BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 8(15)

Cloning, expression and bioinformatics analysis of ATP sulfurylase from *Acidithiobacillus ferrooxidans* ATCC 23270 in *Escherichia coli*

Michael L Jaramillo[1], Michel Abanto[1], Ruth L Quispe[1], Julio Calderón[1], Luís J del Valle[2], Miguel Talledo[1] & Pablo Ramírez[1]*

[1]Laboratory of Molecular Microbiology and Biotechnology, Faculty of Biological Sciences, Universidad Nacional Mayor de San Marcos, Lima – Peru; [2]Centre d'Enginyeria Biotecnologica i Molecular (CEBIM), Departament d'Enginyeria Química, ETSEIB, Universitat Politècnica de Catalunya, Barcelona, Spain; Pablo Ramirez – Email: pabloramirezroca@gmail.com; Phone: +51 15238492; *Corresponding author

Received July 10, 2012; Accepted July 11, 2012; Published August 03, 2012

Abstract:
Molecular studies of enzymes involved in sulfite oxidation in *Acidithiobacillus ferrooxidans* have not yet been developed, especially in the ATP sulfurylase (ATPS) of these acidophilus tiobacilli that have importance in biomining. This enzyme synthesizes ATP and sulfate from adenosine phosphosulfate (APS) and pyrophosphate (PPi), final stage of the sulfite oxidation by these organisms in order to obtain energy. The *atpS* gene (1674 bp) encoding the ATPS from *Acidithiobacillus ferrooxidans* ATCC 23270 was amplified using PCR, cloned in the pET101-TOPO plasmid, sequenced and expressed in *Escherichia coli* obtaining a 63.5 kDa ATPS recombinant protein according to SDS-PAGE analysis. The bioinformatics and phylogenetic analyses determined that the ATPS from *A. ferrooxidans* presents ATP sulfurylase (ATS) and APS kinase (ASK) domains similar to ATPS of *Aquifex aeolicus*, probably of a more ancestral origin. Enzyme activity towards ATP formation was determined by quantification of ATP formed from *E. coli* cell extracts, using a bioluminescence assay based on light emission by the luciferase enzyme. Our results demonstrate that the recombinant ATP sulfurylase from *A. ferrooxidans* presents an enzymatic activity for the formation of ATP and sulfate, and possibly is a bifunctional enzyme due to its high homology to the ASK domain from *A. aeolicus* and true kinases.

Background:
ATP sulfurylase enzyme (ATPS, sulfate adenylyltransferase, EC 3.7.7.4) produces APS and AMP from sulfate and ATP or viceversa. Its physiological function is dependent on the metabolism of the organism **[1]**. This enzyme is widely distributed in all living organisms: Archaea, Bacteria, protista, yeasts, filamentous fungi, plants, animals and humans **[2-5]**. Variations in size of this enzyme are due to that ATP sulfurylase domain (ATS) can be bound to the APS kinase domain or similar to APS kinase (ASK). Exceptionally, ATP sulfurylase has been found to bind to other enzymes such as pyrophosphatase and APS reductase **[5]**. ATPS is involved in the assimilatory and dissimilatory reduction of sulfate catalyzing the activation of inorganic sulfate by ATP to form adenosine-5'-phosphosulfate (APS) and pyrophosphate **[3]**. The assimilation pathway is performed by aerobic organisms for amino acid synthesis such as cysteine, and the dissimilation pathway is carried out by prokaryotes, which in the absence of molecular oxygen they use sulfate as a terminal electron acceptor for cellular respiration, releasing hydrogen sulfide (H_2S) **[6]**. Moreover, ATPS is involved in the indirect oxidation of sulfite, which could be present in a variety of chemolithoautotrophic and phototrophic bacteria, allowing for energy conservation by a chemiosmotic mechanism or by a substrate-level mechanism **[7]** where sulphite would be oxidized via a reverse adenosine phosphosulfate reductase (APR) activity according to: $H_2SO_3 + AMP \leftrightarrow APS + 2e^- + 2H^+$, with participation of ATPS (sulfate adenylyltransferase) in the final step according to: $APS + PP_i \leftrightarrow ATP + SO_4^{2-}$.

Chemolithoautotrophic bacteria such as *A. ferrooxidans* would be responsible for the oxidation of sulphite in mining environments, where there is a large quantity of reduced inorganic sulfur compounds (RISC). In previous research the genomic sequence of *A. ferrooxidans* has been analyzed by bioinformatics software **[8]** and there are potential gene sequences that would encode the ATP sulfurylase enzyme. It is postulated that ATP sulfurylase is a monomer or homooligomer in organisms that oxidize reduced sulfur compounds and presents a subunit molecular weight (MW) ranging from 41 to 69 kDa, in a similar fashion to the ATPS sequence involved in sulfate assimilation found in *Penicillium chrysogenum* and *Saccharomyces cerevisiae* (58 kDa and 63.7 kDa subunits, respectively). ATPS from hyperthermophile chemolithotrophic *Aquifex aeolicus* is an ortholog of the filamentous fungi enzyme, with a functional ASK domain (APS kinase domain) **[9, 10]** and presents high similarity to the true APS kinases in the "P-loop" region.

By contrast, ATPS from chemolithotrophic *Riftia pachyptila* symbiont bacteria presents only the ATS domain. It is postulated that ATPS present in chemolithotrophic organisms is more similar to ancestral ATPS that gave rise to all homooligomeric ATP sulfurylases (ATPSs) **[1,10]**. Other homooligomeric ATPSs are present in higher eukaryotic organisms with domains in inverted orientation (ASK domain at the N-terminus and ATS domain at the C-terminus) different in relation to the ATPS from *A. aeolicus* and *fungi* **[10]**. Their presence has been reported in several animals, such as the marine invertebrate *Urechis carpo* **[11]**, fruitfly **[12]**, mouse **[13]** and humans **[14]**, and express a single polypeptide, named PAPS synthetase **[15]** that participates in a sulfation process. Other ATPSs only have the ATS domain, and are found in the bacteria *Thermus thermophilus* HB, in a similar way to the *Riftia* symbiont **[16]**, *Bacillus subtilis*, *Chromatium vinosum* and *Geobacillus*, cyanobacterium *Synechocystis sp*., archaeon *Archaeoglobus fulgidus* **[3]** and plants (*Arabidopsis thaliana*). ATPSs of *Desulfuvibrio* species are metalloproteins that bind cobalt and zinc, and present a C-X2-C-X8-CXH characteristic sequence (metal binding site) **[7]**. Homooligomeric ATPSs are not similar to heterooligomeric ATPSs found in *E. coli* and responsible for sulfate assimilation **[3, 7]**.

Homooligomeric ATPSs have V blocks in the ATS domain, II and IV blocks are rich in basic amino acids suggesting that they participate in the binding of $MgATP^{2-}$and SO_4^{2-} (Sperling *et al.* 1998). In the ATPS from *Riftia* bacterial symbiont has been determined that the phosphosulfate motif ^{199}QXRN^{202} and the motif ^{205}HXXH^{208} **[1, 17]** are the active site and the substrate-binding site, respectively. It presents a mobile loop of the active site that includes an extended sequence: ^{224}(hp)$_{3}$HXhpXGXXKXXDhpXXXXR^{243} (hp = hydrophobic residues) **[1]**, where the Asp234 in the ATPS from *P. chrysogenum* (Asp237 in the *Riftia* bacterial symbiont and Asp207 in *Aquifex*) orients Arg199 (Arg172 in *Aquifex*) in the sulfate-binding pocket which is necessary for sulfurylase activity **[10]**, and these motifs are conserved in all homooligomeric ATPSs. In the present study, the ATPS enzyme that would participate in the oxidative metabolism of sulfite in the bioleaching bacterium *A. ferrooxidans* was characterized at the molecular level. The *atpS* gene was cloned and expressed in *E. coli* showing a high activity towards the synthesis of ATP, similar to other studied chemolithotrophic bacteria, demonstrating a high similarity to the bifunctional ATPS from *A. aeolicus*. Later studies about this enzyme could be used in the acid water control because one of the products of the enzymatic reaction is the sulphate ion that reacts with hydrogen and generates sulfuric acid, an important environmental polluting agent.

