	Occurrence, phylogeny and evolution of ribulose- 1,5-bisphosphate carboxylase/oxygenase genes in obligately chemolithoautotrophic sulfur-oxidizing bacteria of the genera <i>Thiomicrospira</i> and <i>Thioalkalimicrobium</i>						
	Tatjana P. Tourova, ¹ Elizaveta M. Spiridonova, ² Ivan A. Berg, ² Boris B. Kuznetsov ³ and Dimitry Yu. Sorokin ^{1,4}						
Correspondence Tatjana P. Tourova ttour@biengi.ac.ru	¹ Institute of Microbiology, Russian Academy of Sciences, p-t 60-letiya Oktyabrya, 7/2, Moscow, Russia						
	² Department of Microbiology, Moscow State University, Moscow, Russia						
	³ Centre Bioengineering, Russian Academy of Sciences, Moscow, Russia						
	⁴ Department of Biotechnology, Delft University of Technology, Delft, The Netherlands						
	The occurrence of the different genes encoding ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO), the key enzyme of the Calvin–Benson–Bassham cycle of autotrophic CO ₂ fixation, was investigated in the members of the genus <i>Thiomicrospira</i> and the relative genus <i>Thioalkalimicrobium</i> , all obligately chemolithoautotrophic sulfur-oxidizing <i>Gammaproteobacteria</i> . The <i>cbbL</i> gene encoding the 'green-like' form I RubisCO large subunit was found in all analysed species, while the <i>cbbM</i> gene encoding form II RubisCO was present only in <i>Thiomicrospira</i> species. Furthermore, species belonging to the <i>Thiomicrospira crunogena</i> 16S rRNA-based phylogenetic cluster also possessed two genes of green-like form I RubisCO, <i>cbbL-1</i> and <i>cbbL-2</i> . Both 16S-rRNA- and <i>cbbL</i> -based phylogenies of the						
Received 18 November 2005	Thiomicrospira-Thioalkalimicrobium-Hydrogenovibrio group were congruent, thus supporting its						
Revised 15 February 2006	monophyletic origin. On the other hand, it also supports the necessity for taxonomy reorganization of this group into a new family with four genera						
Accepted / March 2000	uns group into a new latinity with four genera.						

INTRODUCTION

The *Thiomicrospira* genus includes obligately chemolithoautotrophic sulfur-oxidizing bacteria (SOB) isolated from saline, mostly marine, habitats and belonging to the *Gammaproteobacteria* (Robertson & Kuenen 1999; Brinkhoff *et al.*, 1999c), except *Thiomicrospira denitrificans* which belongs to the *Epsilonproteobacteria* and needs to be reclassified. Currently, ten species are recognized in this genus. It also includes some not yet validly described strains. The two major features discriminating this genus from other SOB are salt tolerance and high growth rates. In fact, *Thiomicrospira crunogena* is known as the fastest-growing mesophilic chemolithoautotroph (Jannasch *et al.*, 1985). According to our experience with SOB in salt lakes, halophilic *Thiomicrospira* sp. outcompeted *Halothiobacillus* spp.

The GenBank/EMBL/DDBJ accession numbers for the sequences reported in this paper are DQ272527-DQ272537.

with the same salinity profile at micro-oxic conditions and with sulfide instead of thiosulfate as a substrate (unpublished data).

Recent phylogenetic 16S rRNA-based analysis and descriptions of new SOB related to Thiomicrospira (Brinkhoff et al., 1999a, b; Takai et al., 2004) have demonstrated (i) that the Thiomicrospira genus is heterogeneous, containing at least two different groups, clustering either with Thiomicrospira pelophila (Kuenen & Veldkamp, 1972; Wood & Kelly, 1989, 1993) or with Thiomicrospira crunogena, and (ii) that the genus itself is a member of a bigger group of closely related SOB. In particular, the haloalkaliphilic SOB of the genus Thioalkalimicrobium (Sorokin et al., 2001, 2002) and the hydrogen-oxidizing genus Hydrogenovibrio (Nishihara et al., 1991) are members of Thiomicrospira pelophila and Thiomicrospira crunogena clusters, respectively. Moreover, the SOB symbionts of marine clams Bathymodiolus and Calyptogena are firmly related to the whole Thiomicrospira-Thioalkalimicrobium-Hydrogenovibrio group ('Thiomicrospira group'). This evidence of their divergence clearly demands reorganization of the genus Thiomicrospira into

Abbreviations: RSCU, relative synonymous codon usage; RubisCO, ribulose-1,5-bisphosphate carboxylase/oxygenase; SOB, sulfur-oxidizing bacteria.

two genera based on *Thiomicrospira pelophila* and *Thiomicrospira crunogena* clusters, and creation of a new family '*Thiomicrospiraceae*', which would include at least five separate genera: the two *Thiomicrospira*-based genera, *Thioalkalimicrobium*, *Hydrogenovibrio* and, perhaps, the SOB symbionts of marine clams.

The obligate autotrophy of the '*Thiomicrospira* group' representatives is based on the high activity of the Calvin–Benson–Bassham cycle of inorganic carbon assimilation with ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO) as the key enzyme. Although the potential for heterotrophy has been currently claimed for three strains of the *Thiomicrospira crunogena* cluster (Takai *et al.*, 2004), the evidence given is not very convincing. Interestingly, *Thiomicrospira denitrificans*, a member of the *Epsilonproteobacteria*, uses the reductive tricarboxylic acid cycle for autotrophic CO₂ assimilation (Hugler *et al.*, 2005). This correlates with the anaerobic nature of this bacterium, which separates it from other *Thiomicrospira* species.

RubisCO exists in two distinct forms. Form I RubisCO is composed of eight large subunits and eight small subunits (L_8S_8), which are encoded by the *cbbL* and *cbbS* genes, respectively. This form is widely distributed in CO₂-fixing organisms, including all higher plants, algae, cyanobacteria and many autotrophic bacteria. Form I RubisCO is divided into two major groups, termed 'green-like' and 'red-like' (Delwiche & Palmer, 1996). Form II RubisCO is composed of large subunits only (L_x), encoded by the *cbbM* gene. This form is so far restricted to several phototrophic purple bacteria, aerobic and facultatively anaerobic chemoautotrophic bacteria, and dinoflagellates. In addition to these well-recognized forms, two novel types, forms III and IV RubisCO, have recently been revealed after the complete genome sequencing of some archaea and bacteria.

