alkalilimnicola ehrlichei*, which inhabit alkaline hypersaline lakes. These species probably occur at the chemocline within these environments, oscillating between anaerobic and microaerobic conditions, and experiencing variable CO₂ levels dependent on positioning within the chemocline interface. Even though these lakes are alkaline, and bicarbonate and carbonate rather than CO₂ may dominate the inorganic carbon pool, the CO₂ may be sufficiently high and growth relatively slow to allow fixation by the Form IAq Rubisco but not high enough to favour the use of a Form II enzyme. *Nitrococcus mobilis* in this group is a nitrite-oxidizing bacterium and contains a single Form IA Rubisco and α-carboxysome genes. This species inhabits marine surface-water environments and as such experiences a more aerobic and reduced CO₂ environment. The presence of carboxysomes and a CCM is consistent with this.

*Allochromatium vinosum* has been shown to have a dominant Form IAq Rubisco (Viale et al., 1989; Kobayashi et al., 1991). However, it is also apparent that it contains a second Form IAc Rubisco which is probably associated with the potential to make carboxysomes. The observation that the Form IAc genes seem weakly expressed may indicate that this is a gene that is not used for significant growth; however, it may also indicate that it is only expressed at low CO₂ levels which have not yet been tested. However, the fact that it occurs indicates that other species in this group may have flexibility to grow in medium and low CO₂ conditions.

The presence of Form IAq and Form II Rubiscos in the halotolerant sulphur-oxidizing species *Halothiobacillus*
neopolitanus would indicate a constant potential for this species to inhabit higher O_2 and lower CO_2 environments than for the Form IAq species, as well as photosynthesis under anaerobic and high CO_2 conditions. By contrast, the acid-tolerant and iron-oxidizer Acidithiobacillus ferrooxidans possesses both a Form IAc and a Form IAq Rubisco. This would suggest a potential adaptation for variable CO_2 and O_2 levels, giving it the ability to fix carbon both at low CO_2 and ambient O_2 as well as medium CO_2 and lower O_2 levels, but an absence from anaerobic environments.

Finally, within the Thiocrichales, Hydrogenovibrio marinus and Thiomicrospira crunogenea both have a complement of Form II, Form IAq, and Form IAc Rubiscos. As mentioned above, both these species have provided evidence for the dynamic nature of the expression and function of these genes in relation to the environment. In Hydrogenovibrio, although no cell physiology has been conducted, Form II Rubisco was expressed at very high CO_2 (15%), Form IAq was expressed at 2% CO_2, and Form IAc became dominant when grown at 0.15% CO_2. Somewhat unexpectedly, all three forms were expressed at 0.03% CO_2 (Yoshizawa et al., 2004). In Thiomicrospira, Rubisco expression has not been examined, but the cell physiology responded to growth at limiting CO_2 by increasing the affinity of CO_2 fixation for external inorganic carbon and actively accumulating inorganic carbon (Dobrinski et al., 2005). This was interpreted as clear evidence for a CCM associated with carboxysomes function. In the case of Thiomicrospira, a CCM may serve to increase the supply of CO_2 in a hydrothermal-vent habitat. It was speculated that temperature differences between the bottom water (2 °C) and hydrothermal efflux (35 °C) may create turbulent eddies in the water column which could expose cells to oscillations in CO_2 (ranging from 2 μM to 20 μM) and O_2 levels. In both these species, the Form II Rubiscos are presumably expressed in the more anaerobic environments at high CO_2.

The Rubisco genes have been surveyed within a range of obligate sulphur-oxidizing chemolithothrophs closely related to Thiomicrospira crunogenea and Hydrogenovibrio marinus (Tourova et al., 2006). There is a high occurrence of Form IAc Rubiscos in these species, indicating a general use of CCMs and carboxysomes to acquire CO_2 in this larger group. One interesting fact is that Thiomicrospira crunogenea is known as the fastest-growing mesophilic chemolithoautotroph (Jannasch et al., 1985), and indeed high growth rates are common for many of these sulphur-oxidizing bacteria (Tourova et al., 2006). The employment of CCMs is often most critical in fast-growing species where uptake of external inorganic carbon can exceed the supply rate to the surface of the cell, resulting in Ci-limiting conditions and the need to use both CO_2 and HCO_3− as carbon sources (Badger, 1987).