Methodology:

Strains, plasmids and culture conditions

A. ferrooxidans ATCC 23270 was cultured in a modified 9K medium at pH 1.5, in a rotary shaker at 28ºC for 7 days **[18].** *E. coli* TOP10 and BL21Star™ (DE3) (Invitrogen) were cultured on Luria Bertani (LB). For growth of transformant *E. coli* strains in LB medium agar was added to a final concentration of 1.8% (w/v) and ampicillin (100 µg/ ml). PET101/D-TOPO plasmid was used for gene cloning obtaining a pET101-*atpS* plasmid.

DNA manipulation

The extraction of genomic DNA from *A. ferrooxidans* was carrying out by using a Miniprep Wizard Genomic Purification Kit (Promega) following the supplier´s instructions. Previously, the cells were concentrated by centrifugation at 13000 x *g* for 2 min and then washed 3 times with acidic water (pH 1.72) and 10 mM sodium citrate (pH 8.0). Genomic DNA was stored at -20ºC. The *atpS* gene encoding ATP sulfurylase was directly searched analyzing the genome sequence from *A. ferrooxidans* ATCC 23270 using the tBlastn algorithm v2.0 **[19]**. The gene was amplified with the: atpS-N-term primer: (5'-CAC CAT GCC ATC CAT TCC TCA GGT TG-3') with the addition of nucleotides CACC at the 5' end to allow ligation to the vector and the atpS-C-term primer (5'-CGC CTC CTC CGC GAT GCG TTT CAC-3'). Cloning of the *atpS* gene was carried out using a Champion™ pET directional TOPO expression kit (Invitrogen) according to the supplier´s instructions. The *atpS* gene (1674 bp) from *A. ferrooxidans* was amplified by PCR using the Thermal Ace Polymerase (Invitrogen) and the following PCR program: 3 minutes at 95ºC followed by 30 cycles at 95ºC for 30 s, 60ºC for 30 s and 72ºC for 2.5 minutes and finally to 72ºC for 15 minutes. For cloning, the amplified fragment was linked to the pET101/D-TOPO plasmid (5753 bp) resulting in the pET101-*atpS* plasmid. Then, chemically competent cells (*E. coli* TOP10) were transformed with the pET101-*atpS* plasmid. The recombinant clones were selected growing in LB agar + ampicillin (100 µg/ml) and incubating for 18 h at 37°C. The presence of the insert in *E. coli* TOP10 transformants were verified by colony PCR using the N-terminal primer of the vector (T7-TopoF-vector) and the atpS-C-term primer of *atpS* gene. The pET101-*atpS* plasmid of *E. coli* TOP10 positive recombinant was extracted using the plasmid DNA Minipreparation method with cetyl trimethyl ammonium bromide (CTAB).

Briefly, 4.5 ml of an overnight culture grown in Luria broth plus ampicillin was centrifuged, and the resulting pellet was resuspended in 200 µl of lysozyme (50 mg/ml) and 4 µl of STET and incubated for 5 min at room temperature. After that, the suspension was placed in water at 100ºC for 45 s, centrifuged at 13000 x *g* for 10 min and 8µl of RNAse (10 mg/ml) was added, following incubation for 15 min at room temperature. 8 µl of 5% CTAB was added and incubated for 3 min at room temperature. The pellet obtained by centrifugation at 13000 x *g* for 10 min was resuspended in en 300 µl of 1.2 M NaCl. The plasmid DNA

was then precipitated with 750 µl of absolute ethanol, incubated for 5 min at room temperature and centrifuged at 13000 x *g* for 10 min. The obtained pellet was resuspended in 70% ethanol and centrifuged at 13000 x *g* for 5 min. Supernatant was discarded and the pellet was dried for 30 min. Then, it was resuspended in 10 µl of ultrapure water.

Genomic DNA, plasmid and PCR products were visualized by 1% agarose gel electrophoresis (Gibco BRL®) in TAE 0,5X. The molecular weight marker used was the 1kb Plus DNA Ladder (Invitrogen). Ethidium bromide (0.5 µg/ml) was used for a 5 min gel staining. Gels were visualized with a UV transilluminator at 320 nm. The *atpS* gene obtained from plasmid pET101-*atpS* was sequenced to verify the correct insertion of the gene in the plasmid and to obtain the nucleotide sequence of the gene. Sequencing was done using the following primers: T7-TopoF-vector (5'-TAA TAC GAC TCA CTA TAG GG-3'), T7-TopoR-vector (5'-TAG TTA TTG CTC AGC GGT GG-3 ') and internal primers atpS-R673 (5'-AGT CTT CGT ATA ATG GGG-3'), atpS-F1005 (5'-ATT ACG GGT TCC TTG GAG-3 ') and atpS-R1075 (5 '-GTA GCG ACC GGC CTG ACG-3'). Sequences were viewed and assembled using the SeqAssem software **[20]**. Accession number for the *atpS* gene sequence of *A. ferrooxidans* ATCC 23270: FM177944.

Expression of recombinant ATPS protein of A. ferrooxidans in E. coli BL21 Star™ (DE3)

E. coli BL21 Star™ (DE3) transformation with pET101-*atpS* plasmid was made according to manufacturer´s instructions. The induction-expression analysis of recombinant *E. coli* was conducted in the presence of 1 mM IPTG at 37ºC for 3 hours with constant agitation (DO_{600nm} of 0.6). *E. coli* BL21 Star™ (DE3) without plasmid was used as a control strain. Recombinant protein expression was analyzed from total cell extracts by SDS-PAGE **[18]**. For determination of MW was used marker Page Ruler™ Protein Ladder (Fermentas).

ATPS enzymatic activity of A. ferrooxidans by bioluminescence

The protein concentrations of *E. coli* BL21 strain Star™ (DE3) cell extracts was analyzed under induction and no induction conditions with IPTG and were determined using the Bradford reagent **[21]** to an absorbance at 595 nm. A bioluminescent test ATP Determination Kit (Molecular Probes) was used to determine the ATP sulfurylase activity according to manufacturer's instructions. It produced the standard reaction solution (10 ml) containing: H_2O, 8.9 ml reaction buffer (20X), 0.5 ml; DTT (0.1 M), 0.1 ml; D-luciferin (10 mM), 0.5 ml; and luciferase (5 mg/ml), 2.5 µl. A 100 µl reaction was started with the addition of luciferase. Testing with extracts de *E. coli* to the reaction was made by addition of 0.125 µl of APS (8 nmol/µl) and 0.5 µl of PPi (40 pmol/µl) for ATP production (according to the reaction: PPi + APS → ATP + SO_4^{2-}). Subsequently, ATP produced was detected by bioluminescence (according to the reaction: Luciferin + ATP + O_2 → oxyluciferin + AMP + pyrophosphate + CO_2 + light). The following controls were used: a) Control without luciferase, APS and PPi, b) Control with APS and PPi without luciferase and c) Control of the positive reaction with 5 µl of ATP (5 µM). Relative light units (RLU) readings were carried out in a microplate reader luminometer Synergy ™ HT Multi-Detection Microplate Reader (Bio-Tek, USA). Readings were controlled by KC4 v3.0 software with PowerReports (Bio-Tek). All RLU measurements were performed in duplicate at a wavelength of 530 nm for 15 min of reaction at 25°C. The quantity of ATP produced was calculated from the URL of standard ATP concentrations (0.25 to 10 µM) in duplicate at a wavelength of 530 nm at 25ºC. An enzymatic unit produces 1.0 µmol of ATP from APS and PPi per minute at 30ºC.

Bioinformatics and phylogenetic analysis of the amino acid sequence of the ATPS of A. ferrooxidans

From the amino acid sequence of the ATPS of *A. ferrooxidans* obtained in this work (CAQ76453) was determined the molecular weight and *pI* using the ProtParam program **[22]**. PSLpred software **[23]** and PSORTb v.2.0 **[24]** were used to determine the subcellular localization. The presence of signal peptide was determined using SignalP v1.1 software **[25].** For comparison of *A. ferrooxidans* ATPS, amino acid sequences of other ATPS were obtained from Genbank **[26].** Sequences were aligned with Clustal W **[27]** and edited with Bioedit **[28]** or BOXSHADE **[29].** For structural alignment of *A. ferrooxidans* ATPS sequence, the Cn3D software was used **[30].** Previously, the structure of ATPSs from *A. aeolicus* (MMDB ID 44452, PDB: 2GKS_A) and APS kinase from *P. chrysogenum* (MMDB ID: 21216, PDB: IM7G_A) were obtained from NCBI. Phylogenetic analysis of the ATPS from *A. ferrooxidans,* was made using the amino acid sequences of homooligomeric ATPSs obtained from NCBI database **[31]** (http://www.ncbi.nlm.nih.gov), including the sequence for *A. ferrooxidans* ATPS (CAQ76453). Sequence annealing was made using the Clustal X program **[32]** and edited with the BioEdit program **[28]**. For the construction of phylogenetic trees the MEGA v.5.05 program was used **[33]**. Gaps were excluded and the Neighbor-joining method **[34]** with Poisson correction model was used. Consistency of tree was evaluated by bootstrap analysis of 1000 repetitions.