Some bacteria have been found to possess more than one set of RubisCO genes. Cupriavidus necator H16 (formerly Alcaligenes eutrophus H16) possesses two sets of almost identical genes which encode the red-like form I enzyme (Kusian et al., 1995), Allochromatium vinosum (formerly Chromatium vinosum) has two sets of divergent genes which encode the green-like form I enzyme (Viale et al., 1989) and Rhodobacter azotoformans has two sets of genes which encode both green- and red-like form I enzymes (Uchino & Yokota, 2003). Furthermore, Halothiobacillus neapolitanus (formerly Thiobacillus neapolitanus), Thiomonas intermedia (formerly Thiobacillus intermedius), Thiobacillus denitrificans, Rhodobacter sphaeroides and Rhodobacter capsulatus have genes for both form I and form II enzymes (Gibson & Tabita, 1977a, b; Shively et al., 1986; English et al., 1992; Stoner & Shively, 1993; Paoli et al., 1995). Acidithiobacillus ferrooxidans (formerly Thiobacillus ferrooxidans) and Hydrogenovibrio marinus have three different sets of RubisCO genes, two encoding the green-like form I enzyme, whereas the third one encodes a form II enzyme (Kusano et al., 1991; Heinhorst et al., 2002; Yaguchi et al., 1994; Nishihara et al., 1998).

In this work the phylogenetic diversity and evolution of the RubisCO genes of some *Thiomicrospira* and all *Thioalka-limicrobium* species has been analysed with the aim of obtaining additional insight into the relatedness of different species and their clusters within this SOB group.

METHODS

Bacterial strains. Three type strains of the genus *Thiomicrospira* were obtained from DSMZ. The cultures were grown on mineral medium with thiosulfate as a substrate as described by Brinkhoff *et al.* (1999a), except for *Thiomicrospira pelophila* where 100 μ g vitamin B₁₂ l⁻¹ was added. Three type strains belonging to the *Thioalkalimicrobium* genus were maintained in our laboratory and grown on a mineral medium with thiosulfate at pH 10, as described previously (Sorokin *et al.*, 2001).

DNA isolation and PCR amplification. DNA extraction was performed as described previously (Boulygina *et al.*, 2002). PCR was used to amplify the fragments of bacterial genes encoding the large subunit of red- and green-like form I RubisCO (*cbbL*) and form II RubisCO (*cbbM*) using specially developed and previously tested primer pairs (Spiridonova *et al.*, 2004). PCR products were analysed by electrophoresis in 1.0% agarose gel stained with ethidium bromide and documented by the BioDoc Analyse System (Biometra). PCR products were purified through low-melting-point agarose using Wizard PCR Preps kit (Promega).

Cloning and sequencing of the PCR fragments. Purified PCR products were cloned using the pGEM-T vector system (Promega). Plasmid DNA was extracted and purified using the Wizard MiniPrep kit (Promega). Clones containing appropriately sized inserts were sequenced from universal M13 forward and reverse primers (Sambrook *et al.*, 1989). Sequencing was performed with an ABI 3730 using the Big Dye Terminator v.3.1 sequencing reaction kit (Applied Biosystems).

Phylogenetic analysis. The preliminary analysis of the new sequences was done using BLAST from the NCBI server (www. ncbi.nlm.nih.gov/blast/). The nucleotide and inferred amino acid sequences were aligned with sequences from GenBank using CLUSTAL W (Thompson *et al.*, 1994). Phylogenetic trees were reconstructed using four different algorithms: neighbour-joining (Saitou & Nei, 1987) in the TREECONW program package (Van de Peer & De Wachter, 1994), and maximum-parsimony (Fitch, 1971), distance matrix (Fitch & Margoliash, 1967) and maximum-likelihood (Felsenstein, 1981) using PHYLIP 3.5c software (Felsenstein, 1993).

The relative synonymous codon usage (RSCU) values of the RubisCO genes were calculated using CODONW software (John Peden, www. molbiol.ox.ac.uk/cu). To investigate the major trends in codon usage in different species, CODONW was used to carry out a correspondent analysis. This resulted in a point in the codon space for each species, the positions of which sometimes suggested codon usage bias.

Levels of synonymous (dS) and non-synonymous (dN) nucleotide diversities were calculated with the YN00 program (PAML package; Yang, 2000) using the method of Yang & Nielsen (2000).

RESULTS

Detection and amplification of the RubisCO genes

Using the specific primer set for the green-like *cbbL* gene, fragments of about 800 bp were amplified from the DNAs of

the type strains of *Thiomicrospira crunogena*, *Thiomicrospira kuenenii*, *Thiomicrospira pelophila*, *Thioalkalimicrobium aerophilum*, *Thioalkalimicrobium cyclicum* and *Thioalkalimicrobium sibiricum*. The use of the primer set specific for red-like *cbbL* gave negative results for all species. The application of the *cbbM*-amplifying primer set resulted in the appearance of a PCR product of about 800 bp only for *Thiomicrospira* species.

The clones prepared from the *cbbL* PCR fragments yielded a single sequence-type for all three *Thioalkalimicrobium* species and *Thiomicrospira pelophila*, but two sequence-types for *Thiomicrospira crunogena* and *Thiomicrospira kuenenii*. The clones prepared from the *cbbM* PCR fragments yielded only one sequence-type for all *Thiomicrospira* species.

The results of the BLAST analysis indicated that all newly obtained sequences belonged to the RubisCO family. *Thiomicrospira crunogena* and *Thiomicrospira kuenenii* possessed three genes (*cbbL-1*, *cbbL-2* and *cbbM*), *Thiomicrospira pelophila* possessed two genes (*cbbL* and *cbbM*) and the *Thioalkalimicrobium* species possessed only a single RubisCO gene (*cbbL*).

Phylogenetic analysis of the cbbL sequences

The nucleotide and amino acid sequences of *cbbL* were aligned, the positions with gaps and ambiguous sequences were removed and the remaining 738 nt and 246 aa were used for further phylogenetic analysis. The phylogenetic trees constructed by neighbour-joining (Fig. 1), maximumparsimony, Fitch-Margoliash and maximum-likelihood algorithms (data not shown) were similar with minor exceptions both for nucleotide-based and amino-acidbased phylogenetic trees. In both nucleotide- and aminoacid-based trees, cbbL-2 of Thiomicrospira crunogena and Thiomicrospira kuenenii, and cbbL of Thiomicrospira pelophila and three Thioalkalimicrobium species formed a single cluster with high bootstrap values (100 and 97%) based on nucleotides and amino acids, respectively) with the cbbL-2 of the marine hydrogen-oxidizing SOB Hydrogenovibrio marinus. Similarly, cbbL-1 of Thiomicrospira crunogena and Thiomicrospira kuenenii formed a single cluster with high bootstrap values (100 % for both nucleotide- and amino-acid-based trees) with cbbL-1 of Hydrogenovibrio marinus.

