Genes and pathways for CO₂ fixation in the obligate, chemolithoautotrophic acidophile, Acidithiobacillus ferrooxidans, Carbon fixation in A. ferrooxidans

Mario Esparza¹,4, Juan Pablo Cárdenas¹, Botho Bowien³, Eugenia Jedlicki², David S Holmes¹*

Abstract

Background: Acidithiobacillus ferrooxidans is chemolithoautotrophic γ-proteobacterium that thrives at extremely low pH (pH 1-2). Although a substantial amount of information is available regarding CO₂ uptake and fixation in a variety of facultative autotrophs, less is known about the processes in obligate autotrophs, especially those living in extremely acidic conditions, prompting the present study.

Results: Four gene clusters (termed cbb1-4) in the A. ferrooxidans genome are predicted to encode enzymes and structural proteins involved in carbon assimilation via the Calvin-Benson-Bassham (CBB) cycle including form I of ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO, EC 4.1.1.39) and the CO₂-concentrating carboxysomes. RT-PCR experiments demonstrated that each gene cluster is a single transcriptional unit and thus is an operon. Operon cbb1 is divergently transcribed from a gene, cbbR, encoding the LysR-type transcriptional regulator CbbR that has been shown in many organisms to regulate the expression of RubisCO genes. Sigma⁷⁰-like -10 and -35 promoter boxes and potential CbbR-binding sites (T-N₁₁-A/TNA-N₁₁-TNA) were predicted in the upstream regions of the four operons. Electrophoretic mobility shift assays (EMSAs) confirmed that purified CbbR is able to bind to the upstream regions of the cbb1, cbb2 and cbb3 operons, demonstrating that the predicted CbbR-binding sites are functional in vitro. However, CbbR failed to bind to the upstream region of the cbb4 operon that contains cbbP, encoding phosphoribulokinase (EC 2.7.1.19). Thus, other factors not present in the assay may be required for binding or the region lacks a functional CbbR-binding site. The cbb3 operon contains genes predicted to encode anthranilate synthase components I and II, catalyzing the formation of anthranilate and pyruvate from chorismate. This suggests a novel regulatory connection between CO₂ fixation and tryptophan biosynthesis. The presence of a form II RubisCO could promote the ability of A. ferrooxidans to fix CO₂ at different concentrations of CO₂.

Conclusions: A. ferrooxidans has features of cbb gene organization for CO₂-assimilating functions that are characteristic of obligate chemolithoautotrophs and distinguish this group from facultative autotrophs. The most conspicuous difference is a separate operon for the cbbP gene. It is hypothesized that this organization may provide greater flexibility in the regulation of expression of genes involved in inorganic carbon assimilation.

Background

Acidithiobacillus ferrooxidans is a mesophilic, obligately chemolithoautotrophic, γ-proteobacterium that gains energy and reducing power from the oxidation of ferrous iron and reduced inorganic sulfur compounds (RISCs) [1]. It grows optimally at pH 2, although growth as low as pH 1 has been reported [2]. The microorganism is a key player in the solubilization of copper in industrial bioleaching operations and makes an important contribution to the biogeochemical cycling of nutrients and metals in pristine and manmade acidic environments. In such environments, CO₂ would be expected to exist preferentially as a dissolved gas in
equilibrium with the atmosphere and not in the bicarbonate form typically found at circum-neutral pHs [3].

*A. ferrooxidans* has previously been shown [4,5] to have candidate genes (*cbbL* and *cbbS*) for the large and small subunits of ribulose-1,5-bisphosphate carboxylase/oxygenase (RubisCO, EC 4.1.1.39) that catalyses CO₂ fixation by the Calvin-Benson-Bassham (CBB) cycle in many organisms [6]. *cbbL* and *cbbS* are linked to genes predicted to encode carboxysome shell proteins [7] and a divergently transcribed gene encoding the LysR-type transcription regulator CbbR [4]. The intergenic region between *cbbR* and *cbbL* is predicted to harbor binding sites for CbbR [4]. In addition, microarray transcript profiling experiments have detected differential expression of several genes in *A. ferrooxidans* potentially involved in the CBB cycle depending on the growth substrate used [8].

These observations taken together, suggest that, in *A. ferrooxidans*, CbbR can regulate the expression of RubisCO and the carboxysome genes and therefore is linked to genes involved in the CBB cycle depending on the growth substrate used [8].

Methods

Bacterial strains and culture conditions

Information regarding bacterial strains and plasmids used in this study is provided in Table 1. *A. ferrooxidans* was cultured in 9 K medium (adjusted to pH 3.5 with H₂SO₄) containing 5 g/l elemental sulfur at 30°C under aerobic conditions on a rotary shaker at 150 rpm as described previously [21]. *Escherichia coli* harboring plasmids was grown at 37°C in LB broth with ampicillin (Amp: 100 μg/ml).

General DNA techniques and sequencing of DNA

*A. ferrooxidans* cultures were centrifuged at 800 × g to remove solid sulfur precipitates prior to cell harvest. Unattached cells were pelleted at 8000 × g for 10 min. The cell pellet was resuspended in 9 K salt solution for washing and washed cells were collected by centrifugation at 8000 × g for 10 min as described previously [21]. Standard procedures [22] were employed to isolate genomic and plasmid DNA from bacteria, to transform plasmid DNA into *E. coli*, and for general DNA handling. Restriction endonucleases and DNA-modifying enzymes were used as recommended by the manufacturers. Plasmid DNA was prepared by means of the QiAprep Spin Mini Kit (Qiagen). Polymerase chain reaction (PCR) products were amplified with Taq DNA polymerase (Fermentas) and purified from agarose gels using the QiaEx DNA Purification Kit (Qiagen). Each PCR reaction contained in a volume of 25 μl 1 ng of template DNA, 0.5 μM of required primers and 0.2 mM of each deoxyribonucleotide in 1× PCR buffer containing 1.5 mM MgCl₂ (Fermentas). PCR conditions were as follows: initial denaturing step at 95°C for 5 min followed by 30 amplification cycles (denaturation at 95°C for 30 s, annealing at the appropriate temperature depending on the specific primer pairs for 20 s, elongation at 72°C) and a final elongation step at 72°C for 10 min. DNA sequencing of pBAD-\(cbbR\) was carried out by the Göttingen Genomics Laboratory (Göttingen, Germany).

Isolation of RNA and RT-PCR

Total RNA was isolated from cells of *A. ferrooxidans* grown to mid-log phase in 9 K medium supplemented with sulfur, as described previously [23]. The RNA...
preparation was treated with DNase I (Fermentas) before proceeding with the cDNA synthesis step. One microgram of total cellular RNA was used for each reaction. Reverse transcription-PCR (RT-PCR) was performed on purified RNA using the One-Step RT-PCR kit (Qiagen). The sequences of the RT and PCR primers used are provided in Table 2. As controls, reactions were carried out that included RNA but lacked reverse transcriptase to assess genomic DNA contamination and that lacked RNA but contained 1 ng of genomic DNA.