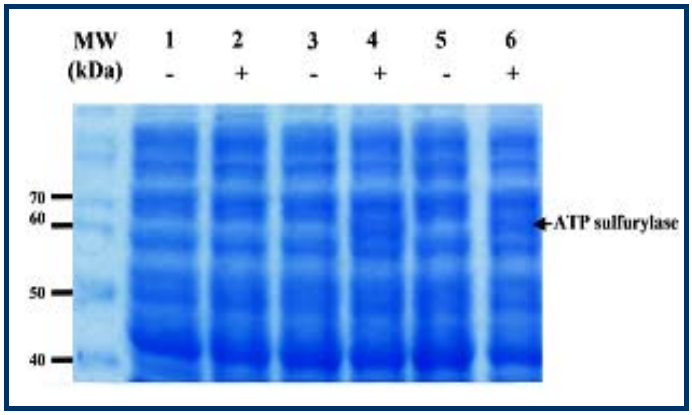

Figure 1: *In vivo* overexpression of the ATP sulfurylase from *A. ferrooxidans* 23270 in *E. coli*. The molecular weight standard is indicated on the left in kilodaltons. 1 and 2, total protein extracts of *E. coli* BL21 Start (DE3) with and without IPTG, **4** and **6**, induction of protein expression of the ATP sulfurylase in recombinant clones BL21-22 and BL21-42 containing the *atpS* gene with histidine tail. The recombinant plasmids were used to transform the *E. coli* BL21 Star (DE3) strain. In the logarithmic growth phase, strains were grown in the presence (2, 4 & 6) or absence (1, 3 and 5) of 1 mM IPTG for 3 h. Samples of total protein (1 to 6) were separated by SDS-polyacrylamide electrophoresis and stained with Coomassie blue. The right arrow indicates the band corresponding to the induction of ATPS protein bind to 6xHis at the C-terminal domain.

BIOINFORMATION

open access

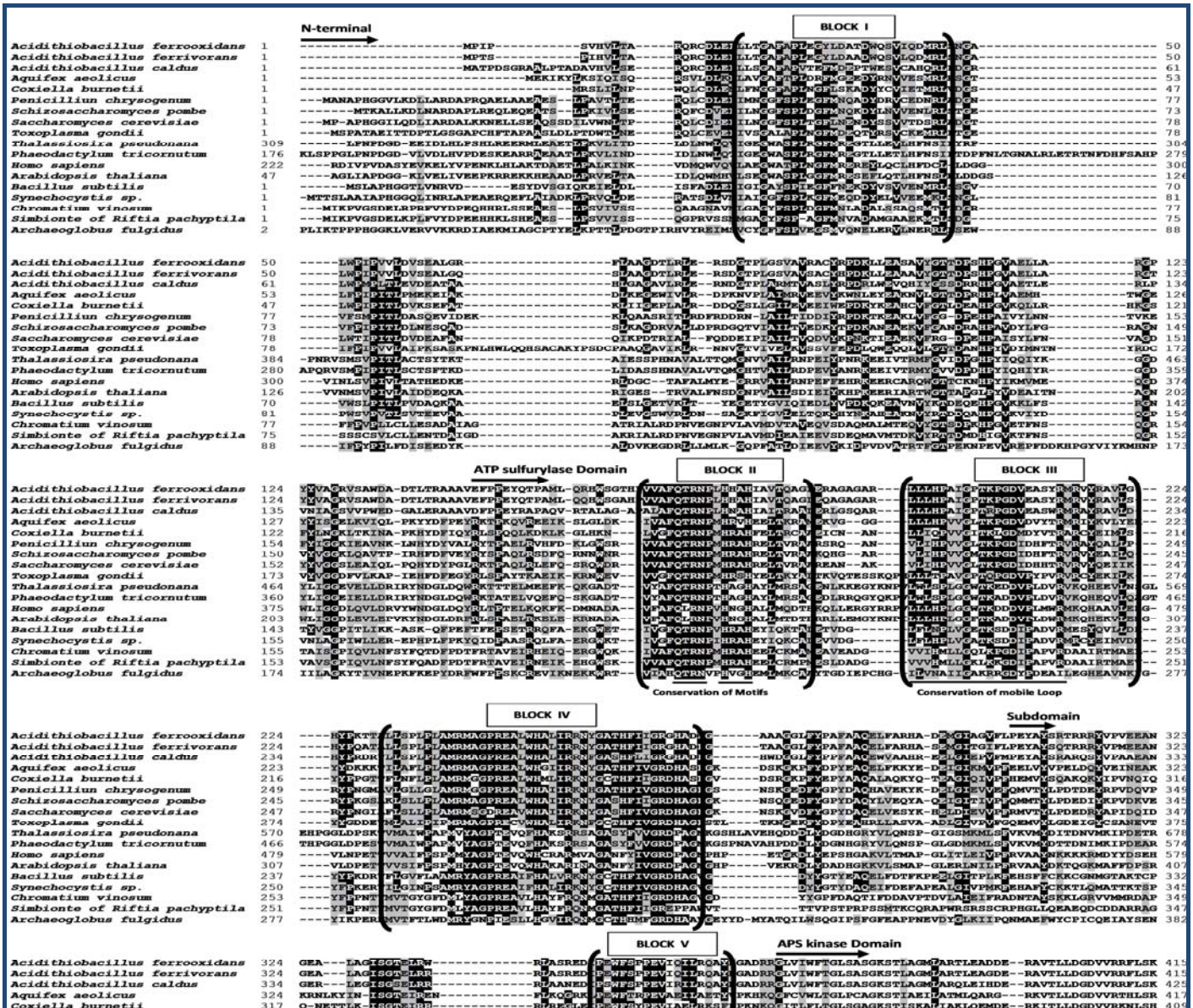

Figure 2: Multiple alignment of the ATPS from *A. ferrooxidans* with other homooligomerics ATP sulfurylase presenting 5 conserved regions or blocks in the ATS domain. Amino acids were aligned using Clustal W. Residues identical in most sequences are indicated in black and similar residues are indicated in gray. The arrows indicate the site of initiation of the N-terminal, domain ATP sulfurylase, Domain APS kinase and subdomain of the ATPS from *A. ferrooxidans*. *A. ferrooxidans (CAQ76453), A. ferrivorans (YP004784724), A. caldus (EET28427), A. aeolicus (O67174), C. burnetti (AAO90244), P. chrysogenum (Q12650), Sch. pombe (P78937), S. cerevisiae (P08536), T. gondii (EEB04289), T. pseudonana (EED89391), P. tricomutum (EEC48794), H. sapiens (O43252), A. thaliana (NP188929), B. subtilis (O34764), Synechocystis sp. (P74241), Chromatium vinosum (O66036), Simbionte of Riftia pachyptila (AAA72419), A. fulgidus (O28606).*

Result:

Identification of ATPS gene in the A. ferrooxidans ATCC 23270 genome

An open reading frame (ORF) for the *atpS* gene (1674 bp) encoding ATP sulfurylase (ATPS) was located at nucleotide positions 2327613 to 2329286 in the genome of *A. ferrooxidans* by bioinformatic analysis. Comparing this sequence with other ATPS of GenBank indicates the presence of an ATP sulfurylase domain (ATS) conserved typical of this type of enzyme, which is associated with an APS kinase domain (ASK) as occurs in some ATPSs **[15]**. The expected size of amplified *atpS* gene was obtained (1674 bp) and cloned in the pET101-TOPO vector

resulting in the plasmid pET101-*atpS*, which was used to transform *E. coli* cells. Secuencing of *atpS* gene was performed from the plasmid extracted from recombinant *E. coli* TOP10.