However, for the '*Thiomicrospira* group' as a whole, the trees based on nucleotides and amino acids differed in their topologies and in the lengths of the branches (compare Fig. 1a and b). In the nucleotide-based tree, *cbbL-1* and *cbbL-2* formed a single group with an 87 % bootstrap value, clustering with *rbcL* of the cyanobacterium *Prochlorococcus marinus* with an 84 % bootstrap value, with very long branching. The closest neighbour of this cluster was the major cluster of RubisCO genes of the cyanobacteria *Prochlorothrix hollandica, Synechococcus* sp. PCC 6301, *Synechococcus* sp. PCC 7002 and *Anabaena* sp. PCC 7120 with a 92 % bootstrap value. In the tree based on amino acids,

cbbL-1 and *cbbL-2* of the '*Thiomicrospira* group' formed two separate clusters with uncertain branching, whereas *Prochlorococcus marinus* clustered with *Synechococcus* sp. WH7803 at a high bootstrap value (99%) with almost equal branch length. Also, the relatedness of the '*Thiomicrospira* group' and the major cluster of cyanobacteria was not detected in the latter tree.

The results of *cbbL* analysis within the 'Thiomicrospira group' correlated with its phylogenetic clustering based on 16S rRNA analysis (Fig. 2). In this tree, Thiomicrospira pelophila formed a cluster with Thioalkalimicrobium species ('Thiomicrospira pelophila cluster'), whereas Thiomicrospira crunogena and Thiomicrospira kuenenii formed another cluster with Hydrogenovibrio marinus (Thiomicrospira crunogena cluster). In compliance with these data, single *cbbL* genes of the Thioalkalimicrobium species were closely related (88.9% nucleotide similarity and 97.2% amino acid identity) to a single Thiomicrospira pelophila cbbL gene, whereas the cbbL-2 and cbbL-1 genes of Thiomicrospira crunogena and Thiomicrospira kuenenii were closer to cbbL-2 (87.7-90.1% nucleotide similarity and 96.8-98.0% amino acid identity) and cbbL-1 (85.2-92.0% nucleotide similarity and 89.6–93.0% amino acid identity) of Hydrogenovibrio marinus, respectively. However, a closer relatedness of cbbL genes of this whole group to cyanobacterial *rbcL* genes contradicts its phylogenetic position within the Gammaproteobacteria in a 16S-rRNA-based tree.

Phylogenetic analysis of cbbM

The nucleotide and amino acid sequences of *cbbM* were aligned, the positions with gaps and ambiguous sequences were removed and the remaining 777 nt and 259 aa were used for the phylogenetic analysis. The topologies of the nucleotide-based and amino-acid-based phylogenetic trees constructed by neighbour-joining (Fig. 3), maximumparsimony, Fitch–Margoliash and maximum-likelihood algorithms (data not shown) were similar for each method with minor exceptions.

In the nucleotide-based tree, the *cbbM* gene of *Thiomicro-spira* species and of *Hydrogenovibrio marinus* formed a single cluster (at 100 % bootstrap value). This cluster fell into the radiation of thioautotrophic *Beta-* and *Gammaproteobacteria* belonging to the genera *Thiobacillus*, *Acidithiobacillus*, *Halothiobacillus* and *Thiomonas*. In the amino-acid-based tree, only *Thiomicrospira pelophila* and *Thiomicrospira crunogena* remained in the thioautotrophic *cbbM* cluster, whereas *Thiomicrospira kuenenii* and *Hydrogenovibrio marinus* with almost identical amino acids sequences (98·1 %) formed a separate branch with an uncertain branching point position.

Nucleotide composition and codon usage

Genes acquired by horizontal transfer often have atypical GC content, codon bias and repetitive elements (Medigue *et al.*, 1991). Therefore, it was interesting to compare the

2162



Fig. 1. The phylogenetic position of the '*Thiomicrospira* group' species in *cbbL* molecular trees based on sequence analysis of (a) nucleotides and (b) translated amino acids. The sequences determined in this study are marked by bold type. Tree topography and evolutionary distances are given by the neighbour-joining method with Jukes and Cantor (for nucleotides) and Poisson (for amino acids) corrections. Numbers at the nodes indicate the percentage bootstrap values for the clade of this group in 1000 replications (the values for the maximum-parsimony method are given in parentheses).



Fig. 2. The phylogenetic position of '*Thiomicrospira* group' species in the 16S rRNA molecular tree. The sequences for which RubisCO genes were determined are marked by underlining. Tree topography and evolutionary distances are given by the neighbour-joining method with Jukes and Cantor distances. Numbers at the nodes indicate the percentage bootstrap values for the clade of this group in 1000 replications (the values for the maximum-parsimony method are given in parentheses).

GC content and codon usage of the RubisCO genes within the '*Thiomicrospira* group' to detect the role of possible gene transfer in their evolution. The total GC content of all analysed RubisCO genes was close to the genomic GC content for each species of the group (45·1–48·9 against 42.0–49.6 mol%, respectively). The GC₃ content (third position of codons) of RubisCO gene sequences (30.3-37.0 mol%) was lower than the total GC content of the RubisCO genes and the overall genomic GC content.



Fig. 3. The phylogenetic position of '*Thiomicrospira* group' species in *cbbM* molecular trees based on sequence analysis of (a) nucleotides and (b) translated amino acids. The sequences determined in this study are marked by bold type. Tree topography and evolutionary distances are given by the neighbour-joining method with Jukes and Cantor (for nucleotides) and Poisson (for amino acids) corrections. Numbers at the nodes indicate the percentage bootstrap values for the clade of this group in 1000 replications (the values for the maximum-parsimony method are given in parentheses).

Codon usage analysis was carried out on the RSCU data. Correspondence analysis of the results (Fig. 4) identified major trends in codon usage: the y axis is associated with GC₃ (Musto *et al.*, 1998), whereas the x axis is correlated with the frequencies of codons ending in C or U versus A or G (Fennoy & Bailey-Serres, 1993). The codon usage of all species of the *'Thiomicrospira* group' was typical of AT-biased micro-organisms in which codons with an A or T in the third position are used preferentially (Ohtaka & Ishikawa, 1993). In general, the codon usage of all RubisCO genes of the *'Thiomicrospira* group' was almost identical and differed from the RubisCO genes of other autotrophs in the RSCU correspondence analysis plot (Fig. 4). Thus analysis of codon usage did not show any intra-group bias that might be the result of gene transfer in these species.