Cloning and expression of cbbR
A DNA fragment corresponding to the coding region of cbbR of A. ferrooxidans was amplified by PCR using primers (Integrated DNA Technologies) cbbRfw and cbbRrev (Table 2). The amplified product was cloned into the expression vector pBAD-TOPO (Invitrogen) according to the manufacturer’s instruction. The resulting plasmid pBAD-cbbR was introduced by electroporation into E. coli TOP10 (Invitrogen) competent cells [22]. E. coli was grown at 37°C in 10 ml LB containing 100 μg/ml ampicillin to an OD600 of 0.8. Overproduction of the recombinant His6-tagged CbbR protein was initiated by adding arabinose to a final concentration of 0.1% (w/v) with continued shaking at 200 rpm for 12 h.

Purification of CbbR
Cells from 1.5 l of induced culture were harvested by centrifugation (8,000 × g for 10 min at 4°C) and at -20°C. After thawing the cell pellet was resuspended in 40 ml denaturing buffer containing 6 M guanidine-HCl, 100 mM NaH2PO4 and 10 mM Tris-HCl, pH 8.0, and incubated at room temperature with continuous stirring for about 30 min until inclusion body proteins were solubilized. Any remaining insoluble material was removed by centrifugation at 18,000 × g and 7°C for 20 min. The resulting supernatant was filtered through a 0.45-μm membrane and the recombinant protein subsequently purified by affinity chromatography on a 2.5 cm 0.5 M membrane and the recombinant protein sub-

fractions were monitored at 280 nm. Fractions with the highest protein content were pooled, dialysed twice against 50 mM HEPES-NaOH, pH 7.8 containing 200 mM KCl, 10 mM MgCl2, 1 mM dithiothreitol, 0.05 mM phenylmethylsulfonyl fluoride and 50% (w/v) glycerol. The final protein concentration was 4 mg/ml. Protein preparations were analyzed by SDS-polyacrylamide gel electrophoresis in 12% (w/v) polyacrylamide slab gels under reducing conditions in the presence of 100 mM β-mercaptoethanol. Gels were stained with Coomassie Brilliant Blue R-250. Protein contents were determined using the method of Bradford [24], with bovine serum albumin as a standard. CbbR was stored at -20°C.

Production of antisera to CbbR
Multiple intradermal injections were applied to immunize a female Californian giant rabbit (3.0 kg) as described by [25]. A fresh CbbR preparation (0.5 ml; 1 mg/ml) was emulsified in one volume of complete Freund adjuvant (Commonwealth Serum Laboratories, Melbourne, Australia). The emulsion was prepared under aseptic conditions and 1.0 ml was initially injected into four sites on the back of the animal. Booster injections were given in the same way 75 days after the primary immunization, except that incomplete Freund adjuvant was used. The immune response was monitored by Western Blotting assays with serum separated from test blood samples (1.0 to 2.0 ml) that were obtained from an ear vein every 15 to 20 days after each immunization.

Electrophoretic mobility shift assays (EMSA)
DNA fragments containing the four potential cbb operon promoter regions were amplified by PCR and simultaneously biotinylated using the biotin 5’-labelled primers (Table 2). DNA-binding assays were performed at 30°C in a final volume of 17 μl containing 12 mM HEPES-NaOH, pH 7.9, 4 mM Tris-HCl, pH 7.9, 1 mM EDTA, 60 mM KCl, 1 mM dithiothreitol, 10% (w/v) glycerol, 5 μg/μl of bovine serum albumin and 2 μg/μl of poly(dI-dC). The indicated amount of CbbR protein (~290 μM) was incubated with the biotin end-labeled target DNA (20 pmol) for 15 min. A 50-fold excess of unlabeled DNA probe was used to challenge the labeled probe. In supershift experiments, a 1:500 dilution of CbbR-specific antiserum was added to the reaction after DNA binding of CbbR and incubated for an additional 15 min. After the binding reactions, samples were loaded onto a low-ionic strength non-denaturing polyacrylamide gel (4.8% [w/v], which had been prerun at a constant current of 200 mA for more than 90 min, and electrophoresed at 150 mA for about 60 min in 0.5× TBE buffer (89 mM Tris base, 89 mM boric acid and 2 mM EDTA). The separated CbbR-DNA complexes were
Table 2 Sequences of primers used in co-transcription RT-PCR assays, cloning experiments and EMSA assays

| Primers used in RT-PCR assays | Primers used in EMSA assays and cloning experiments |
|-------------------------------|-----------------------------------------------|
| **cbb1 operon**               |                                               |
| **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** | **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** |
| 1 | cbbR | CACGCGCTGGTGCTCGA | 2 | cbbR | CTAGACTTTTTTATGCGGCTGTTT |
| 3 | cbbL1 | CTGCGAATCGTGCTCGG | 4 | csoS2 | GCGCAGAAGCCAGGACCTT |
| 5 | csoS2 | CCTATGCTGGCGCTGCCAAC | 6 | csoS3 | GTGCATGACGCCAGGCCC |
| 7 | csoS3 | GTCAGCGGCTCAAGGCCC | 8 | csoS1A | GCGCCTCGTATCG |
| 9 | csoS1B | GGGACAGATCTCAGGGG | 11 | parA | AGTGAAGCCCGACGACGCC |
| 10 | bfrA | CGCGCAGAAGAGTTACAAGCCTTG | 12 | parA | CTGTCATGACGGGATCG |
| 13 | parA | CTGATCGAACCCTGAGGATCGG | 14 | parA | GTGCGATTTGATATGCCG |
| 15 | parA | GTGCTGCGGTTGAAGGGGT | 16 | parA | CAGGACGTTCCTGGGCTG |
| 18 | parA | GTGGAGTTCGATAATGGCGGAG | 17 | parA | AAGCACCTACGCGCGCTG |
| 21 | bioDa | CACGCGCTGGTGCTCGA | 22 | bioDa | GAGGACGTTCCTGGGCTG |
| **cbb2 operon**               |                                               |
| **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** | **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** |
| 1 | tatCa | ACGACGGCGTCTAGAACCGCC | 2 | cbbL2 | CTCCGGCCTCATACTTTTTTACGG |
| 3 | cbbL2 | CATCGAAGAGGGCGTGAGGCC | 4 | cbbL2 | GCGCAGAAGCCAGGACCTT |
| 7 | cbbS2 | GTGCATGACGCCAGGCCC | 8 | cbbS2 | GCGCAGAAGCCAGGACCTT |
| 11 | cbbS2 | GCGACGGCGTCTAGAACCGCC | 12 | cbbS2 | GCGCAGAAGCCAGGACCTT |
| 14 | cbbS2 | GCTCCGGCCTCATACTTTTTTACGG |
| **cbb3 operon**               |                                               |
| **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** | **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** |
| 1 | hyp4a | TACGAAAGGGCCCTCCCGC | 2 | hyp3 | CGACGCGACATGGGATCTT |
| 3 | hyp3 | CGCGGCAACGCGCCTGCCAT | 4 | hyp3 | CGACGCGACATGGGATCTT |
| 5 | cbbT | ATCGGAAATCGCTCGGCCC | 6 | cbbT | TTGGGAGCGGACAACACCA |
| 7 | cbbK | CCTACGATCGCTCGGCCC | 9 | cbbK | AAGCCTACGCGCGCTG |
| 8 | pykA | TTGATCTCCTCAACCATGCG | 11 | pykA | GACGCTGACGGGATCTG |
| 10 | yfeAa | GCGGAGCCTTGTCCCTCGG | 14 | yfeAa | GCCGCGCCTGACCCAGC |
| 12 | yfeAa | GCCGCGCCTGACCCAGC |
| **cbb4 operon**               |                                               |
| **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** | **Number** | **Gene** | **Forward primer (5′-3′)** | **Reverse primer (5′-3′)** |
| 1 | ompAa | GGATTTTCCTATTTTTGCGGG | 3 | sahA | GCCGATCGACGGGACCTT |
| 2 | metK | TTGGGAGCGGACAACACCA | 5 | metK | AAGCCTACGCGCGCTG |
| 4 | sahA | CTTGCCGCGGCGCTCGG | 7 | cbbP | GACGCTGACGGGATCTG |
| 6 | metF | CACGCGCTCCTCGTCTGAGGAC | 9 | cbbP | GCCGCGCCTGACCCAGC |
| 8 | cbbP | AGTGGGCGCGGCGCTGAGGAC | 10 | cbbP | GCCGCGCCTGACCCAGC |