Rubisco genes present in different environments

There have been a number of studies surveying the presence of the different Rubisco form genes in bacterial assemblages from different environments, which have made some inferences about their phylogenetic, metabolic, and ecological significance. Sampling of marine phytoplankton communities has shown the presence of Form IA and IB Rubiscos associated with α- and β-cyanobacteria, together with Form ID Rubiscos associated with non-green algae such as diatoms and coccolithophorids (Pichard et al., 1997a, b; Paul et al., 1999). Analysis of deep-sea environments, including hydrothermal vents, has identified the presence of Form III Rubisco genes from archaea and Rubisco Forms II, IAc and IAq, IB, IC, and ID (Elsaied and Naganuma, 2001; Elsaied et al., 2007). The Form II and Form IA Rubiscos appeared most highly related to members of the Thiobacilli, Thiocrichales, where two or three Rubisco forms may be present in the same genome (see above). Rubisco genes have also been found in bacteria from an anoxic hypersaline basin (van der Wielen, 2006). Here there was evidence for Form I and II Rubiscos most closely related to obligate chemolithoautotroph Thiobacillus species and Hydrogenovibrio marinus. Sampling from a groundwater aquifer, which is anaerobic to microaerobic, has found the presence of Form II and Form IA Rubiscos again most closely related to obligate chemolithoautotrophs such a Halothiobacillus, Acidithiobacillus, and Thiobacillus species (Alfreider et al., 2003).

Terrestrial sites such as soils (Selesi et al., 2005; Tolli and King, 2005) and volcanic deposits (Nanba et al., 2004) have also been sampled. These reveal the presence of facultative and obligate chemolithotrophs. Agricultural and pine forest soils showed a narrow diversity of Form IIA Rubiscos related to Nitrobacter species, and a high diversity of Form IC Rubiscos related to a range of alpha- and beta-proteobacteria (Selesi et al., 2005; Tolli and King, 2005). Volcanic deposits showed Form IC Rubisco related to a number of facultative chemolithothrophs from the alpha-proteobacteria. However, microbial mats overlying the volcanic material were dominated by Form IA Rubiscos related to Thiobacilli species (Nanba et al., 2004).

Conclusions

It is apparent that genes for all five forms of Form I and Form II Rubisco are widely distributed among proteobacteria and cyanobacteria. This is most obvious in proteobacteria where up to three different forms can exist within the one bacterial genome. There are two ways to view the presence of different and multiple Rubiscos within an organism. First, it can be viewed as a simple result of phylogeny and stochastic inheritance. However, it can
also be viewed as a specific result of the adaptive evolution to particular ecological environments that each bacterium has undergone. This review has attempted to examine the extent to which the complement of Rubisco genes within a single bacterial species may fit with the environmental conditions and lifestyle it is adapted to. Bacteria with single Rubisco genes can be seen as specialists, with CO2 fixation by the CBB cycle occurring under a fairly constant and predictable set of environmental conditions relating to CO2 and O2 levels. Multiple Rubisco gene sets can be found in bacteria experiencing a more flexible lifestyle with respect to autotrophic and heterotrophic metabolism, as well as large fluctuations in the CO2 and O2 within the environment. It is interesting to note that this indicates a contrasting difference in evolutionary strategies employed by obligate photoautotrophs such as cyanobacteria, algae, and plants which possess a single Rubisco form, and proteobacteria. In obligate photoautotrophs, adaptation of Rubisco is based on the evolution of the kinetic properties of a single Rubisco and the flexible implementation of some form of CCM. The notion that bacteria may employ a number of different Rubisco forms, together with the implementation of a CCM to achieve evolutionary adaptation and acclimation is quite different. However, in reaching any clear conclusions about the role of different Rubisco forms in most species, it is evident there is a real lack of good data which describe the ecophysiology of proteobacteria in relation to the CO2 and O2 levels in their environments, and more studies of this nature relating gene expression of Rubisco form to environment and physiology are needed.

Supplementary data

**Figure S1.** Alignment and phylogenetic tree of Bacterial Rubisco Form I and Form II Large Subunit amino acid sequences. Alignments and phylogenetic tree were produced as described in Figure 2.

**Figure S2.** Alignment and phylogenetic tree of Bacterial Rubisco Form IA and IB Large Subunit amino acid sequences. Alignments and phylogenetic tree were produced as described in Figure 2.

**Figure S3.** Alignment and phylogenetic tree of Bacterial Rubisco Form I Small Subunit amino acid sequences. Alignments and phylogenetic tree were produced as described in Figure 2.

**Table S1.** The presence of various Form I, II, III and IV Rubiscos in Bacteria, Eukaryotes and archaea.

**Table S2.** The species names and amino acid sequences for Rubisco large and small subunits used for alignments and phylogenetic trees used in the paper.

**Table S3.** The kinetic properties of Form I and II Rubiscos found in Bacteria and Eukaryotes.

References