Based on codon usage, the closest neighbours of the '*Thiomicrospira* group' were *Anabaena* sp. PCC 7120 and its immediate relatives. Moreover, the unusual position of the other cyanobacterium *Prochlorococcus marinus* on the RSCU correspondence analysis plot correlated with its unusual clustering in the *cbbL* nucleotide-based phylogenetic tree.

Synonymous and non-synonymous substitution analysis

Since synonymous (silent) mutations are largely invisible to natural selection, whereas non-synonymous (amino-acidchanging) mutations may be under strong selective pressure, comparison of the rates of fixation of these two types of mutation provides a powerful tool for understanding



Fig. 4. (a) Correspondence analysis of RSCU in 'Thiomicrospira group' species for cbbL-2 genes of the studied group (large •), *cbbL-1* genes of the studied group (\bigcirc) and *cbbL* genes of all other micro-organisms (small \bullet). Taa, *Thioalkalimicrobium* aerophilum; Tas, Thioalkalimicrobium sibiricum; Tac, Thioalkalimicrobium cyclicum; Tcrl-2, Thiomicrospira crunogena cbbL-2; Tkul-2, Thiomicrospira kuenenii cbbL-2; Tpel-2, Thiomicrospira pelophila cbbL-2; Hmal-2, Hydrogenovibrio marinus cbbL-2; Tcrl-1, Thiomicrospira crunogena cbbL-1; Tkul-1, Thiomicrospira kuenenii cbbL-1; Hmal-1, Hydrogenovibrio marinus cbbL-1; Afel-1, Acidithiobacillus ferrooxidans cbbL-1; Afel-2, Acidithiobacillus ferrooxidans cbbL-2; Tbde, Thiobacillus denitrificans; Htne, Halothiobacillus neapolitanus; Tmin, Thiomonas intermedia; Tsae, Thioalkalispira microaerophila; Avil-1, Allochromatium vinosum cbbL-1; Avil-2, Allochromatium vinosum cbbL-2; Tvni, Thioalkalivibrio nitratireducens; Tvpa, Thioalkalivibrio paradoxus; Tvth, Thioalkalivibrio thiocyanoxidans; Tvve, Thioalkalivibrio versutus; Tvja, Thioalkalivibrio jannaschii; Tvnit, Thioalkalivibrio nitratis; Tvde, Thioalkalivibrio denitrificans; Tvthi, Thialkalivibrio thiocyanodenitrificans; Esp, Ectothiorhodospira shaposhnikovii; Hhps, Hydrogenophaga pseudoflava; Ana, Anabaena sp. PCC 7120; Pho, Prochlorothrix hollandica; Pro, Prochlorococcus marinus; Syn, Synechococcus sp. PCC 7002; Syne, Synechococcus sp. PCC 6301; Synec, Synechococcus sp. WH 7803; Thpa, Thioclava pacifica; Rhca, Rhodobacter capsulatus; Hpps, Hydrogenophilus thermoluteolus. (b) Correspondence analysis of RSCU in 'Thiomicrospira group' species for cbbM genes of the studied group (\triangle) and *cbbM* genes of all other micro-organisms (\triangle). Tcr, *Thiomicrospira crunogena*; Tku, *Thiomicrospira kuenenii*; Tpe, Thiomicrospira pelophila; Hma, Hydrogenovibrio marinus; Rru, Rhodospirillum rubrum; Mma, Magnetospirillum magnetotacticum; Mag, magnetite-containing magnetic vibrio MV-1; Tbde, Thiobacillus denitrificans; Afe, Acidithiobacillus ferrooxidans; Rppa, Rhodopseudomonas palustris DH1; Rppal, Rhodopseudomonas palustris DCP3; Rhca, Rhodobacter capsulatus; Rhsp. Rhodobacter sphaeroides; Htne, Halothiobacillus neapolitanus; Tmin, Thiomonas intermedia; Sym1, Riftia pachyptila endosymbiont; Din1, Symbiodinium sp.; Din2, Gonyaulax polyhedra. Genes are plotted at their coordinates on the two axes produced by the analysis - the axes are described in Results.

the mechanisms of DNA sequence evolution (Yang & Nielsen, 2000). Therefore, synonymous and non-synonymous nucleotide substitution rates (dS and dN) and their ratio ($\omega = dN/$ d*S*) for the nucleotide sequences of *cbbL* and *cbbM* genes of the '*Thiomicrospira group*' were calculated (Table 1). dN was lower than dS in all cases with clear evidence of selective

constraint on amino acid replacements. This suggests that the RubisCO genes have evolved in all lineages under negative or purifying selection.

For *cbbL-2* in the genomes of all species of the group under study the ω values within the *Thiomicrospira pelophila* and Thiomicrospira crunogena clusters (0.0140-0.0291 and 0.0132-0.0200, respectively; see Table 1) were half those between these clusters (0.0294–0.0646). The high ω value (0.3363) for the almost identical *cbbL* sequences for the pair Thioalkalimicrobium aerophilum-Thioalkalimicrobium sibiricum could be considered as an exception. The intracluster ω values for *cbbL-1* of the *Thiomicrospira crunogena* cluster (0.0486-0.1135) were about 4- to 5-times higher than for *cbbL-2*. Moreover, the intra-cluster ω values for *cbbL-1* of this group were even higher than the inter-cluster ω values for *cbbL-2* between the *Thiomicrospira pelophila* and Thiomicrospira crunogena clusters, especially for the pair *Thiomicrospira kuenenii–Hydrogenovibrio marinus* (0.1135). This could be explained by the fact that intra-cluster dNvalues for cbbL-1 (0.0345-0.0624) were comparable with inter-cluster dN values for cbbL-2 (0.0524-0.0670), whereas dS values at the intra-cluster level were comparable for both *cbbL-1* and *cbbL-2* (0·3039–0·8602 and 0·3885–0·8611, respectively).