**Primers used in EMSA assays and cloning experiments**

| Letter designation | Forward primer (5′-3′) | Reverse primer (5′-3′) |
|-------------------|------------------------|------------------------|
| (a) Pcb1 | CGGACGCGCAAGATCTTGGATTTT | (b) Pcb1 | CTGCCGCTCCTATTTTTTAA |
| (c) cbbR | TCTATCGCTAGCAGACCTT | (d) cbbR | GCGGCATCCCTCTCTCCCAT |
| (e) Pcb2 | CGGACGCGCAAGATCTTGGATTTT | (f) Pcb2 | CGGACGCGCAAGATCTTGGATTTT |
| (g) Pcb3 | CATTGACACAGGATCTTGGATTTT | (h) Pcb3 | ATCGGACGCTGACGGGATCTG |
| (i) Pcb4 | TTTGGGCGCGGCGCTGAGGAC | (j) Pcb4 | GCCGCGCCTGACCCAGC |

*Genes lie outside the operons and are not shown in Figure 2.*
electrophoretically transferred from the gel onto a nylon membrane in semi-dry blotting apparatus (Biometa, Göttingen). CbbR-DNA binding was detected using a streptavidin-horseradish peroxidase conjugate and a chemiluminescent substrate (Pierce) followed by autoradiography.

Bioinformatic analyses

Metabolic pathways involved in CO₂ assimilation were retrieved from KEGG http://www.genome.ad.jp/kegg/. Protein sequences derived from known genes involved in CO₂ assimilation were used as query sequences to search the genome sequence of *A. ferrooxidans* ATCC 23270, using TBLastN and BlastP, respectively, with default parameters. When a prospective candidate gene was identified, its predicted protein sequence was then used to formulate a BlastP http://www.ncbi.nlm.nih.gov/search of the nonredundant database at NCBI. Only bidirectional best hits were accepted as evidence for putative orthologs. Candidate genes and their translated proteins were further characterized employing the following bioinformatic tools: ClustalW [26] for primary structure similarity relations, PSI-PRED [27] for secondary structure predictions, Prosite [28] for motif predictions, ProDom [29] and Pfam [30] for domain predictions. Information regarding the organization of genes in *A. ferrooxidans* was obtained from [2]. Logos were generated using the web-based application available at http://weblogo.berkeley.edu/logo.cgi. The height of each letter in bits corresponds to its relative abundance at each position. Promoters of the σ⁷₀-type and rho-independent transcriptional stops were predicted for operons *cbb1-4* using the programs BPROM http://www.softberry.com and Transtern [31], respectively.

The organization of gene clusters in facultative and obligate autotrophs involved in the CBB cycle was derived from information available in IMG-JGI http://www.jgi.doe.gov/ and MicrobesOnline http://www.microbesonline.org/, with additional information added for *H. marinus* [18] and *A. ferrooxidans*, *Acidithioacillus caldus* and *Acidithioacillus thiooxidans* (this study). The phylogenetic cladogram of these bacteria was constructed from 16 S rRNA sequences obtained from KEGG Orthology K01977 http://www.genome.jp/kegg/ko.html and from GenBank http://www.ncbi.nlm.nih.gov/ for *A. caldus* (GI454888), *A. thiooxidans* (GI454888) and *H. marinus* (GI3882094). 16 S rRNA alignments were carried out using ClustalW and the cladogram was constructed by the NJ method using the program MEGA 4.0 [32]. The robustness of the tree was evaluated by bootstrapping using 1000 replicas. The tree was rooted using the 16 S rRNA of the *e*-proteobacterium *Helicobacter pylori*.

Results, Discussion and Conclusions

The genome of *A. ferrooxidans* ATCC 23270 encodes CbbR, a LysR-type transcription factor

A gene *cbbR* was predicted in the genome of *A. ferrooxidans* ATCC 23270 (type strain) that potentially encodes a protein with significant amino acid sequence similarity and domain structure to other well-documented CbbRs of the LysR family of transcription factors (Additional file 1). *cbbR* is divergently transcribed from *cbbL1*, a gene predicted to encode the large subunit of form I RubisCO. The genetic linkage between *cbbR* and *cbbL1* is known to be conserved in a number of autotrophic bacteria that fix CO₂ via the CBB cycle such as *Acidithiobacillus ferrooxidans* Fe1 [4], *Hydrogenophilus thermoluteolus* [33], *Nitrosomonas europaea* [19], *Rhodobacter sphaeroides* [34], *Rhodobacter capsulatus* [35], *R. eutropha* H16 [36], *Rhodospirillum rubrum* [17], *Thiobacillus denitrificans* [14] and *Xanthobacter flavus* [9]. We here extend this list to include: *Alkalilimnicola ehrlichii*, *Halorhodospira halophila*, *Methyllobium petroleiphilum*, *Nitrobacter winogradskyi*, *Nitrosococcus oceani*, *Nitrosospira multiformis*, *Thiomicrospira crunobena* and *Xanthobacter autotrophicus* (Additional file 2).

The *cbbR-cbbL1* intergenic region of *A. ferrooxidans* strain Fe1 has been shown to contain divergent σ⁷₀-type promoters and to exhibit two CbbR binding sites that partially overlap these promoters ([4], Figure 1A). The binding sites conform to the pseudo-palindromic motif TNA-N₇-TNA [13] that is a subset of the consensus LysR-type transcription factor binding site T-N₁₁₋₁₋₁ [37]. Logos were derived from a multigenome comparison of the *cbbR-cbbL1* intergenic region of a number of bacteria (Additional file 3) and were aligned with the CbbR sites of *A. ferrooxidans* strain Fe1, allowing the prediction of the CbbR binding sites of *A. ferrooxidans* ATCC 27230 (Figure 1B and 1C).