Expression of recombinant A. ferrooxidans ATCC 23270 ATPS protein in E. coli

The cells of recombinant *E. coli* BL21 Star™ (DE3) on induction conditions with IPTG expressed a recombinant ATPS protein of approximately 63.5 kDa **(Figure 1).** *E. coli* BL21 Star™ (DE3) without plasmid was subjected to the same experiment and did not express the protein in any condition. The weight of the monomer of *A. ferrooxidans* ATPS is 60.49 kDa based on their amino acid sequence.

Enzymatic activity assay of the ATPS protein of A. ferrooxidans

Bioluminescence carried out at a wavelength of 530 nm at 25ºC for 15 min of reaction showed enzyme activity in cell extracts of recombinant *E. coli* strain BL21 (+IPTG) expressing ATPS, which is higher compared with those obtained from extracts of recombinant *E. coli* BL21 (-IPTG), and from negative controls. Specific activity of recombinant ATPS from *A. ferrooxidans* was calculated from cell extracts of *E. coli*, obtaining 106.9 µM ATP/mg/min.

Bioinformatics analysis of the amino acid sequence of ATPS from A. ferrooxidans

The biochemical characteristics of the ATP sulfurylase from *A. ferrooxidans* was determined from its amino acid sequence (CAQ76453) using bioinformatics programs. According to these results the ATPS of *A. ferrooxidan*s would be a soluble protein of cytoplasmic localization similar to other homooligomeric ATPSs that have been described. It has a molecular weight de 60.49 kDa and has not signal peptide. The comparison of the amino acid sequence of ATP sulfurylase from *A. ferrooxidans* (CAQ76453) with other ATP sulfurylase, shows 93% identity and 95% similarity with ATP sulfurylase from *A. ferrivorans* (YP_004784724), 70% identity and 80% similarity with ATP sulfurylase from *A. caldus* ATCC 51756 (EET28427) and 44% identity and 60% similarity with the ATP sulfurylase from *A. aeolicus* (O67164), but also has similarity to ATPS found in *Coxiella burnetii* RSA493, fungi and yeasts. The high similarity of ATPS from *A. ferrooxidans* with homooligomeric ATP sulfurylases is demonstrated by multiple alignments using the Clustal W program **(Figure 2).** From the alignment there was deduced residues comprising the N-terminal domain (1-144), the central catalytic ATS domain (145-304), C-terminal APS kinase domain (residues 369-557) and a subdomain (305-368) which is considered part of the ATS domain by Yu *et al.* (2007) **[10].** The ATS domain of ATPS from *A. ferrooxidans* presents 5 blocks highly conserved in other homooligomeric ATPSs **[3] (Figure 2).** The active site of ATPS from *A. ferrooxidans* has conserved amino acids presenting the phosphosulfate-binding motif ^{170}QXRN^{173} and the ^{176}HXXH^{179} motif found in Block II and is similar to the *R. pachyptila* symbiont **[1],** *A. aeolicus* **[10]** and *P. chrysogenum* **[4].** Furthermore, the mobile loop at the active site of ATPS from *A. ferrooxidans* found in block III presents the consensus sequence: ^{196}(hp)$_3$HXhpXGXXKXXDhpXXXXR^{215}, which is conserved among various families of homoligomeric ATP sulfurylases (hp = hydrophobic residues). The ASK domain of the ATPS from *A. ferrooxidans* has a high similarity with the ASK domain of the bifunctional ATPS from *A. aeolicus* and is identical to the active site of the true kinases (^{376}GLSASGKST^{384}) **(Figure 3).**

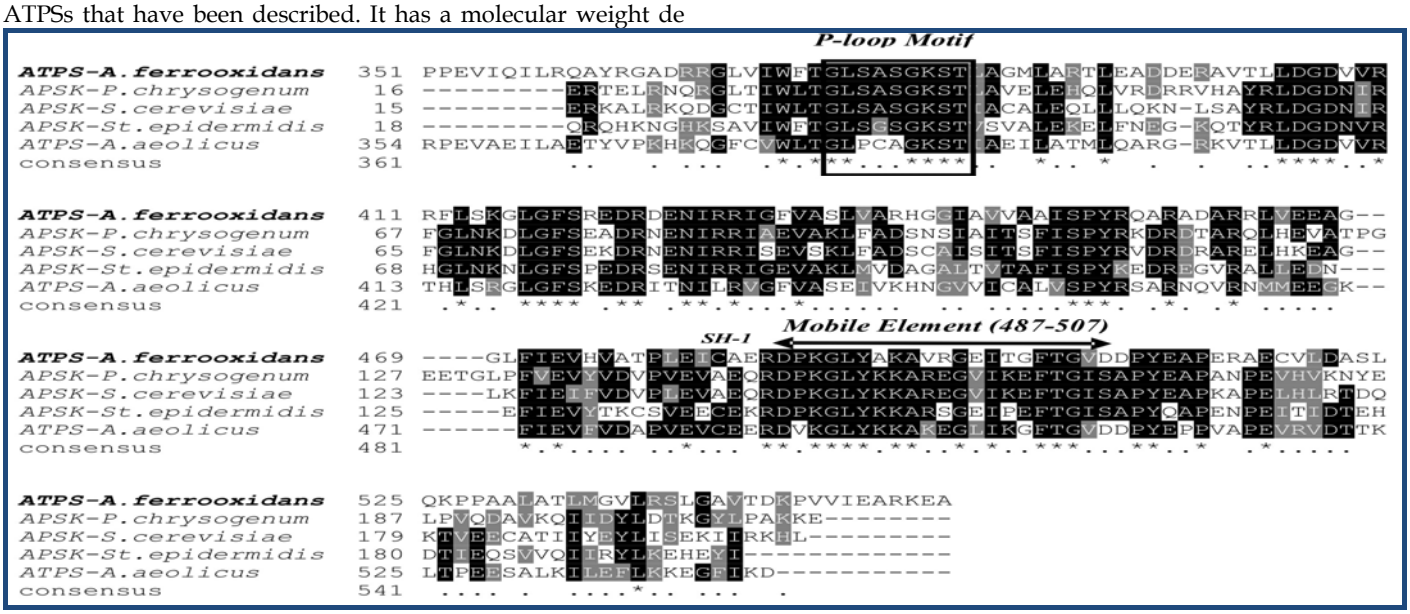

Figure 3: Multiple alignment of the APS kinase domain of ATPS from *A. ferrooxidans* (CAQ76453) and *A. aeolicus* (2GKS) with the APS kinase (APSK) from *P. chrysogenum* (1M7H), *S. cerevisiae* (CAA46252) and *St. epidermidis* (Q5HL02). Identical amino acids (black) and similar (gray) are indicated. The arrow indicates the conserved component of the ATPS from *A. ferrooxidans*: ^{487}DPKGLYAKAVRGEITGFTGVD^{507}. The black square indicates a P-loop region of the APS kinase that has identical sequence GLSASGKST in the ATPS from *A. ferrooxidans*.

Phylogenetic analysis of ATP sulfurylase

The construction of a phylogenetic tree by the Neighbor-Joining method shows that the homooligomeric ATPSs comprise fourth subgroups **(Figure 4).** *Subgroup 1*: This subgroup consists of those who have only the ATS domain and are found in most Gram positive bacteria, Gram negative bacteria, archaea, protista and plants. Our results reveal the presence of certain clades, such as cyanobacteria (*Nostoc*, *Anabaena*, *Trichodesmiun*,

Synechococcus and *Synechocystis*) with a bootstrap of 100%. *Brassica oleracea*, *Arabidopsis thaliana*, *Glycine max*, *Zea mays* and *Solanum suberosum* plants are grouped with a bootstrap of 100%. The ATPSs of *Phaedoactilum tricornutum*, *Thalassiosira pseudonana*, *Prochlorococcus marinus* and, *Chlamydomonas reinhardtii* are grouped (bootstrap 100%). *Subgroup 2*: Consist of ATPSs that have both ATS domains in the N-terminus and the ASK or ASK-like domain at the C-terminus found in fungi, yeasts, *Dictyostelium discoideum* amoebae, *Toxoplasma gondii* protozoa and some bacteria such as *A. ferrooxidans*, *A. ferrivorans* and *A. caldus*. ATPS from genus *Acidithiobacillus* is grouped with ATPS of fungi and of other bacteria that have both domains being located at the base of all ATPSs of this subgroup but with bootstrap of 49%. The filamentous fungi, *Aspergillus niger* and *P. chrysogenum* and yeast *S. cerevisiae*, *Schizosaccharomyces pombe* and *Candida albicans* form a clade with a bootstrap of 95%. Bacteria of the genus *Sulfitobacter*, *Roseobacter*, *Jannaschia*, *Oceanicola batsensis, Silicibacter pomeroyi* and *Rhodobacter sphaeroides* form a clade with high bootstrap value (100%). *Subgroup 3*: Consist of ATPS with both domains but with a different arrangement (ASK domain in the N-terminus and ATS domain in the C-terminus) which is present specifically in higher eukaryotes (Metazoa: animals and human). The ATPS of higher eukaryotic organisms (*Mus musculus, Cavia porcellus, Urechis caupo* and *Homo sapiens*) represent a clade with a bootstrap value of 99%. *Subgroup 4*: Consist de ATPSs with three domains (ASK domain in the N-terminus, ATS domain and Pyrophosphatase domain in the C-terminus) and are found in *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* diatoms and *Phytophthora infestans* forming a cluster (bootstrap 100%) meaning an evolutionary advantage in the way of the sulfate in these organisms.