The ω value for *cbbM* genes of the pair *Thiomicrospira kuenenii–Hydrogenovibrio* marinus was relatively low

(0.0513) because of the low dN (0.0117) and dS (0.2282) values (Table 1). The dN and dS values for cbbM genes of other combinations were within the ranges of 0.0720–0.2740 and 1.4505–2.3291, respectively, which is much higher compared with the values for cbbL-1 and cbbL-2. However, the ω value for the pair *Thiomicrospira pelophila–Thiomicrospira crunogena* was much lower (0.0309) than the ω values for other pairs (0.1192–0.1823). This is a result of a relatively low non-synonymous nucleotide substitution rate (dN = 0.0720) for this pair in contrast to a high synonymous nucleotide substitution rate (dS = 2.3291).

These results demonstrated that the synonymous and nonsynonymous nucleotide substitution rates among the *cbbL*-*1*, *cbbL*-2 and *cbbM* genes of the '*Thiomicrospira* group' were different. This might be explained by the proposal that the selection pressure for *cbbL*-2 was higher than for *cbbL*-1 and *cbbM* and, therefore, by the higher significance of RubisCO encoded by the *cbbL*-2 gene.

DISCUSSION

The usage of functional genes encoding key metabolic enzymes as molecular markers is becoming common practice in phylogenetic studies. In the case of RubisCO it has been shown that phylogenetic reconstructions based on its analysis differ significantly from the results of traditional 16S rRNA-based studies for autotrophic organisms in

Table 1. Synonymous and non-synonymous nucleotide substitution ratios among three sets of RubisCO genes of the 'Thiomicrospira group'

Species pair		cbbL-1			cbbL-2			cbbM		
	dN	dS	dN/dS	dN	dS	dN/dS	dN	dS	dN/dS	
Thioalkalimicrobium aerophilum and Thioalkalimicrobium sibiricum		_	-	0.0055	0.0162	0.3363	-	-	-	
Thioalkalimicrobium aerophilum and Thioalkalimicrobium cyclicum		_	-	0.0055	0.3896	0.0140	_	_	-	
Thioalkalimicrobium aerophilum and Thiomicrospira pelophila		_	-	0.0145	0.6886	0.0210	_	_	-	
Thioalkalimicrobium aerophilum and Thiomicrospira crunogena		_	-	0.0524	1.7508	0.0299	_	_	-	
Thioalkalimicrobium aerophilum and Thiomicrospira kuenenii		_	_	0.0550	1.3241	0.0416	_	_	_	
Thioalkalimicrobium aerophilum and Hydrogenovibrio marinus		-	-	0.0585	1.0473	0.0559	-	-	-	
Thioalkalimicrobium cyclicum and Thioalkalimicrobium sibiricum		_	-	0.0110	0.3885	0.0282	-	_	_	
Thioalkalimicrobium cyclicum and Thiomicrospira pelophila		_	-	0.0201	0.7407	0.0271	-	_	_	
Thioalkalimicrobium cyclicum and Thiomicrospira crunogena		-	-	0.0585	1.3927	0.0420	-	-	-	
Thioalkalimicrobium cyclicum and Thiomicrospira kuenenii		_	-	0.0611	1.6184	0.0378	-	_	_	
Thioalkalimicrobium cyclicum and Hydrogenovibrio marinus		-	-	0.0647	1.1346	0.0570	-	-	-	
Thioalkalimicrobium sibiricum and Thiomicrospira pelophila		-	-	0.0200	0.6863	0.0291	-	-	-	
Thioalkalimicrobium sibiricum and Thiomicrospira crunogena		-	-	0.0583	1.9817	0.0294	-	-	-	
Thioalkalimicrobium sibiricum and Thiomicrospira kuenenii		_	-	0.0610	1.3102	0.0465	-	_	_	
Thioalkalimicrobium sibiricum and Hydrogenovibrio marinus		-	-	0.0645	0.9981	0.0646	_	-	_	
Thiomicrospira pelophila and Thiomicrospira crunogena		_	-	0.0640	1.5795	0.0405	0.0720	2.3291	0.0309	
Thiomicrospira pelophila and Thiomicrospira kuenenii		-	-	0.0670	1.1564	0.0579	0.2695	1.9061	0.1414	
Thiomicrospira pelophila and Hydrogenovibrio marinus		_	-	0.0661	1.1417	0.0579	0.2644	1.4505	0.1823	
Thiomicrospira crunogena and Thiomicrospira kuenenii		0.8226	0.0486	0.0090	0.6803	0.0132	0.2740	1.5779	0.1737	
Thiomicrospira kuenenii and Hydrogenovibrio marinus		0.3039	0.1135	0.0110	0.5477	0.0200	0.0117	0.2282	0.0513	
Thiomicrospira crunogena and Hydrogenovibrio marinus		0.8602	0.0726	0.0163	0.8611	0.0190	0.2675	2.2445	0.1192	

general (Watson & Tabita, 1997) and for particular groups, for example haloalkaliphilic SOB of the genus *Thioalkalivibrio* (Tourova *et al.*, 2005).

In contrast to the latter case, phylogenetic analysis of the '*Thiomicrospira* group' demonstrated a good correlation between 16S-rRNA- and RubisCO-based results. First of all, the analysis of the RubisCO genes showed a monophyletic origin of the group (including the previously studied genus *Hydrogenovibrio*), evident from the high probability of clustering of their RubisCO genes at the nucleotide level. The separate branching of *cbbL-1* of the *Thiomicrospira crunogena* cluster and *cbbM* of the pair *Thiomicrospira kuenenii–Hydrogenovibrio marinus* at the amino acid level could be explained by an increased rate of non-synonymous nucleotide replacements in their RubisCO genes (Table 1).

A two-subgroup division of the 'Thiomicrospira group' based on 16S rRNA phylogeny correlated with form I RubisCO gene analysis. In particular, cbbL-1 genes were found only in the Thiomicrospira crunogena cluster, and the topology of cbbL-2- and 16S-rRNA-based trees was very similar. At the same time, subdivision of the group was not evident on the basis of the *cbbM* gene analysis, which might be a result of a different pathway of evolution of this RubisCO form and a different selection pressure on the form II enzyme for different species of the group, depending on their ecological niches. According to a recent hypothesis, form II might be the more ancient type of the enzyme, optimally functioning under anaerobic conditions and high CO₂ concentration. In this case, form I can be considered as an aerotolerant descendant of form II (Watson & Tabita, 1997; Elsaied & Naganuma, 2001). Form II enzymes are conservative and uniform (Hernandez et al., 1996) in contrast to the more recent form I which has evolved into two types, green-like and red-like, according to their amino acid sequences. According to all the evidence presented above, it could be proposed that comparison of *cbbM* gene sequences allows us to trace only distant relatedness inside the 'Thiomicrospira group', but not more recent divergence of the group for two phylogenetic clusters. It is interesting to note that in the *cbbM*-based phylogenetic tree, the 'Thiomicrospira group' forms a monophyletic cluster with the other SOB of the genera Acidithiobacillus, Halothiobacillus, Thiobacillus and Thiomonas, which are currently assigned to the Gamma- and Betaproteobacteria based on 16S rRNA gene sequence analysis. It suggests a possible common origin of these chemolithoautotrophs with similar metabolism. In contrast, this common origin was not evident from the analysis of *cbbL* genes and this could be explained by lateral gene transfer of *cbbL* to the ancestor of the 'Thiomicrospira group' (see below).