Organization and expression of gene clusters predicted to be involved in CO₂ fixation and associated pathways of central carbon metabolism

A cluster of 16 genes, termed *cbb1*, was predicted to be involved in CO₂ fixation. RT-PCR experiments showed that *cbb1* is transcribed as a single unit and thus can be considered to be an operon (Figure 2A). Operon *cbb1* consists of *cbbL1* and *cbbS1*, potentially encoding the large and small subunits of form I RubisCO, seven *cso* genes predicted to be involved in α-carboxysome formation, two genes (*cbbQ1* and *cbbO1*) presumed to be involved in RubisCO activation and *cbbA*, potentially encoding a fructose-1,6-bisphosphate aldolase. Gene descriptions are provided in Table 3.

Three additional gene clusters termed *cbb2* (four genes), *cbb3* (twelve genes) and *cbb4* (five genes) were
identified that are predicted to encode functions related to CO₂ fixation and central carbon metabolism (Table 3). RT-PCR experiments revealed that gene clusters \textit{cbb2}, \textit{cbb3} and \textit{cbb4} are transcribed as single units, respectively, and thus constitute operons (Figure 2B-D). \textit{cbb2} contains genes (\textit{cbbL2} and \textit{cbbS2}) encoding additional copies of the large and small subunit of form I\textit{Aq} RubisCO and associated RubisCO activation genes (\textit{cbbQ2} and \textit{cbbO2}) (Figure 2B). The deduced amino acid sequences of these genes are similar but not identical to the corresponding proteins encoded in the \textit{cbb1} operon; \textit{CbbL1} and \textit{CbbL2} exhibit 84\% amino acid sequence identity, whereas \textit{CbbS1} and \textit{CbbS2} share 56\% identity and \textit{CbbQ1} and \textit{CbbO1} have 84\% and 59\% identity with \textit{CbbQ2} and \textit{CbbO2}, respectively.

Genes predicted to be encoded by operons \textit{cbb3} and \textit{cbb4} are listed in Table 3 and their organization within these operons is shown in Figure 2.

The two enzymes that are unique to the CBB cycle are RubisCO (encoded by operons \textit{cbb1} and \textit{cbb2}) and phosphoribulokinase (encoded by operon \textit{cbb4}). RuBisCO catalyzes the first step of the cycle, the carboxylation of ribulose-1,5-bisphosphate (RuBP) with CO₂. Phosphoribulokinase catalyzes the last step of the cycle which is the regeneration of the CO₂ acceptor molecule, RuBP, by phosphorylation of ribulose 5-phosphate with ATP. Other steps of the cycle, encoded in operon \textit{cbb3}, are catalyzed by enzymes common to glycolytic and gluconeogenic pathways in central carbon metabolism [8,36].

Promoters of the \(\sigma^{70}\)-type and rho-independent transcriptional stops were predicted for operons \textit{cbb1-4} (Figure 2). In addition, potential CbbR-binding sites were identified in the four operons based on the detection of conserved TNA-N\(_7\)-TNA and T-N\(_{11}\)-A motifs described above for operon \textit{cbb1} (Figure 2).

\textbf{CbbR binds \textit{in vitro} to the predicted \(\sigma^{70}\)-like promoter regions of operons \textit{cbb1-4}}

Binding of CbbR to DNA fragments containing the predicted promoters of the four operons \textit{cbb1-4} was evaluated \textit{in vitro} by electrophoretic mobility shift assays (EMSAs). For this purpose the \textit{cbbR} gene was cloned and expressed in \textit{E. coli}. Purified CbbR was used to prepare antisera (anti-CbbR antibodies) whose activity was checked by Western blotting against purified CbbR (data not shown). Biotin-labeled promoter DNA for the EMSA assays was prepared by PCR using primers specified in Table 2 and whose locations within the four operons are shown in Figure 2.

Results showed that CbbR was able to retard the promoter regions of the \textit{cbb1}, \textit{cbb2} and \textit{cbb3} operons but not the \textit{cbb4} operon (Figure 3). When a 50-fold molar excess of unlabelled fragment was included in the binding assay retardation of the labelled fragments was abolished. Furthermore, the addition of anti-CbbR antibodies to the reaction produced a supershift in migration, indicating that the shift was caused specifically by the binding of CbbR.

Binding of CbbR to the predicted promoter regions of operons \textit{cbb1-3} suggests that it is involved in their regulation. The reason for the failure of CbbR to retard the DNA fragment containing the predicted promoter of the \textit{cbb4} operon is not known. Perhaps this fragment
requires the presence of additional factors for CbbR binding that are not present in the in vitro cocktail used for the EMSA analysis. Alternatively, the predicted CbbR binding site is not functional.

**Gene organization of the cbb operons**

The cbb3 operon includes not only genes involved in carbon assimilation but also harbors genes with similarity to trpE and trpG that are predicted to encode the components I and II of anthranilate synthase, the first enzyme of the tryptophan biosynthesis pathway. Anthranilate synthase catalyzes the conversion of chorismate to anthranilate with the concomitant release of pyruvate [38,39]. In some cases, this conversion can be accomplished by TrpE alone [40].

In order to determine if the association between trpEG and the cbb genes is restricted to *A. ferrooxidans*, an examination of gene organization was carried out in all sequenced genomes of facultative and obligate autotrophic proteobacteria. Twenty-six proteobacterial organisms (11 α-, 7 β- and 8 γ-) were analyzed, including 10 obligate autotrophs. Linkage between trpE/G and cbbE and/or cbbZ was found in all sequenced obligate autotrophs, all of which belong to the β- or