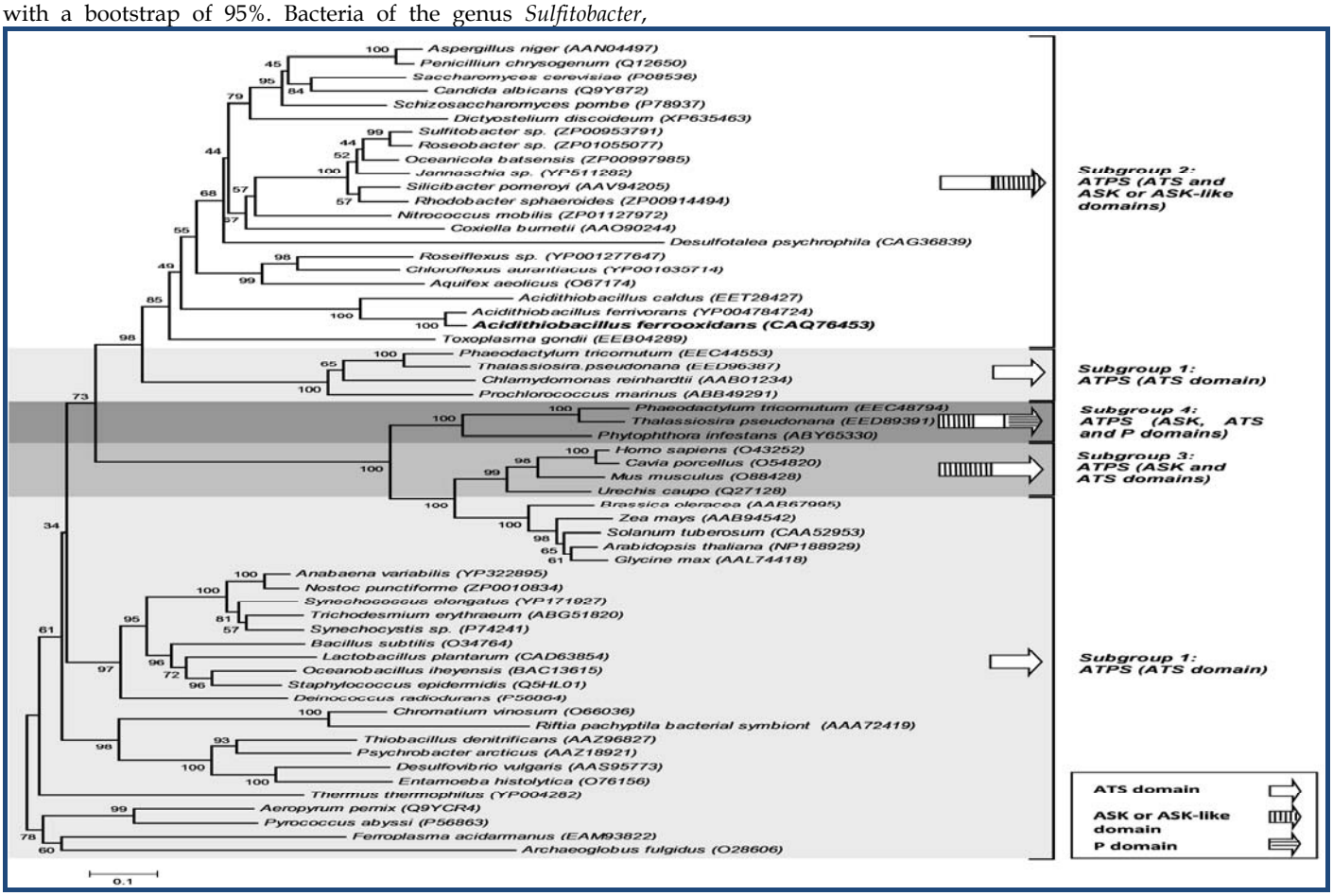

Figure 4: Phylogenetic tree of the homooligomerics ATPS rooted with the ATPS de archaea building by Neighbor Joining method from multiple alignment of amino acid sequences of the ATPS. An analysis of 1000 bootstrap replicates, using the Poisson correction model for Mega v.5.05 program. There are fourth subgroups de ATPS: 1) Some bacteria, archeae, algae, *Entamoeba histolytica*, *Phaeodactylum tricornutum, Thalassiosira pseudonana* diatomeas and plants with ATPS only with domain (sulfurylase: ATS), 2) fungi, yeasts, *Dictyostelium discoideum* amoebae, *Toxoplasma gondii* protozoa and some bacteria with the ATPS with both domains (sulfurylase: ATS and kinase: ASK o ASK-like) and 3) eukaryotic higher animals and man have the ATPS with both domains in the reverse order (kinase: ASK and sulfurylase: ATS), and 4) *Phaeodactylum tricornutum* and *Thalassiosira pseudonana* diatoms and *Phytophthora infestans* have the ATPS with three domains (kinase: ASK, sulfurylase: ATS, and Pyrophosphatase domain: P). Tree based on a comparison of 342 amino acids of the ATS domain of *A. ferrooxidans*.

Discussion:

The enzymes ATP sulfurylases (ATPSs) belong to a superfamily of proteins and are widely distributed among microorganisms, protista, plants, animals and humans **[2, 4, 5, 35]** involved in desassimilative sulfate reduction, activation of sulfate assimilation and sulfur oxidation. The direct oxidation pathway of sulfite is the most widely distributed, but photolithotrophic and chemolithotrophic bacteria belonging to the β- and γ-