Our data have increased the range of bacteria possessing multiple sets of *cbbL* genes. Among the reasons for the appearance of 'multi-copy' genes, duplication (with a probability of further selective loss of one copy) and lateral gene transfer are currently recognized. For almost identical copies, as in the case of two *cbbL* genes in *Cupriavidus* *necator* H16 and *Acidithiobacillus ferrooxidans* Fe1, a recent duplication event is suggested (Kusano *et al.*, 1991; Kusian *et al.*, 1995). For the *cbbL* copies in *Allochromatium vinosum*, with significant sequence divergence but a common GC composition and codon usage, a more ancient duplication event is hypothesized (Viale *et al.*, 1989; Kobayashi *et al.*, 1991). On the other hand, the presence of two *cbbL* copies in *Acidithiobacillus ferrooxidans* ATCC 23270 with significant differences in nucleotide sequence, GC ratios, and codon usage suggests lateral gene transfer as a mechanism of their origin (Heinhorst *et al.*, 2002).

Analysis of the three sets of RubisCO genetic clusters present in the genome of the hydrogen-oxidizing member of the *'Thiomicrospira* group', *Hydrogenovibrio marinus*, allows us to suggest the following method of their origin: the ancestor of this species, possessing the *cbbM* gene cluster, acquired *cbbLS-2* genes by lateral transfer, which, after duplication and rearrangement of other genes of the *cbbM* cluster, generated the *cbbL-1* gene cluster (Yoshizawa *et al.*, 2004).

In general our results are consistent with such a scenario. Taking into consideration evolutionary distances and codon usage, it might be suggested that, among modern autotrophs, cyanobacteria and not other photo- or chemoautotrophic bacteria could be the most probable donors of the *'Thiomicrospira* group' *cbbL-2* gene pool. The presence of distant but undoubtedly related *cbbL-1* and *cbbL-2* genes, and their identical GC content and codon usage in investigated SOB species, suggest the occurrence of a gene duplication event in the ancestral form of the *'Thiomicrospira* group'. Recent evolution of the ancestor that acquired all three types of RubisCO genes has resulted in selective loss of the *cbbM* and *cbbL-1* in its alkaliphilic descendants (genus *Thioalkalimicrobium*) and the *cbbL-1* gene in *Thiomicrospira pelophila*.

Such a loss might be a result of different catalytic properties of various forms of RubisCO. In particular, immunoblotting analysis revealed different expression of the three types of RubisCO genes in Hydrogenovibrio marinus depending on CO_2 content. *cbbM* is exclusively expressed at high CO_2 content (15%), both *cbbM* and *cbbL-1* are expressed at intermediate CO₂ concentrations, while expression of the *cbbL-2* gene starts at low levels of CO₂, approaching its recent atmospheric content, when all three forms are present (Yoshizawa et al., 2004). This example demonstrates that possession of isoenzymes with slightly different metabolic properties, such as three different forms of RubisCO in Hydrogenovibrio marinus, gives the bacterium a certain ecological advantage and flexibility. Assuming the close similarity of the genes encoding these forms in Hydrogenovibrio marinus and the Thiomicrospira species of the Thiomicrospira crunogena cluster, it seems likely that these SOB species used the same mechanism of adaptation to environmental changes. Remarkably high chemolithoautotrophic growth rates common for this SOB cluster might be a result of the presence of several forms of this key enzyme with variable catalytic properties. Depending on conditions, for example the ratio of CO_2 to O_2 , one or another form might have been underexpressed and, eventually, even dropped completely, as probably happened within the haloalkaliphilic genus *Thioalkalimicrobium*, possessing a single form I RubisCO best adapted to modern atmospheric conditions (Yoshizawa *et al.*, 2004). The anaerobic phototrophic SOB *Allochromatium vinosum* probably represents an intermediate stage in such an evolutionary course, possessing two different *cbbL* copies, one of which is practically not expressed (Kobayashi *et al.*, 1991).

The *cbbL-1* gene encoding RubisCO form I appears to be a reserve enzyme for the investigated SOB group. If this is so, the rate of non-synonymous nucleotide replacements in this gene must increase with decreased selection pressure. The complete lack of *cbbL-1* and *cbbM* in the genus *Thioalkalimicrobium* (obligate alkaliphiles) and *Thiomicrospira pelophila* (alkalitolerant bacterium) (see Sorokin & Kuenen, 2005) might have something to do with adaptation to high carbonate alkalinity: one of the possible explanations could be low actual CO₂ concentration at pH above 8.

Taxonomic implications

The use of phylogenetic analysis of genes other than those for 16S rRNA as molecular markers in bacterial taxonomy is not yet customary despite the obvious advantage of such information for genes encoding key metabolic enzymes vitally important for an organism's survival. However, inclusion of additional molecular markers might help to solve some complicated taxonomic and evolutionary problems, such as the current example of the 'Thiomicrospira group'. Both the 16S rRNA and RubisCO gene sequence analyses strongly support the necessity of taxonomic revision of this group, more specifically, dividing it into four genera within a new monophyletic family, the 'Thiomicrospiraceae'. The genera Thioalkalimicrobium (Sorokin et al., 2001) and Hydrogenovibrio (Nishihara et al., 1991) are sufficiently separated from Thiomicrospira physiologically and genetically, but Thiomicrospira has to be divided into two genera based on the Thiomicrospira crunogena and Thiomicrospira pelophila clusters, while 'Thiomicrospira denitrificans', a member of the Epsilonproteobacteria, should certainly be removed from the group. The possibility of including the symbiotic SOB into this new family should also be considered, but this requires more data on their physiology and RubisCO-based phylogeny.

ACKNOWLEDGEMENTS

This work was supported by the Russian Foundation for Basic Research (grant 05-04-48064).