---

**Figure 2** Organization and co-transcription of four cbb gene clusters in *A. ferrooxidans* ATCC 23270. (A) cbb1 (B) cbb2 (C) cbb3 and (D) cbb4. The following are represented in each of the panels A-E: (a) nucleotide sequences of the predicted α-35-like promoter region (-10 and -35 sites in italics) and potential CbbR-binding sites in grey boxes with the LysR-type TNA-N7-TNA and T-N11-A consensus binding sites in bold letters, (b) gene organization of the respective operons with predicted rho-independent transcriptional stop sites indicated as stem-loop symbols, (c) locations of PCR primers used for RT-PCR experiments (indicated by numbers) or EMSA assays (indicated by letters) and (d) gel electrophoresis of fragments amplified by RT-PCR using purified cellular RNA as template. A 1-kb scale bar is shown. One of the T-N11-A consensus binding sites in the cbb4 operon is part of a larger pseudo-palindrome indicated by inverted arrows. Predicted gene functions are provided in Table 3.
| Operon | Accession | Gene name | Predicted function | Best BlastP hit | % Similarity | Score | E-value | Domains and motifs |
|--------|-----------|-----------|-------------------|----------------|-------------|-------|---------|-------------------|
| cbb1   | ACK78724.1| cbbR      | LysR family transcriptional regulatory protein CbbR | Nitrococcus mobilis | 76          | 363   | 7e-99   | PD462572, PD756396, Pfam03466, Pfam00126, COG0583 |
|        | ACK79627.1| cbbL1     | Ribulose bisphosphate carboxylase large subunit 1 [4.1.1.39] | Halothiobacillus neapolitanus | 94          | 882   | 0       | PD417314, PD000044, Pfam00016, Pfam02788, COG1850 |
|        | ACK77836.1| cbbS1     | Ribulose bisphosphate carboxylase small subunit 1 [4.1.1.39] | Methylcococcus capsulatus | 80          | 161   | 8e-39   | PD000290, Pfam00101, COG4451 |
|        | ACK78689.1| csoS2     | Carboxysome structural peptide | Thiobacillus denitrificans | 59          | 325   | 9e-87   | PD579361, tat signal peptide |
|        | ACK80925.1| csoS3     | Carboxysome structural peptide | Thiobacillus denitrificans | 65          | 537   | 5e-151  | PD191834, Pfam08936 |
|        | ACK80352.1| csoS4A    | Carboxysome peptide A | Thiobacillus denitrificans | 93          | 139   | 6e-32   | PD12510, Pfam03319, COG4576, tat signal peptide |
|        | ACK79436.1| csoS4B    | Carboxysome peptide B | Thiobacillus denitrificans | 82          | 119   | 7e-26   | PD12510, Pfam03319, COG4576 |
|        | ACK78722.1| csoS1C    | Microcompartments protein | Nitrosomonas europaea | 97          | 142   | 6e-33   | PD003442, Pfam00936, COG4577 |
|        | ACK78722.1| csoS1A    | Microcompartments protein | Nitrosomonas europaea | 97          | 144   | 1e-33   | PD003442, Pfam00936, COG4577 |
|        | ACK78722.1| csoS1B    | Microcompartments protein | Nitrosomonas europaea | 95          | 146   | 3e-34   | PD003442, Pfam00936, COG4577 |
|        | ACK79096.1| bfrA      | Bacterioferritin | Thiobacillus denitrificans | 70          | 135   | 6e-31   | PDA00179, Pfam00210, COG1633 |
|        | ACK79723.1| hyp1      | Hypothetical protein | Thiobacillus denitrificans | 81          | 68    | 2e-10   | PDA1E015 |
|        | ACK80576.1| parA      | Partition protein A | Thiobacillus denitrificans | 72          | 196   | 6e-49   | PD194671, Pfam01656, COG1192 |
|        | ACK78664.1| hyp2      | Hypothetical protein | Acidithiobacillus ferrooxidans | 100         | 156   | 1e-09   | |
|        | ACK80060.1| cbbQ1     | Rubisco activation protein | Nitrosomonas europaea | 92          | 489   | 5e-137  | PD490543, Pfam04046, Pfam07728, COG0714, COG5271 |
|        | ACK80817.1| cbbO1     | Rubisco activation protein | Thiobacillus denitrificans | 74          | 940   | 0       | PD140693, PD679436, Pfam00092, COG4867, COG4548 |
|        | ACK80290.1| cbbA      | Fructose-bisphosphate aldolase [4.1.2.13] | Bradyrhizobium sp. | 61          | 295   | 3e-78   | PD002376, PD030418, Pfam01116, Pfam07876, COG191 |
| cbb2   | ACK80366.1| cbbL2     | Ribulose bisphosphate carboxylase/oxygenase large subunit 2 [4.1.1.39] | Thiobacillus denitrificans | 97          | 920   | 0       | PD417314, PD000044, Pfam00016, Pfam02788, COG1850 |
|        | ACK79774.1| cbbS2     | Ribulose bisphosphate carboxylase/oxygenase small subunit 2 [4.1.1.39] | Thiobacillus denitrificans | 88          | 203   | 3e-51   | PD000290, Pfam00101, COG4541 |
|        | ACK80953.1| cbbQ2     | Rubisco activation protein | Nitrosomonas europaea | 92          | 483   | 6e-135  | PD490543, PD372819, Pfam08406, Pfam07728, COG1850 |
|        | ACK78928.1| cbbO2     | Rubisco activation protein | Thiobacillus denitrificans | 76          | 965   | 0       | PD140693, PD025507, COG4548 |
| cbb3   | ACK80740.1| hyp3      | Hypothetical protein | Thiobacillus denitrificans | 49          | 149   | 8e-9    | PD796582 |
|        | ACK78212.1| suhB      | Inositol-phosphate phosphatase [3.1.3.25] | Methylcococcus capsulatus | 66          | 646   | 8e-66   | PD001491, PD013702, pfam00059, pfam00016, COG0483, COG1218 |
|        | ACK80404.1| cbbF      | Fructose-1,6-bisphosphatase [3.1.3.11] | Mariprofundus ferrooxydans | 71          | 823   | 3e-86   | PD007014, PD63173, Pfam03320, COG1494 |
|        | ACK79091.1| cbbT      | Transketolase [2.2.1.1] | Methylcococcus capsulatus | 75          | 2264  | 0.0     | PD308336, pfam00456, pfam02779, COG4005, COG0021 |
|        | ACK78716.1| cbbG      | Glyceraldehyde-3-phosphate dehydrogenase type I [1.2.1.-] | Burkholderia thailandensis | 82          | 1189  | 1e-128  | PD959395, PD859695, pfam02800, pfam00044, COG0057 |
Table 3 Predicted genes of cbb operons (Continued)

| Operon cbb4 | CbbK         | Phosphoglycerate kinase [2.7.2.3] | Alcanivorax borkumensis | 80  | 1296 | 6e-141 | PD008619, PDA014E1, pfam00162, COG0126 |
|             | CbbA         | Fructose-bisphosphate aldolase [4.1.2.13] | Nitrosococcus oceanis | 79  | 1491 | 2e-163 | PD983049, PD745602, pfam00224, pfam02887, COG0469 |
|             | CbbE         | Ribulose-5-phosphate 3-epimerase [5.1.3.1] | Herminimonas anseunoxidans | 90  | 1474 | 1e-161 | PD875785, PD002378, pfam01116, COG0191 |
|             | CbbZ         | Phosphoglycerate phosphatase [3.1.3.18] | Thiodacillus denitrificans | 80  | 753  | 2e-78   | PD0035683, PD591639, pfam00834, COG0036 |
|             | CbbP         | Ribulose-5-phosphate 3-epimerase [5.1.3.1] | Herminimonas anseunoxidans | 64  | 484  | 4e-47   | PD946755, PDA11895, pfam00702, COG0546, COG0367 |
|             | CbbP         | Phosphoribulokinase [2.7.1.19] | Thiodacillus denitrificans | 77  | 1569 | 2e-172  | PD005777, PD105823, pfam00425, pfam04715, COG0147, COG1169 |
|             | CbbZ         | S-adenosylmethionine synthetase [2.5.1.6] | Ralstonia eutropha | 86  | 591  | 2e-167  | PD499406, PD606972, pfam02773, pfam02772, COG0192 |
|             | CbbP         | S-adenosyl-L-homocysteine hydrolase [3.3.1.1] | Pseudomonas stutzeri | 88  | 748  | 0        | PD730548, PD551162, pfam05221, pfam00670, COG0499 |
|             | CbbP         | 5,10-methylenetetrahydrofolate reductase [1.7.99.5] | Methylcoccus capsulatus | 69  | 306  | 1e-81   | PD756524, PD763008, pfam02219, COG0685 |
|             | CbbP         | Phosphoribulokinase [2.7.1.19] | Nitrosococcus oceanis | 78  | 402  | 2e-110  | PD739884, PD015803, pfam00485, COG3954 |
|             | CbbP         | Phosphoribulokinase [2.7.1.19] | Polaronomas naphthalenivorans | 81  | 759  | 1e-81   | PD898898, pfam01569, COG0474, CD03386, CD00127 |