Proteobacteria also have an indirect oxidation pathway **[7]**, where ATPS participate in the final oxidation reaction, producing ATP and SO_4^{2-} from APS and PPi **[1]**. We have cloned and overexpressed the ATP sulfurylase from *A. ferrooxidans* in *E. coli* as a recombinant protein of approximately 63.5 kDa bound to six histidines **(Figure 1),** being the monomer molecular weight of 60.49 kDa protein (based on its 557 amino acids) similar to the monomer molecular weight of ATP sulfurylase (62.8 kDa) of the thermophilic chemolithotroph *A. aeolicus* **[9]**. Probably, the ATPS from *A. ferrooxidans* has a dimeric structure similar to chemolithotroph *A. aeolicus* **[9]** because it has conserved amino acids involved in the dimerization in the ATPS from *A. aeolicus* (results not shown), whereas in the ATPS from *P. chrysogenum* the subunits dimerize and then form a triad of dimers **[36]**. These results are consistent with the proposal by Sperling (1998) *et al*. **[3]**, which each subunit of the monomeric or homooligomeric ATPSs would range from 41 to 69 kDa. The specific activity of ATPS from *A. ferrooxidans* for the synthesis of ATP determined from crude *E. coli* extracts is 106.9 units/mg/min. The specific activity for the ATPS purified from *A. aeolicus* was 13.1 units/mg proteins **[9]**. In contrast, The ATPS from *T. denitrificans* synthesizes around 400 units/g cell weight of ATP and the levels in the *R. pachyptila* symbiont is similar assuming that the bacterium represents 50% of trophosomal tissue **[9]**. The ATPS of *A. ferrooxidans* is functional in the direction of ATP synthesis involved in the oxidative metabolism of sulfite, and possibly with a low activity in the direction of APS synthesis (experiments not conducted) similar to the ATPS from *A. aeolicus* and the *R. pachyptila* symbiont **[1, 9]**. In organisms such as sulfate assimilators *P. chrysogenum* and *S. cerevisiae*, ATPS activity by the molybdolisis method was 60 units/mg of protein and 39 units/mg of protein, respectively **[37]**. However, by the luminometer method the reported activity was 140 units/mg proteins for ATPS from *S. cerevisiae* due to the addition of APS in the filtration step of purification **[38]**. Moreover, in higher eukaryotes, the ATPS that integrates human PAPS synthetase has an activity of 18.7 units/mg **[5]**. In plants, ATPS activity in leaves of spinach (*Spinacia oleracea* L.) is 23.1 nmol ATP/mg/min **[39]** and *Arabidopsis thaliana* is 0.012 units/mg **[40].** Some chemolithotrophic bacteria such as *Acidithiobacillus* could have three enzymes: sulfite-dependent cytochrome oxidase (APS), adenylyl phosphate (APAT) and ATP sulfurylase (ATPS), which could catalyze the reaction of terminal sulfate production in the complete oxidation of reduced sulfur compounds while others seem to possess one or two enzymes (**9**). The gene for the APS reductase (*cysH* gene) involved in the pathways of sulfate reduction and sulfide oxidation in the biological sulfur cycle **[41]** has been reported in *A. ferrooxidans*. The organisms that possess this APS reductase pathway are widely distributed in natural environments with high concentrations of sulfide or other reduced sulfur compounds and use this pathway to generate ATP or by attaching electrons to reduce CO_2 **[42]**. Although the conserved hypothetical gene (*orf2*) embedded in the *hdr* locus of sulfur oxidizers could also be involved, that strongly suggests that these microorganisms have a novel sulfur oxidation pathway in which sulfite is hypothesized to be produced in the cytoplasm by heterodisulfide reductase that in turn would catalyze the oxidation of sulfite to APS. Furthermore, the enzyme responsible for the second step in this pathways, has been reported by quantitative RT-PCR analysis of the *atpS* gene expression of *A. ferrooxidans* in the presence reduced inorganic sulfur compounds oxidation **[43]** and the presence of the an ORF to ATPS in the genome of the new specie described A. *ferrivorans* **[44],** which along with our results in the functionality of ATPS, leads us to suspect the involvement of ATPS from *A. ferrooxidans* in the oxidation of sulfite through the reverse path of APS reductase or a novel sulfur oxidation pathway.

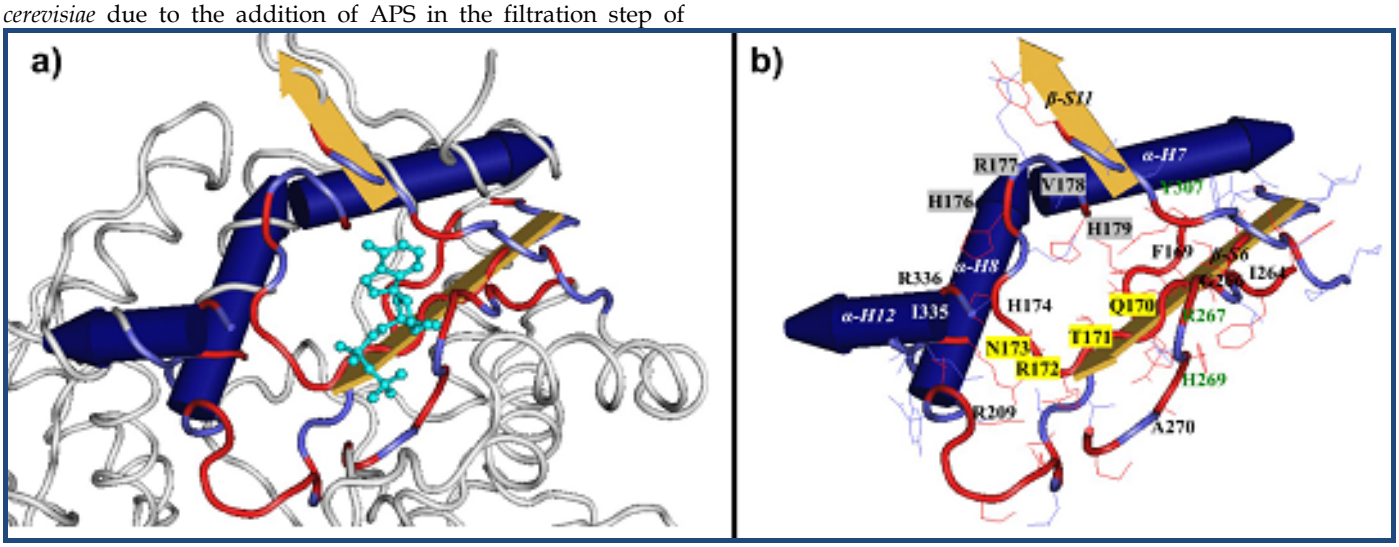

Figure 5: Structural alignment of the ATPS from *A. ferrooxidans* (CAQ76453) with ATPS of *A. aeolicus* (MMDB: ID44452; PDB: 2GKS). **a**) Active site of the ATS domain with an ADP molecule, and **b**) most important amino acid residues of the active site for function and binding nucleotides are marked with numbers in the ATPS sequence from *A. ferrooxidans*. The molecule in sky-blue is ADP. Identical amino acids are shown in red and similar ones in blue. The HXXH motif (in gray) stabilizes PPi and QXRN motif (in yellow) is the catalytic motif. Residues R267, H269 and Y307 (in green) are joined with the ATP.

The ATS domain of ATPS from *A. ferrooxidans* has five highly conserved regions or blocks similar to all the homooligomeric ATPSs from archaea to higher eukaryotes **(Figure 2),** with blocks II and IV rich in basic amino acids that participate in the binding of $MgATP^{2-}$and SO_4^{2-} **[3],** corresponding these blocks to domain II of homooligomeric ATPSs. Block II presents the ^{170}QXRN^{173}, ^{176}HXXH^{179} motifs and Block III presents the sequence of the active site loop

^{196}(hp)$_3$HXhpXGXXKXXDhpXXXXR^{215} motif **(Figure 2),** both blocks are important for enzyme activity **[1, 4].** An alignment with the structure of ATPS from *Aquifex aeolicus* shows that the active site loop motif presents Asp209 amino acid described in the ATPS from *P. chrysogenum* (Asp234) and *A. aeolicus* (Asp207) that would help to guide the Arg172 (Arg172 also in *A. aeolicus*) in the sulfate-binding site that is required for sulfurylase activity **[10]** and this amino acid would regulate the catalytic activity and preference for the of APS as substrate for the synthesis of ATP as in other chemolithotrophic bacteria and allow an open conformation of the loop for more time, compared to the Asp234 of the enzyme from *P. crhysogenum* (sulfate assimilator) that regulates the activity to active or inactive forms **[10].** The amino acids Arg172, Gln170 and Ala270 form the sulfate-binding site and amino acids Arg267 and His269 would interact with ribose. Whereas, Arg267 and Tyr307 (Leu307 in *A. aeolicus*) would be responsible for the preference of this enzyme by ATP over other nucleosides. These amino acids have been identified in the crystallized form of ATPS from *A. aeolicus* **(Figure 5).** However, so far it has not been determined which residues are involved in determining the direction of the physiological reaction of the ATPS enzyme. Possibly minor perturbations are responsible for the optimization of the kinetic properties for the physiologically relevant direction **[10].** The ATPS of the genus *Desulfuvibrio* **[45]** and some other ATPSs (*Archaeoglobus fulgidus*, *Entamoeba histolytica* and *Bacillus subtilis*) have the ability to bind metals, but the ATPS of *A. ferrooxidans* does not present the sequence responsible for binding to metal C-X2-C-X8-CXH **[7]** (results not shown).