REFERENCES

Boulygina, E. S., Kuznetsov, B. B., Marusina, A. I., Tourova, T. P., Kravchenko, I. K., Bykova, S. A., Kolganova, T. V. & Galchenko, V. F. (2002). The study of nucleotide sequences of *nifH* genes from some methanotrophic bacteria. *Microbiology* 71, 425–432.

Brinkhoff, T., Muyzer, G., Wirsen, C. O. & Kuever, J. (1999a). *Thiomicrospira kuenenii* sp. nov. and *Thiomicrospira frisia* sp. nov., two mesophilic obligately chemolithoautotrophic sulfur-oxidizing bacteria isolated from an intertidal mud flat. *Int J Syst Bacteriol* **49**, 385–392.

Brinkhoff, T., Muyzer, G., Wirsen, C. O. & Kuever, J. (1999b). *Thiomicrospira chilensis* sp. nov., a mesophilic obligately chemolithoautotrophic sulfur-oxidizing bacterium isolated from a *Thioploca* mat. *Int J Syst Bacteriol* **49**, 875–879.

Brinkhoff, T., Sievert, S. M., Kuever, J. & Muyzer, G. (1999c). Distribution and diversity of sulfur-oxidizing *Thiomicrospira* spp. at a shallow-water hydrothermal vent in the Aegean Sea (Milos, Greece). *Appl Environ Microbiol* **65**, 3843–3849.

Delwiche, C. F. & Palmer, J. D. (1996). Rampant horizontal transfer and duplication of RubisCO genes in eubacteria and plastids. *Mol Biol Evol* 13, 873–882.

Elsaied, H. & Naganuma, T. (2001). Phylogenetic diversity of ribulose-1,5-bisphosphate carboxylase/oxygenase large-subunit genes from deep-sea microorganisms. *Appl Environ Microbiol* 67, 1751–1765.

English, R. S., Williams, C. A., Lorbach, S. C. & Shively, J. M. (1992). Two forms of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Thiobacillus denitrificans. FEMS Microbiol Lett* 94, 111–119.

Felsenstein, J. (1981). Evolutionary trees from DNA sequences: a maximum likelihood approach. J Mol Evol 17, 368–376.

Felsenstein, J. (1993). PHYLIP (phylogeny inference package), version 3.53c. Department of Genetics, University of Washington, Seattle, USA.

Fennoy, S. L. & Bailey-Serres, J. (1993). Synonymous codon usage in *Zea mays* L. nuclear genes is varied by levels of C and G-ending codons. *Nucleic Acids Res* 21, 5294–5300.

Fitch, W. M. (1971). Toward defining the course of evolution: minimum change for a specified tree topology. *Syst Zool* 20, 406–416.

Fitch, W. M. & Margoliash, E. (1967). Construction of phylogenetic trees. *Science* 155, 279–284.

Gibson, J. L. & Tabita, F. R. (1977a). Different molecular forms of ribulose-1,5-bisphosphate carboxylase from *Rhodopseudomonas* sphaeroides. J Biol Chem 252, 943–949.

Gibson, J. L. & Tabita, F. R. (1977b). Isolation and preliminary characterization of two forms of ribulose 1,5-bisphosphate carboxylase from *Rhodopseudomonas capsulata*. J Bacteriol 132, 818–823.

Heinhorst, S., Baker, S. H., Johnson, D. R., Davies, P. S. & Cannon, G. C. (2002). Two copies of form I RubisCO genes in *Acidithiobacillus ferrooxidans* ATCC 23270. *Curr Microbiol* **45**, 115–117.

Hernandez, J. M., Baker, S. H., Lorbach, S. C., Shively, J. M. & Tabita, F. R. (1996). Deduced amino acid sequence, functional expression, and unique enzymatic properties of the form I and form II ribulose bisphosphate carboxylase/oxygenase from the chemo-autotrophic bacterium *Thiobacillus denitrificans. J Bacteriol* 178, 347–356.

Hugler, M., Wirsen, C. O., Fuchs, G., Taylor, C. D. & Sievert, S. M. (2005). Evidence for autotrophic CO_2 fixation via the reductive tricarboxylic acid cycle by members of the ε subdivision of proteobacteria. *J Bacteriol* 187, 3020–3027.

Jannasch, H. W., Wirsen, C. O., Nelson, D. C. & Robertson, L. A. (1985). *Thiomicrospira crunogena* sp. nov., a colorless sulfur-oxidizing bacterium from a deep-sea hydrothermal vent. *Int J Syst Bacteriol* **35**, 422–424.

Kobayashi, H., Viale, A. M., Takabe, T., Akazawa, T., Wada, K., Shinozaki, K., Kobayashi, K. & Sugiura, M. (1991). Sequence and

expression of genes encoding the large and small subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase from *Chromatium vinosum*. *Gene* **97**, 55–62.

Kuenen, J. G. & Veldkamp, H. (1972). Thiomicrospira pelophila, gen. n., sp. n., a new obligately chemolithotrophic colourless sulfur bacterium. Antonie van Leeuwenhoek 38, 241–256.

Kusano, T., Takeshima, T., Inoue, C. & Sugawara, K. (1991). Evidence for two sets of structural genes coding for ribulose bisphosphate carboxylase in *Thiobacillus ferrooxidans*. J Bacteriol 173, 7313–7323.

Kusian, B., Bednarski, R., Husemann, M. & Bowien, B. (1995). Characterization of the duplicate ribulose-1,5-bisphosphate carboxylase genes and *cbb* promoters of *Alcaligenes eutrophus*. *J Bacteriol* 177, 4442–4450.

Medigue, C., Rouxel, T., Vigier, P., Henaut, A. & Danchin, A. (1991). Evidence for horizontal gene transfer in *Escherichia coli* speciation. *J Mol Biol* 222, 851–856.

Musto, H., Romero, H. & Rodriguez-Maseda, H. (1998). Heterogeneity in codon usage in the flatworm *Schistosoma mansoni*. J Mol Evol 46, 159–167.

Nishihara, H., Igarashi, Y. & Kodama, T. (1991). *Hydrogenovibrio marinus* gen. nov., sp. nov., a marine obligately chemolithoautotrophic hydrogen-oxidizing bacterium. *Int J Syst Bacteriol* **41**, 130–133.

Nishihara, H., Yaguchi, T., Chung, S. Y., Suzuki, K., Yanagi, M., Yamasato, K., Kodama, T. & Igarashi, Y. (1998). Phylogenetic position of an obligately chemoautotrophic, marine hydrogen-oxidizing bacterium, *Hydrogenovibrio marinus*, on the basis of 16S rRNA gene sequences and two form I RubisCO gene sequences. *Arch Microbiol* 169, 364–368.