* The sequence and annotation of the complete A. ferrooxidans strain ATCC 23270 genome is available at the Comprehensive Microbial Resource (CMR) (J. Craig Venter Institute, http://www.jcvi.org) and in GenBank/EMBL/DDBJ accession number CP001219.

a Proposed gene name.
b Proposed enzyme activity with EC number if available
c Organism with the best BlastP hit to the candidate gene.
d Percentage of similarity (% Sim) of candidate gene to that found in the organism listed in row (c).
e Score of BlastP match.
f E value of BlastP match.
g Motif and domains identified in the candidate proteins: CD, Conserved Domains; COG, Clusters of Orthologous Groups of Proteins; Pfam, protein families; PD, Prodom (protein domains); PS, Prosite tat signal peptide

Figure 3 Binding of CbbR to the promoter regions of the operons cbb1-4 using the EMSA assay in the presence (+) or absence (-) of competing 50x excess of unlabelled probe DNA (P[50x]) or antibodies to CbbR (anti-CbbR). Abbreviations: P*, probe DNA; S, shift; SS, supershift.
\(\gamma\)-proteobacteria divisions (Figure 4, Table 4), whereas only 4 out of 14 facultative heterotrophs were detected with this clustering. These four exceptions are found only in the \(\beta\)- or \(\gamma\)-proteobacteria and none in the \(\alpha\)-proteobacterial division (Figure 4, Table 4). This suggests a previously unreported linkage between genes encoding CBB cycle associated enzymes and \(trpEG\) or \(trpE\) that is most conserved in obligate autotrophs of the \(\beta\)- and \(\gamma\)-proteobacteria.

We hypothesize that in \(A. \text{ferrooxidans}\) production of pyruvate via anthranilate synthase activity provides a novel network connection between the CBB cycle on the one hand and general central carbon metabolism including the incomplete (‘horseshoe’-like) TCA [2] on the other hand. Consistent with this idea is the presence of a predicted \(pykA\) upstream of \(trpEG\) in the \(cbb3\) operon. \(PykA\) is predicted to encode pyruvate kinase that catalyzes the conversion of phosphoenolpyruvate (PEP) to pyruvate. In addition to supplying pyruvate, \(PykA\) could also reduce the level of intracellular PEP. PEP has been shown to be a ligand of CbbR in \(R. \text{marinus}\) H16, promoting its binding to target DNA sites and consequently effecting the regulation of \(cbb\) genes [40]. If PEP carries out a similar function in \(A. \text{ferrooxidans}\), the depletion of PEP via \(PykA\) activity could provide a means for feedback control of operons that are regulated by CbbR, including the auto-regulation of operon \(cbb3\).

The organization of \(cbb\) genes in \(A. \text{ferrooxidans}\) exhibits similarities with obligate autotrophs that distinguish this group from facultative autotrophs. For example, \(A. \text{ferrooxidans}\), contains three or more gene clusters dedicated to carbon assimilation. This is similar to other obligate autotrophic \(\gamma\)-proteobacteria including \(A. \text{caldus}\), \(A. \text{thiooxidans}\), \(H. \text{marinus}\), \(N. \text{oceanic}\) and \(T. \text{crunogena}\). In obligate autotrophic \(\beta\)-proteobacteria such as \(N. \text{eurgens}\), \(N. \text{marin}\) and \(T. \text{multiform}\), and \(T. \text{denitrifica}\) this contrasts with facultative autotrophs that contain only one or two \(cbb\) clusters (Figure 4, Table 4), with some exceptions, e.g. the \(\alpha\)-proteobacteria \(B. \text{sp}.,\) \(N. \text{winogradskii}\), \(R. \text{sphaeroides}\) and \(R. \text{pulis}\) and the \(\beta\)-proteobacteria \(R. \text{eutropha}\) which contain unique, but duplicated, \(cbb\) clusters. Multiple \(cbb\) clusters could provide obligate autotrophs with a greater flexibility in regulating CO\(_2\) fixation compared to facultative autotrophs. For example, this flexibility may be necessary to adjust carbon assimilation in response to changing environmental concentrations of CO\(_2\) [18], whereas facultative autotrophs might be able to circumvent this need by exploiting organic carbon sources in times of low CO\(_2\) concentrations.

Another characteristic of \(cbb\) gene organization in \(A. \text{ferrooxidans}\) is the lack of linkage of the phosphoribulokinase gene, \(cbbP\), with other \(cbb\) genes (Figure 4, Table 4) as has previously been reported for the deep-sea vent obligate chemolithoautotroph \(T. \text{crunogena}\) XCL-2 and for several other obligate autotrophs [20,41]; we now extend this list to include \(A. \text{ferrooxidans}\) ATCC 23270 and ATCC 53993, \(A. \text{caldus}\), \(A. \text{thiooxidans}\) \(H. \text{marinus}\), and \(T. \text{eurgens}\) (Figure 4, Table 4). In contrast, in all sequenced facultative autotrophs \(cbbP\) is associated with other \(cbb\) genes (Figure 4, Table 4).

In obligate autotrophs, the contextual disconnection of \(cbbP\) from \(cbbLS\) could provide greater flexibility for CO\(_2\) fixation by allowing RubisCO to be differentially expressed according to environmental and/or metabolic requirements without simultaneously expressing the remaining CBB cycle genes, many of which carry out functions in addition to carbon fixation. This is in sharp contrast to the organization found in most facultative autotrophs, where \(cbbP\) is usually juxtaposed to \(cbbLS\) and other genes of the CBB cycle facilitating their coordinate repression during heterotrophic growth [13,20,34,36,41].

**Model for predicted enzymes and pathways involved in CO\(_2\) fixation**

A model is proposed for \(C_i\) fixation in \(A. \text{ferrooxidans}\) based on the predicted roles of the genes encoded in the \(cbb\) operons (Figure 5). In contrast to most facultative autotrophs, the main focus of regulation of the CBB cycle in \(A. \text{ferrooxidans}\) may be the CO\(_2\) fixation reaction itself catalyzed by RubisCO, rather than at the level of the other CBB cycle enzymes. This hypothesis is supported by the observation that the genes encoding RubisCO and RubisCo accessory proteins, are clustered in operons that do not contain \(cbbP\) nor \(cbb\) that encode the main CBB enzymes. \(cbbP\) is also separated from the rest of the \(cbb\) genes in the \(cbb4\) operon, with an apparent absence of CbbR binding to its promoter. We suggest that the promoters for the \(cbb1\), \(cbb2\) and \(cbb3\) operons have different affinities for CbbR and may thus exhibit different regulation patterns, possibly associated with the environmental availability of CO\(_2\). The \(cbb1\) operon, containing \(cbbLS-cso\), is predicted to serve at low CO\(_2\) concentrations because carboxysomes have been shown to improve RubisCO catalytic efficiency by concentrating CO\(_2\) [6,13]. In contrast, the \(cbb2\) operon, containing \(cbbLS\)-cqo, is predicted to be used when higher concentrations of CO\(_2\) are available since carboxysome synthesis is energetically and materially expensive [18].