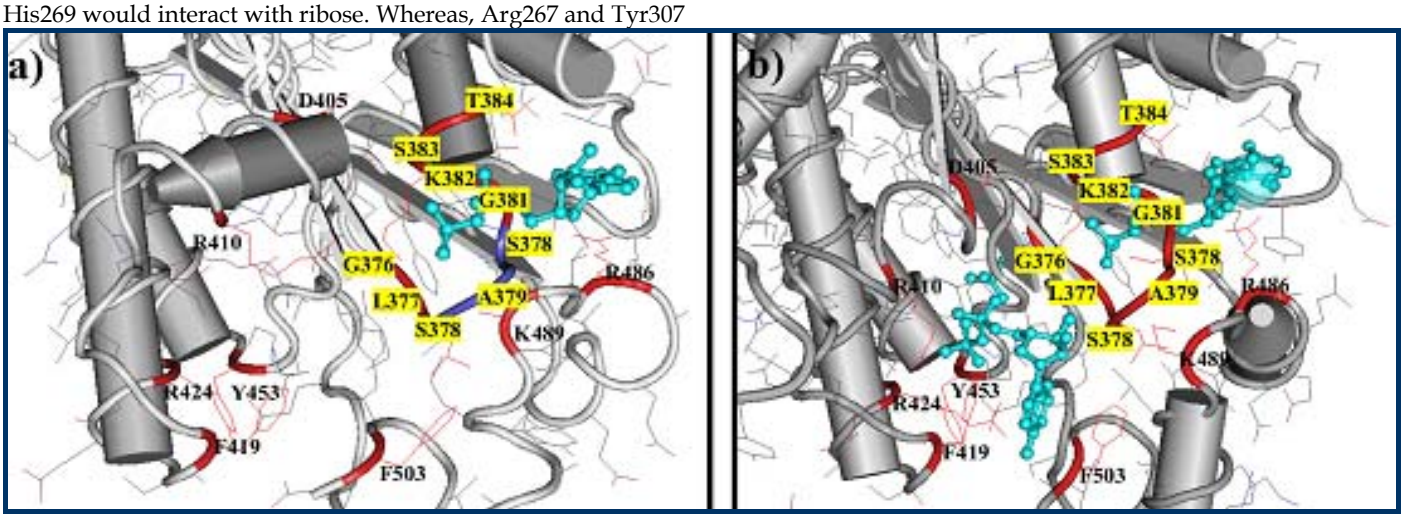

Figure 6: Comparative structural alignment of the ASK domain of ATPS from *A. ferrooxidans* (CAQ76453) with: **a**) ASK domain of ATPS from *A. aeolicus* (MMDB ID 44452, PDB: 2GKS_A); **b**) APS kinase from *P. chrysogenum* (MMDB ID: 21216, PDB: IM7G_A). It is shown the catalytic site of APS kinase domain with the P-loop motif (yellow). Red color indicates identical amino acids and blue indicates similar ones blue in the activity of true APS kinase. In **a**) the molecule lies in the ADP and in **b**) it does in the ATP (in sky blue). The numbering is based on the sequence of ATPS from *A. ferrooxidans*.

The ASK domain activity of ATPS from *A. ferrooxidans* remains to be established. However, it is possible an APS kinase activity because it has conserved amino acids compared to the ATPS from *A. aeolicus* that presents APS kinase activity **(Figure 3 & Figure 6) [9]**, showing "P-loop" ^{376}GLSASGKST^{384} motifs, whose sequence is identical to that found in true APS kinases (^{32}GLSASGKST^{40}) and is similar to the ATPS from *A. aeolicus* that presents the ^{379}GLPCAGKST^{387} motif. By contrast, in ATPS from filamentous fungi this motif in the ASK-like domain has changed and only conserves four amino acids: ^{403}GYMNSGKDA^{411} **[9]**. Other key residues of the ASK domain presents in the ATPS de *A. ferrooxidans* that have been described in the ATPS from *A. aeolicus* are present: a) Asp405 (Asp407 in *A. aeolicus*) that interacts with Mg^{+2} + ATP; b) Phe419 (Phe421 in *A. aeolicus*) and Phe503 (the same position in *A. aeolicus*), which would bind to the adenine ring of APS; c) the amino acid Arg410 (Arg412 in *A. aeolicus*) and Arg424 (Arg426 in *A. aeolicus*), which would associate the phosphosulfate group of APS, and d) Tyr453 (Tyr455 in *A. aeolicus*), which would help to align the Arg424 (similar to Arg426 in *A. aeolicus*). Additionally, e) the residue Lys489 (similar to *A. aeolicus*), as in true APS kinase, while in *P. chrysogenum* is Arg515 **[46]**. Furthermore, the arrangement of some helices, beta sheets and loops of this domain with true ASK kinase is very similar.

The evolutionary origin of homooligomeric ATPSs has not been determined; they have no similarity to heterooligomeric ATPSs. These two classes of ATPS with the same function are probably originated by convergent evolution **[3].** The phylogenetic tree of homooligomeric ATPSs rooted with archaeal ATPSs leads us to suppose that probably the evolution of this enzyme started from the existence of ancestral *atpS* gene similar to *subgroup 1* of the homooligomeric ATPSs with only ATS domain, presents in archaea and in Gram-positive bacteria mainly. Later, from ATPS ancestral gene would have originated in two subgroups: a) The *subgroup 2* of the homooligomeric ATPSs which has two domains (ATS in the N-terminal and ASK or ASK-like in the C-terminal) and b) *subgroup 3* of the homooligomeric ATPSs who have both domains in the reverse order to *subgroup 2*. Possibly the *subgroup 2* is originated from a fusion of ancestral *atpS* gene with a gene encoding the APS kinase protein giving rise to a primitive bifunctional ATPS, whose homologous representative could be the bifunctional ATPS found recently in chemolithotrophic bacteria *A. aeolicus* from which the ATP sulfurylase of fungi would have evolved **[10].** The ATPS from *A. ferrooxidans* has a high level of homology to the C-terminal domain (kinase domain) of the ATPS from *A. aeolicus* and presents an identical P-loop region to that of true APS kinases **(Figures 3 & 6),** which probably suggests that *A. ferrooxidans*

possesses a functional enzyme of the *subgrupo 2* more ancestral than that in *A. aeolicus*. The most basal location of the enzyme from *A. ferrooxidans* with A. *ferrivorans* and *A. caldus* in the phylogenetic tree would corroborate our hypothesis **(Figure 4).** Moreover, the *subgroup 3* would have originated during the course of evolution after the divergence of an ancestral ATPS similar to *subgroup 1* gave rise first to the ATPS of plants and then by fusion with the gene for APS kinase generated a bifunctional enzyme called PAPS synthetase in metazoans (ASK domain and ATS domain) as postulated earlier **[47]**. Alternatively, the domains of the ATPSs *subgroup 2* or *subgroup 3* probably originated by a rearrangement of domains in the bifunctional enzyme similar to primitive bacterial ATPS of the *subgroup 2* as *A. aeolicus* and *A. ferrooxidans*, and are possibly intermediate forms between the ATPS of fungi and metazoan PAPS synthetases as postulated Hanna (2002) *et al*. **[9]**. Additionally, *subgroup 4* includes ATP sulfurylases with ATS, ASK and pyrophosphatase domains would have originated from *subgroup 3* during the course of evolution. This is postulated the variations of ATPSs are the result of ancestral fusion genes evolving by an assortment of gene fissions, duplications, deletions, and horizontal transfers in different lineages **[48]**. Our phylogenetic tree is consistent with the possibility that the enzyme ATP sulfurylase from the chemolithotrophic *Riftia* symbiont (subgroup 1) is most similar to ancestral ATP sulfurylase from which the family of homooligomeric ATP sulfurylase would have evolved **[1]** but probably there are other ancestral ATPS present in other microorganisms **(Figure 4).**