Ohtaka, C. & Ishikawa, H. (1993). Accumulation of adenine and thymine in *groE*-homologous operon of an intracellular symbiont. *J Mol Evol* **36**, 121–126.

Paoli, G. C., Morgan, N. S., Tabita, F. R. & Shively, J. M. (1995). Expression of the *cbbLcbbS* and *cbbM* genes and distinct organization of the *cbb* Calvin cycle structural genes of *Rhodobacter capsulatus*. *Arch Microbiol* 164, 396–405.

Robertson, L. A. & Kuenen, J. G. (1999). The colourless sulfur bacteria. In *The Prokaryotes: an Evolving Electronic Resource for the Microbiological Community*, 3rd edn, Release 3.0 21 May 1999 (http:// link.springer-ny.com/link/service/books/10125/). New York: Springer.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* 4, 406–425.

Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.

Shively, J. M., Devore, W., Stratford, L., Porter, L., Medlin, L. & Stevens, S. E. (1986). Molecular evolution of the large subunit of ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO). *FEMS Microbiol Lett* **37**, 251–257.

Sorokin, D. Y. & Kuenen, J. G. (2005). Alkaliphilic chemolithotrophs from soda lakes. *FEMS Microbiol Ecol* 52, 287–295.

Sorokin, D. Y., Lysenko, A. M., Mityushina, L. L., Tourova, T. P., Jones, B. E., Rainey, F. A., Robertson, L. A. & Kuenen, G. J. (2001). *Thioalkalimicrobium aerophilum* gen. nov., sp. nov. and *Thioalkalimicrobium sibiricum* sp. nov., and *Thioalkalivibrio versutus* gen. nov., sp. nov., *Thioalkalivibrio nitratis* sp. nov. and *Thioalkalivibrio denitrificans* sp. nov., novel obligately alkaliphilic and obligately chemolithoautotrophic sulfur-oxidizing bacteria from soda lakes. *Int J Syst Evol Microbiol* **51**, 565–580.

Sorokin, D. Y., Gorlenko, V. M., Tourova, T. P., Tsapin, A. I., Nealson, K. H. & Kuenen, G. J. (2002). *Thioalkalimicrobium cyclicum* sp. nov.

and *Thioalkalivibrio jannaschii* sp. nov., novel species of haloalkaliphilic, obligately chemolithoautotrophic sulfur-oxidizing bacteria from hypersaline alkaline Mono Lake (California). *Int J Syst Evol Microbiol* **52**, 913–920.

Spiridonova, E. M., Berg, I. A., Kolganova, T. V., Ivanovsky, R. N., Kuznetsov, B. B. & Tourova, T. P. (2004). An oligonucleotide primer system for amplification of the ribulose-1,5-bisphosphate carboxylase/oxygenase genes of bacteria of various taxonomic groups. *Microbiology* **73**, 316–325.

Stoner, M. T. & Shively, J. M. (1993). Cloning and expression of the D-ribulose-1,5-bisphosphate carboxylase/oxygenase form II gene from *Thiobacillus intermedius* in *Escherichia coli. FEMS Microbiol Lett* **107**, 287–292.

Takai, K., Hirayama, H., Nakagawa, T., Suzuki, Y., Nealson, K. H. & Horikoshi, K. (2004). *Thiomicrospira thermophila* sp. nov., a novel microaerobic, thermotolerant, sulfur-oxidizing chemolithomixo-troph isolated from a deep-sea hydrothermal fumarole in the TOTO caldera, Mariana Arc, Western Pacific. *Int J Syst Evol Microbiol* **54**, 2325–2333.

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **9**, 3251–3270.

Tourova, T. P., Spiridonova, E. M., Berg, I. A., Kuznetsov, B. B. & Sorokin, D. Yu. (2005). Phylogeny of ribulose-1,5-bisphosphate carboxylase/oxygenase genes in haloalkaliphilic obligately autotrophic sulfur-oxidizing bacteria of the genus *Thioalkalivibrio*. *Microbiology* 74, 321–328.

Uchino, Y. & Yokota, A. (2003). 'Green-like' and 'red-like' RubisCO *cbbL* genes in *Rhodobacter azotoformans*. *Mol Biol Evol* 20, 821–830.

Van de Peer, Y. & De Wachter, R. (1994). TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft Windows environment. *Comput Applic Biosci* 10, 569–570.

Viale, A. M., Kobayashi, H. & Akazawa, T. (1989). Expressed genes for plant-type ribulose 1,5-bisphosphate carboxylase/oxygenase in the photosynthetic bacterium *Chromatium vinosum*, which possesses two complete sets of the genes. *J Bacteriol* 171, 2391–2400.

Watson, G. M. F. & Tabita, F. R. (1997). Microbial ribulose 1,5bisphosphate carboxylase/oxygenase: a molecule for phylogenetic and enzymological investigation. *FEMS Microbiol Lett* 146, 13–22.

Wood, A. P. & Kelly, D. P. (1989). Isolation and characterization of *Thiobacillus thyasiris* sp. nov., a novel marine facultative autotroph and the putative symbiont of *Thyasira flexuosa*. Arch Microbiol 152, 160–166.

Wood, A. P. & Kelly, D. P. (1993). Reclassification of *Thiobacillus thyasiris* as *Thiomicrospira thyasirae* comb. nov., an organism exhibiting pleomorphism in response to environmental conditions. Arch Microbiol 159, 45–47.

Yaguchi, T., Chung, S. Y., Igarashi, Y. & Kodama, T. (1994). Cloning and sequencing of the L2 form of RubisCO from a marine obligately autotrophic hydrogen-oxidizing bacterium, *Hydrogenovibrio marinus* strain MH-110. *Biosci Biotechnol Biochem* 58, 1733–1737.

Yang, Z. (2000). PAML (phylogenetic analysis by maximum likelihood), version 3.0. London: University College London.

Yang, Z. & Nielsen, R. (2000). Estimating synonymous and nonsynonymous substitution rates under realistic evolutionary models. *Mol Biol Evol* 17, 32–43.

Yoshizawa, Y., Toyoda, K., Arai, H., Ishii, M. & Igarashi, Y. (2004). CO₂responsive expression and gene organization of three ribulose-1,5bisphosphate carboxylase/oxygenase enzymes and carboxysomes in *Hydrogenovibrio marinus* strain MH-110. *J Bacteriol* **186**, 5685–5691.