The \(cbb3\) operon, containing genes for most CBB cycle enzymes and pyruvate kinase, is proposed to be
Figure 4 Organization of gene clusters involved in the CBB cycle of facultative and obligate autotrophic α-, β- and γ-proteobacteria presented as a phylogenetic cladogram based on 16 S RNA. Numbers refer to bootstrapping results from 1000 trees. Organism names are provided in the text. The asterisk indicates that the respective organism is an obligate autotroph.
responsible for connecting CO₂ fixation with the rest of central carbon metabolism. Except for ubsG and ubsK encoding glyceraldehyde-3-phosphate dehydrogenase, type I and phosphoglycerate kinase respectively, genes of the ubs3 operon have duplicated copies in the genome (data not shown), potentially allowing regulation of the CBB cycle independently of the remaining pathways of central carbon metabolism. For example, some CBB cycle intermediates also form part of gluconeogenesis and glycolysis resulting in the production of pyruvate that is channeled, via the pyruvate dehydrogenase complex, to the incomplete TCA "horseshoe" where the flux of intermediates serves for amino acid biosynthesis (e.g. glutamate). The pyruvate dehydrogenase also provides acetyl-CoA used in fatty acid biosynthesis. In addition, the presence of ubsZ in the ubs3 operon is associated with phosphoglycolate phosphatase activity, responsible for removal of phosphoglycolate, an undesirable product of the oxygenase activity of RubisCO, that must be detoxified preferentially by rechanneling to 3-phosphoglycerate [13,36].

The co-transcriptional connection between the ubs, pypA and trpEG genes in the ubs3 operon may reflect the substrate requirement of anthranilate phosphoribosyltransferase for an activated pentose (5-phosphoribosyl 1-pyrophosphate) in order to proceed to the next step of tryptophan biosynthesis [42]. The production of the activated pentose would be stimulated by the activity of

### Table 4 Characteristics of ubs gene clusters in facultative and obligate autotrophic bacteria

| Organism | Autotrophy status | Phylogenetic classification -proteobacteria | No. copies cbbR genes? | Presence of cso genes? | trpEG associated with ubs? | cbb gene cluster associated with cbbP? | No. cbb gene clusters |
|----------|------------------|------------------------------------------|------------------------|------------------------|----------------------------|-------------------------------------|---------------------|
| Acidithiobacillus ferrooxidans ATCC 23270 and ATCC 33993 | obligate Gamma- | 2 | Yes | Yes | No | 5* |
| Acidithiobacillus thiooxidans ATCC 19377 | obligate Gamma- | 2 | Yes | Yes | No | 5 |
| Acidithiobacillus caldus ATCC 51756 | obligate Gamma- | 2 | Yes | Yes | No | 5 |
| Nitrosomonas europaea ATCC 19718 | obligate Beta- | 1 | No | Yes | No | 4 |
| Nitrosomonas europa ATCC 19707 | obligate Beta- | 1 | Yes | Yes | No | 4 |
| Nitrososoccus oceani ATCC 19707 | obligate Beta- | 1 | No | Yes | No | 4 |
| Thiomicrosphaera crunogena XCL-2 | obligate Gamma- | 3 | Yes | Yes | No | 5 |
| Hydrogenovibrio marinus MH-110 | obligate Gamma- | 2 | Yes | N/D | N/D | 3 |
| Thiobacillus denitrificans ATCC 25259 | obligate Beta- | 2 | Yes | Yes | No | 5 |
| Nitrosospira multiformis ATCC 25196 | obligate Beta- | 1 | No | Yes | No | 4 |
| Methyllobacter capsulatus Bath | obligate methanotroph | 1 | No | Yes | Yes | 3 |
| Nitrospira marina X14 | facultative Alpha- | 3 | Yes | No | Yes | 3 |
| Nitrospira winogradskyi Nb-255 | facultative Alpha- | 3 | Yes | No | Yes | 3 |
| Halothiobacillus halophilus SL1 | facultative Gamma- | 1 | No | Yes¹ | Yes | 2 |
| Alkalilimnica ehrlichii MLHE-1 | facultative Gamma- | 1 | No | Yes² | Yes | 2 |
| Bradyrhizobium sp. BTAI | facultative Alpha- | 2 | Yes | No | Yes | 3 |
| Bradyrhizobium japonicum USDA 110 | facultative Alpha- | 1 | No | No | Yes | 1 |
| Ralstonia eutropha H16 | facultative Beta- | 1 | No | No | Yes | 1 |
| Dechloromonas aromatica RCB | facultative Alpha- | 1 | No | No | Yes | 2 |
| Magnetospirillum magneticum AMB-1 | facultative Alpha- | ? | No | No | Yes | 2 |
| Paracoccus denitrificans PD1222 | facultative Alpha- | 1 | No | No | Yes | 1 |
| Rhodoferax ferrireducens T118 | facultative Beta- | 1 | No | No | Yes | 1 |
| Rhodopseudomonas palustris CGA009 | facultative Alpha- | 2 | No | No | Yes | 3 |
| Rhodospirillum rubrum ATCC 11170 | facultative Alpha- | 1 | No | No | Yes | 1 |
| Sinorhizobium meliloti 1021 | facultative Alpha- | 1 | No | No | Yes | 1 |

*In addition to the four ubs operons described in this paper, a fifth gene cluster containing ubs genes (including a form II RubisCO gene) has recently been detected in A. ferrooxidans (43). Two copies of cbbR and two cbb gene clusters are present on two plasmids; two highly similar operons present in the genome; in these organisms, trpE gene is neighbor to cbbP but not cbbE. M. eutropha H16 posseses a duplicated cbb operon, with similar copies in the chromosome and in a megaplasmid. Data derived from cloned sequences (18). N/D = no data.
Figure 5 Proposed roles of the (A) predicted enzymes and pathways involved in CO₂ fixation in *A. ferrooxidans* linked to (B) gene evidence. Genes are color-coded to match the predicted function of their products. RPI, ribose phosphate isomerase; G-3-P, glyceraldehyde-3-phosphate; DHAP, dihydroxyacetone phosphate; 3-PG, 3-phosphoglycerate; PEP, phosphoenolpyruvate.
the operon. An alternate hypothesis is that the co-transcriptional connection represents a means for pyruvate regeneration since both pykA and trpE/G produce pyruvate.

In addition to the four cbb operons described herein, a fifth gene cluster has recently been detected in *A. ferrooxidans* that includes genes *cbbM, cbbO3* and *cbbO3* predicted to encode form II of RuBisCO and its associated chaperons, respectively [43]. The cluster also contains another putative *cbbR* divergently transcribed from *cbbMQO*. Future work will evaluate the role of this cluster in CO₂ fixation.