During the evolution of the homoligomeric ATPSs, occurred horizontal *atpS* gene transfer events between some organisms whose expression allowed them to adapt their metabolism and lifestyle **[1]**. Similar results were obtained by Patron *et al*. (2008) **[35]** whose proposed that the inheritance of the enzymes of the ATPS, APR and PAPR have multiple origins in lineages that comprise *opisthokonts* (fungi and animals), gene fusions with other enzymes of sulphate assimilation pathway and evidence an eukaryote-to-prokaryote lateral gene transfer. Some organisms (*Thiobacillus denitrificans*, *Phaeodactylum tricornutum*, *Thalassiosira pseudonana*) present yet ATPS with only ATS domain and ATPS with ATS; ASK domains or fusion with pirophosphatase domain. The crystallography of the putative bifunctional ATPS (with ATS and ASK domains) from *Thiobacillus denitrificans* has shown that it only presents APS kinase activity, and exhibits numerous structural and sequence differences in the ATS domain to other ATPSs that render it inactive with respect to ATP sulfurylase activity, probably has unknown function **[49]**. The fusion of ATPS has happened not only with the gene for APS kinase, but with the inorganic pyrophosphatase enzyme in stramenopiles (in the diatom *Thalassiosira pseudonana* and *Phaeodactylum tricornutum*, and the oomycete *Phytophtohora sojae*) and on haptofites (algae such as *Pavlova lutheria* and *Emiliania huxleri*). Furthermore, in *Heterocapsa triquetra* is found the fusion of ATPS to APR, which probably would ensure a rapid transition from APS to the site of its reduction by increasing the production rate of sulfite **[35]**. Genomic analysis of the *atpS* gene region of *A. ferrooxidans* ATCC 23270 (16 kb) shows that the *atpS* gene is not associated with other genes in the metabolism of sulfur compounds especially with the APS reductase involved in the indirect route of oxidation of sulphite, but is associated with other transferases and chaperonins. An ORF that encoded the APS reductase of *A. ferrooxidans* was found in the genome the *A. ferrooxidans* ATCC 23270 strain but not adjacent to the *atpS* gene. By contrast, in the phototrophic sulfur oxidizer *Chromatium vinosum*, genes encoding for ATP sulfurylase and APS reductase (*aprMBA*, *aprM* encodes a membrane anchor protein) form an operon **[50]**. In the sulfate reducing archaeon *A. fulgidus*, the *aprC* gene that encodes a soluble protein with no known function, is inserted between the *atpS* gene (called *sat*) and the *aprBA* gene **[3].** In the genome of *C. tepidum*, the genes for ATP sulfurylase and APS reductase are located adjacent to each other **[7]**.

Conclusion:

We have demonstrated the expression of the gene encoding the enzyme ATP sulfurylase in the chemolithotrophic bacterium *A. ferrooxidans* 23270 that it would participate in the indirect pathway of sulfide oxidation to obtain energy. It presents an homooligomeric ATPS with: **a)** similarity at the sequence level and structure to homooligomeric ATPS with conserved motifs and mobile loop at the active site, **b**) Five blocks present in all homooligomeric ATPS, **c)** Enzyme activity producing ATP from APS and PPi, **d)** Presence of amino acids similar to ATPS from *A. aeolicus* involved in dimer formation, **e)** similarity to *subgroup 2* of the homooligomeric ATPS, and **f**) size protein and cytoplasmic location. Subsequent studies of the ATPS present in *A. ferrooxidans* as purification and crystallization involved in the indirect oxidation of sulfite will be vital in understanding the mechanism of acid drainage generation used in bioleaching processes to improve the recovery rate of metals.

Acknowledgment:

This work was supported in part by grant from CONCYTEC (contract Nº 245-03-CONCYTEC-OAJ). LJdV was supported by grant from AECID (A/9727/07). MLJ was the recipient CONCYTEC Master in Molecular biology scholarship.

References:

[1] Beynon JD *et al. Biochemistry.* 2001 **40**: 14509 [PMID: 11724564]
[2] Geller DH *et al. J Biol Chem.* 1987 **262:** 7374 [PMID: 3034891]
[3] Sperling D *et al. FEMS Microbiol Lett.* 1998 **162:** 257 [PMID: 9627961]
[4] Hanna E *et al. J Biol Chem.* 2004 **279:** 4415 [PMID: 14613928]
[5] Lansdon E *et al. Biochemistry.* 2004 **43:** 4356 [PMID: 15065880]
[6] Bick JA *et al. J Bacteriol.* 2000 **182:** 135 [PMID: 10613872]
[7] Kappler U & Dahl C, *FEMS Microbiol Lett.* 2001 **203:** 1 [PMID: 11557133]
[8] Valdes J *et al. BMC Genomics.* 2003 **4:** 51 [PMID: 14675496]
[9] Hanna E *et al. Arch Biochem Biophys.* 2002 **406:** 275 [PMID: 12361716]
[10] Yu Z *et al. J Mol Biol.* 2007 **365:** 732 [PMID: 17095009]
[11] Rosenthal E & Leustek T**,** *Gene.* 1995 **165:** 243 [PMID: 8522184].
[12] Jullien D *et al. Mech Dev.* 1997 **68:** 179 [PMID: 9431815]
[13] Li H *et al. J Biol Chem.* 1995 **270:** 29453 [PMID: 7493984]
[14] Venkatachalam K *et al. J Biol Chem.* 1998 **273:** 19311 [PMID: 9668121]
[15] Lalor DJ *et al. Protein Eng.* 2003 **16:** 1071 [PMID: 14983089]
[16] Taguchi Y *et al. Acta Crystallogr D Biol Crystallogr.* 2003 **59:** 1645 [PMID: 12925800]

[17] Ullrich TC *et al. EMBO J. 2001* **20:** 316 [PMID: 11157739]
[18] Ramírez P *et al. Appl Environ Microbiol.* 2002 **68:** 1837 [PMID: 11916703]
[19] Altschul S *et al. Nucleic Acids Res.* 1997 **25:** 3389 [PMID: 9254694]
[20] http://www.sequentix.de
[21] Bradford MM, *Anal Biochem.* 1976 **72:** 248 [PMID: 942051]
[22] http://www.expasy.ch/cgi-bin/protparam
[23] Bhasin M *et al. Bioinformatics.* 2005 **21:** 2522 [PMID: 15699023]
[24] http://www.psort.org/psortb/
[25] http://www.cbs.dtu.dk/services/SignalP/
[26] http://www.ncbi.nlm.nih.gov/Ftp/
[27] Thompson J *et al. Nucleic Acids Res.* 1994 **22:** 4673 [PMID: 7984417].
[28] http://www.mbio.ncsu.edu/bioedit/bioedit.html
[29] http://www.ch.embnet.org/software/BOX_form.html
[30] Hogue CW, *Trends Biochem Sci.* 1997 **22:** 314 [PMID: 9270306]
[31] http://www.ncbi.nlm.nih.gov
[32] Thompson J *et al. Nucleic Acids Res.* 1997 **25:** 4876 [PMID: 9396791]
[33] Kumar S *et al. Brief Bioinform.* 2004 **5:** 150 [PMID: 15260895]
[34] Saitou N & Nei M, *Mol Biol Evol.* 1987 **4:** 406 [PMID: 3447015]
[35] Patron N *et al. BMC Evol Biol.* 2008 **8:** 39 [PMID: 18248682]
[36] MacRae IJ *et al. Biochemistry.* 2001 **40:** 6795 [PMID: 11389593]
[37] Foster B *et al. J Biol Chem.* 1994 **269:** 19777 [PMID: 8051058]
[38] Karamohamed S *et al. Protein Expr Purif.* 1999 **15:** 381 [PMID: 10092498]
[39] Lunn JE *et al. Plant Physiol.* 1990 **94:** 1345 [PMID: 16667839]
[40] Leustek T *et al. Plant Physiol.* 1994 **105:** 897 [PMID: 8058839]
[41] Heng C *et al. Curr Microbiol.* 2009 **58:** 586 [PMID: 19225840]
[42] Aue BE & Nelson DC, *J Bacterial.* 1994 **176:** 3723 [PMID: 8206850]
[43] Quatrini R *et al. BMC Genomics.* 2009 **10:** 394 [PMID: 19703284]
[44] Liljeqvist M *et al. J Bacteriol.* 2011 **193:** 4304 [PMID: 21705598]
[45] Gavel O *et al. Biochemistry.* 1998 **37:** 16225 [PMID: 9819214]
[46] Medina D *et al. Arch Biochem Biophys.* 2001 **393:** 51 [PMID: 11516160]
[47] Kurima K *et al. J Biol Chem.* 1999 **274:** 33306 [PMID: 10559207]
[48] Bradley M *et al. J Mol Evol.* 2009 **68:** 1 [PMID: 19067028]
[49] Gay SC *et al. Acta Crystallogr D Biol Crystallogr.* 2009 **65:** 1021 [PMID: 19770499]
[50] Hipp W *et al. Microbiology.* 1997 **143:** 2891 [PMID: 9308173]

Edited by P Kangueane
Citation: Jaramillo *et al.* Bioinformation 8(15): 695-704 (2012)