GENOME ANNOUNCEMENT

Draft Genome Sequence of the Extremely Acidophilic Bacterium Acidithiobacillus caldus ATCC 51756 Reveals Metabolic Versatility in the Genus Acidithiobacillus

Jorge Valdes,1 Raquel Quatrini,1 Kevin Hallberg,2 Mark Dopson,3 Pablo D. T. Valenzuela,1 and David S. Holmes1*

Center for Bioinformatics and Genome Biology, Fundación Ciencia para la Vida and Departamento de Ciencias Biologicas, Facultad de Ciencias Biologicas, Universidad Andres Bello, Santiago, Chile1; School of Biological Sciences, College of Natural Sciences, Bangor University, Bangor LL57 2UW, United Kingdom2; and Department of Molecular Biology, Umeå University, SE-901 87 Umeå, Sweden3

Received 26 June 2009/Accepted 6 July 2009

Acidithiobacillus caldus is an extremely acidophilic, moderately thermophilic, chemolithoautotrophic gammaproteobacterium that derives energy from the oxidation of sulfur and reduced inorganic sulfur compounds. Here we present the draft genome sequence of Acidithiobacillus caldus ATCC 51756 (the type strain of the species), which has permitted the prediction of genes for survival in extremely acidic environments, including genes for sulfur oxidation and nutrient assimilation.

Acidothiobacillus caldus is one of the three recognized species of the genus Acidithiobacillus, which also circumscribes A. thiooxidans and A. ferrooxidans. These bacteria live in extremely acidic environments (pH 1 to 3) typically associated with copper mining operations (bioleaching) (15, 17) and natural acid drainage systems (7). All of these bacteria have the capacity to gain energy by the oxidation of sulfur and reduced inorganic sulfur compounds and to thrive in extremely high concentrations of heavy metals (16). Of the three species, A. ferrooxidans is unique in also being able to obtain energy through the oxidation of ferrous iron, as well as being a facultative anaerobe capable of using ferric iron as an alternative electron acceptor. Acidithiobacilli have been shown to be able to fix atmospheric carbon via the Calvin-Benson-Bassham cycle (1, 11, 21) and to synthesize extracellular polymeric substances that are postulated to promote adhesion to mineral surfaces (3).

As opposed to A. ferrooxidans, for which substantial bioinformatic and experimental evidence exists for these and other properties (4, 14, 19, 20), A. caldus is poorly characterized, although it is known that it cannot carry out iron oxidation or nitrogen fixation (13). In contrast to the other two species of the genus, A. caldus thrives at temperatures up to 45 to 50°C. In order to unravel strategies for energy and nutrient assimilation used by A. caldus to survive and proliferate in extremely acidic environments, we have generated and annotated a draft genome sequence of A. caldus and performed a genome-based metabolic reconstruction to address these questions.

The draft genome sequence of the type strain of A. caldus, ATCC 51756, was determined by a whole-genome shotgun strategy. Genomic libraries of 4 kb and 40 kb were constructed and sequenced, assembled using the Phred/Phrap programs (5), leading to the generation of a draft assembly based on 41,813 high-quality reads. Using Consed (8), assemblies that contained only contig segments with at least twofold coverage were edited and curated. Gene modeling was performed using CRITICA (2) and Glimmer (6). Predicted coding sequences were annotated based on comparisons with public databases (COG, KEGG, Pfam, TIGRFAMs, Uniprot, and NR-NCBI). Automatic metabolic reconstruction was carried out using the PRIAM and Pathways tools for prediction and curation.

The A. caldus ATCC 51756 draft genome sequence has a total of 2,946,159 bp distributed in 139 contigs with an average GC content of 61.4%. Two 5S-16S-23S operons and 47 tRNAs on the draft assembly were identified, as were complete sets of genes for the synthesis of amino acids, nucleotides, and prosthetic groups. As in the case of A. ferrooxidans ATCC 23270 and other chemolithoautotrophic representatives, an incomplete tricarboxylic acid (TCA) cycle was detected, in which genes for the 2-oxoglutarate dehydrogenase enzyme complex were absent. The incomplete TCA cycle has been hypothesized to be an ancient biosynthetic pathway, instead of an energy generation cycle characteristic of the complete TCA cycle (22).

A detailed inspection of the genome sequence revealed the presence of a complete set of genes encoding flagellum formation and chemotaxis. All of the genes of the classical Calvin-Benson-Bassham pathway were predicted, including genes for two RuBisCO (ribulose-1,5-bisphosphate carboxylase oxygenase) enzymes (type I and type II) and carboxysome formation.
genes. In addition, genes for sulfur and reduced inorganic sulfur compound oxidation were identified, including two gene clusters harboring the genes encoding the sulfur oxidation complex SOX (soxYZB-hyp-resB-soxAX-resC and soxYZA-hyp-SoxB), previously characterized in neutrophilic, chemolithoautotrophic bacteria (9); for the sulfur quinone oxidoreductase enzyme (sgr); a sulfur oxy- genase:reductase gene (sor); and genes for a tetrahalothionate hydrolase and a thiosulfate quinone oxidoreductase (dosD) previously characterized in this strain (18). Several terminal oxidases were identified, including two copies of the genes encoding a \( b_o \)-type quinol oxidase (cyoBACD), six copies of the genes encoding a \( b_d \)-type quinol oxidase (cydAB), and one putative \( a_o \)-type quinol oxidase gene cluster termed “quoxBACD.” No gene was detected that could encode rusticyanin, which has been shown to be involved in electron flow during iron oxidation in \( A. \) ferrooxidans ATCC 23270 (10).

Genes for ammonia uptake were predicted, while those for atmospheric nitrogen fixation were not found in the draft genome as expected. In addition, an alternative nitrogen assimilatory pathway in \( A. \) caldus can be proposed based on the presence of a gene cluster encoding a membrane-associated nitrate reductase (narGHI). A similar complex has been shown to carry out nitrogen assimilation in \( M. \) tuberculosis (12).

The information provided in the draft genome sequence of \( A. \) caldus ATCC 51756 reported here will facilitate additional bioinformatic and experimental investigations to elucidate the role of this microorganism in biotreatment and in natural and man-made acidic environments. This information also provides a first overview of the comparative metabolic diversity of the genus \( A. \) acidithiobacillus and generates a more comprehensive picture of the metabolic diversity and adaptability of organisms that dwell in extreme acidic environments.

**Nucleotide sequence accession numbers.** This whole genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the project accession no. ACVD00000000. The version described in this paper is the first version no. ACVD10000000.

This work was supported by Fondecyt 11060164 and 1090451, DI-UNAB 34-06 and DI-UNAB 15-06/I, and a Microsoft Sponsored Research Award.

**REFERENCES**