**Additional material**

**Additional file 1: Prediction of secondary structure elements in CbbB of Acidithiobacillus ferrooxidans.** Above: secondary structure predictions of alpha-helix, beta-sheet, HTH DNA binding domain, oligomerization domain and LysR-substrate like domain. Below: alignment of amino acid sequences from the HTH domain from several bacteria (abbreviations used can be found in Additional File 2) with the pfam domain01026.

**Additional file 2: Alignment and conservation of DNA sequences in the intergenic regions between cbbR and cbbL1 in autotrophic bacteria.** The DNA sequences contain the cbb control elements including the operator, the promoter (pccbL) and the promoter cbbR (pccbR). The CbbR regulator bind to region R (recognition site) and the region A (activation site) of the cbbR operator. The nucleotides conserved (TNA-N20-TNA, T-N11-A) for to bind CbbR are located in intergenic regions RI-1, RI-2 and RI-3. The prediction of the promoter and the sites for to bind σ70 are in the columns (sequences -35 and -10).

**Additional file 3: Sequences used to generate LOGOS of the intergenic region between cbbR and cbbL1.**

**Acknowledgements**

This work was supported by a grant from Fondecyt 1090451, a Microsoft Sponsored Research Award, a Deutscher Akademischer Austausch Dienst (DAAD) scholarship to ME, a CONICYT graduate student grant to J-PC and a grant from the Deutsche Forschungsgemeinschaft to BB.

**Author details**

1. Center for Bioinformatics and Genome Biology, MIFAB, Fundación Ciencia para la Vida and Deppto. de Ciencias Biológicas, Facultad de Ciencias Biológicas, Universidad Andres Bello, Santiago, Chile.
2. ICBM, Faculty of Medicine, University of Chile, Santiago, Chile.
3. Instituto de Microbiología und Genetik, Georg-August-Universität Göttingen, Göttingen, Germany.
4. Departamento de Acuicultura, Facultad de Recursos del Mar, Universidad de Antofagasta, Antofagasta, Chile.

**Authors’ contributions**

DH, EJ and ME conceived the study. ME carried out the experiments. BB and J-PC contributed significantly to the analysis and interpretation of results. DH drafted the manuscript. All authors contributed to the draft and approved the manuscript.

**Received:** 4 March 2010 Accepted: 27 August 2010

**Published: 28 August 2010**

**References**

1. Holmes D, Bonnefoy V: Genetic and bioinformatic insights into iron and sulfur oxidation mechanisms of bioleaching organisms. In Biomining. Edited by: Rawlings DE, Johnson B. D. Springer Berlin Heidelberg, 2007:281-307.
2. Valdés J, Pedrosa I, Quatrini R, Dodson RJ, Terrettta H, Blake R, Eisan JA, Holmes D: Acidithiobacillus ferrooxidans metabolism: from genome sequence to industrial applications. BMC Genomics 2008, 9:597.
3. Ask A Scientist. Carbon dioxide and water. [http://www.newton.dep.anl.gov/askasci/chem03/chem03573.htm]
4. Kusano T, Sugawara K: Specific binding of Thiobacillus ferrooxidans RbcR to the intergenic sequence between the rbc operon and the rbcL gene. J Bacteriol 1993, 175:1019-1025.
5. Tabita FR: Molecular and cellular regulation of autotrophic carbon dioxide fixation in microorganisms. Microbiol Rev 1988, 52:155-189.
6. Price GD, Badger MR, Woodger FJ, Long BM: Advances in understanding the cyanobacterial CO₂-concentrating-mechanism (CCM): functional components, O transporters, diversity, genetic regulation and prospects for engineering into plants. J Exp Bot 2008, 59:1441-1461.
7. Cannon GC, Baker SH, Soyer F, Johnson DR, Bradburne CE, Mehmehn JL, Davies PS, Jiang QL, Heinhorst S, Shively JM: Organization of carboxysome genes in the thiobacilli. Curr Microbiol 2003, 46:115-119.
8. Appia-Ayme C, QR, Denis Y, Denizot F, Silver S, Roberto F, Veloso F, Valdes J, Cárdenas JP, Esparza M, Orellana O, Jedlicki E, Bonnefoy V, Holmes D: Microarray and bioinformatic analysis suggest models for carbon metabolism in the autotroph Acidithiobacillus ferrooxidans. Hydrobiologia 2006, 582:273-280.
9. van den Bergh ER, Dijkhuizen L, Meijer WG: CbbR, a LysR-type transcriptional activator, is required for expression of the autotrophic CO₂ fixation enzymes of Xanthobacter flavus. J Bacteriol 1993, 175:6097-6104.
10. Windhovel U, Bowien B: Identification of cbbR, an activator gene of autotrophic CO₂ fixation in Alcaligenes eutrophus. Mol Microbiol 1991, 5:2695-2705.
11. Sen P, Roy D: Multiple sequences for the regulation of gene expression in Chlorobium tepidum. Curr Microbiol 2003, 46:155-189.
12. Strecker M, Sickinger E, English RS, Shively JM: Calvin cycle genes in Nitrobacter vulneris T3. FEMS Microbiology Letters 1994, 120:45-50.
13. Shively JM, van Keulen G, Meijer WG: Something from almost nothing: carbon dioxide fixation in chemoautotrophs. Annu Rev Microbiol 1998, 52:191-230.
14. Beller HR, Chain PS, Letain TE, Chakicherla A, Larimer FW, Richardson PM, Coleman MA, Wood AP, Kelly DP: The genome sequence of the obligately chemolithoautotrophic, facultatively anaerobic bacterium Thiobacillus denitrificans. J Bacteriol 2006, 188:1473-1488.
15. Gibson J, Tabita FR: Nucleotide sequence and functional analysis of cbbR, a positive regulator of the Calvin cycle operons of Rhodobacter sphaeroides. J Bacteriol 1993, 175:5778-5784.
16. Paoli GC, Soyer F, Shively J, Tabita FR: Rhodobacter capsulatus genes encoding form I ribulose-1,5-biphosphate carboxylase/oxygenase (cbbLS) and neighbouring genes were acquired by a horizontal gene transfer. Microbiology 1998, 144(Pt 11):219-227.
17. Falcón DL, Tabita FR: Complementation analysis and regulation of CO₂ fixation gene expression in a ribulose-1,5-biphosphate carboxylase/oxygenase deletion strain of Rhodospirillum rubrum. J Bacteriol 1993, 175:5066-5077.
18. Toyoda K, Hoshizawa Y, Arai H, Ishi M, Igarashi Y: The role of two CbbRs in the transcriptional regulation of three ribulose-1,5-biphosphate carboxylase/oxygenase genes in Hydrogenobacter marinus strain MH-110. Microbiology 2005, 151:3615-3625.
19. Wei X, Sayavedra-Soto LA, Arp DJ: The transcription of the cbb operon in Nitrosomonas europaea. Microbiology 2004, 150:1869-1879.
