First Insights into the Genome of the Amino Acid-Metabolizing Bacterium *Clostridium litorale* DSM 5388

Anja Poehlein,[a] Hamed S. Alghaithi,[b] Lenin Chandran,[b] Cynthia M. Chibani,[b] Elena Davydova,[b] Karthikeyan Dhamotharan,[b] Wanwan Ge,[b] David A. Gutierrez-Gutierrez,[a] Advait Jagirdar,[b] Bahar Khonsari,[b] Kamal Prakash P. R. Nair,[b] Rolf Daniel[a]

Genomic and Applied Microbiology and Göttingen Genomics Laboratory, Georg-August University, Göttingen, Germany[a]; Members of the Applied Bioinformatics in Microbiology Course of the Microbiology and Biochemistry MSc/PhD program, Georg-August University, Göttingen, Germany[b]

***Clostridium litorale* is a Gram-positive, rod-shaped, and spore-forming bacterium, which is able to use amino acids such as glycine, sarcosine, proline, and betaine as single carbon and energy sources via Stickland reactions. The genome consists of a circular chromosome (3.41 Mb) and a circular plasmid (27 kb).**

Address correspondence to Rolf Daniel, rdaniel@gwdg.de.

The obligate anaerobic spore-forming bacterium *Clostridium litorale* is Gram-positive and rod-shaped. It is able to utilize amino acids such as glycine, sarcosine, proline, and betaine as sole carbon and energy sources via Stickland reactions (1, 2). *C. litorale* belongs, like its closest relative *Eubacterium acidaminophilum*, to cluster XI of clostridia (3, 4). The *C. litorale* type strain DSM 5388 was isolated from anoxic marine mud of the North Sea, Germany (1).

C. litorale was cultivated under anaerobic conditions using a mixture of alanine (30 mM) and betaine (50 mM) as substrate (5). The MasterPure complete DNA purification kit (Epicentre, Madison, USA) was used to isolate chromosomal DNA of *C. litorale* DSM 5388. The extracted DNA was employed to generate 454-shotgun and Illumina-shotgun libraries according to the protocols of the manufacturers. The libraries were sequenced using a 454 GS-FLX system (Titanium chemistry, Roche Life Science, Mannheim, Germany) and Genome Analyzer II (Illumina, San Diego, CA, USA). Sequencing resulted in coverages of 22.31 and 64.42, respectively. The assembly of the genome sequence was performed with the Roche Newbler assembly software 2.9, Mira 3.4, and SPAdes 3.0 (6), and resulted in 22 contigs.

The draft genome of *C. litorale* DSM 5388 comprises a single chromosome (3.41 Mb) and a circular plasmid (27 kb) with an overall G+C content of 41.3%. Automatic gene prediction was performed using the software tools YACOB and GLIMMER (7). Identification of rRNA and tRNA genes was done with RNAmmer and tRNAscan, respectively (8, 9). The IMG ER (integrated microbial genomes-expert review) system (10, 11) was used for automatic annotation, which was subsequently manually curated by using the Swiss-Prot, TREMBL, and InterPro databases (12). The draft genome harbored 14 rRNA operons, 77 tRNA genes, and 2,159 protein-encoding genes with function prediction and 933 genes coding for hypothetical proteins.

The genome of *C. litorale* harbors all genes coding for proteins necessary for degradation of glycine, sarcosine, and betaine. Two putative gene clusters encoding a glycine reductase complex were detected. In addition, one putative gene cluster for a betaine reductase complex and one for a sarcosine reductase complex were identified. These gene clusters showed a structure similar to those identified in *C. sticklandii*, *C. difficile*, *Sporomusa ovata*, or *E. acidaminophilum* (13–15). We could also identify a proline reductase gene cluster, which contains a gene coding for a protein (PrdC) with high sequence similarity to the C subunit of the Rnf complex. This gene is also present in *C. sticklandii* (13). Subunits A (GrdA) and B (GrdB, GrdF, GrdH, PrdB) of all these reductases are selenocysteine-containing proteins. In addition, we detected all genes (*selABC*) and the tRNA necessary for incorporation of selenocysteine into proteins. We could also identify a small selenocysteine containing protein (PrpU), which harbors a CxxU redox active motif. This has been previously only identified in *E. acidaminophilum* (15).

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. JJMM00000000. The version described here is version JJMM01000000.

REFERENCES

1. **Fendrich C, Hippe H, Gottschalk G.** 1990. *Clostridium halophilum* sp. nov. and *C. litorale* sp. nov., an obligate halophilic and a marine species degrading betaine in the Stickland reaction. Arch. Microbiol. **154:** 127–132. http://dx.doi.org/10.1007/BF00423321.
2. **Dietrichs D, Meyer M, Rieth M, Andreesen JR.** 1991. Interaction of selenoprotein PA and the thioredoxin system, components of the NADPH-dependent reduction of glycine in *Eubacterium acidaminophilum* and *Clostridium litorale*. J. Bacteriol. **173:**5983–5991.
3. **Baena S, Fardeau ML, Woo TH, Ollivier B, Labat M, Patel BK.** 1999. Phylogenetic relationships of three amino-acid-utilizing anaerobes, *Selenomonas acidaminovorans*, 'Selenomonas acidaminophila' and *Eubacterium acidaminophilum*, as inferred from partial 16S rDNA nucleotide sequences and proposal of *Thermanaerovibrio acidaminovorans* gen. nov., comb. nov.

and *Anaeromusa acidaminophila* gen. nov., comb. nov. Int. J. Syst. Bacteriol. **49:**969–974. http://dx.doi.org/10.1099/00207713-49-3-969.

4. **Collins MD, Lawson PA, Willems A, Cordoba JJ, Fernandez-Garayzabal J, Garcia P, Cai J, Hippe H, Farrow JA.** 1994. The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations. Int. J. Syst. Bacteriol. **44:**812–826. http://dx.doi.org/10.1099/00207713-44-4-812.

5. **Zindel U, Freudenberg W, Rieth M, Andreesen JR, Schnell J, Widdel F.** 1988. *Eubacterium acidaminophilum* sp. nov., a versatile amino acid-degrading anaerobe producing or utilizing H2 or formate. Description and enzymatic studies. Arch. Microbiol. **150:**254–266. http://dx.doi.org/10.1007/BF00407789.

6. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. J. Comput. Biol. **19:**455–477. http://dx.doi.org/10.1089/cmb.2012.0021.

7. **Tech M, Merkl R.** 2003. YACOP: Enhanced gene prediction obtained by a combination of existing methods. In Silico Biol. **3:**441–451.

8. **Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, Ussery DW.** 2007. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. **35:**3100–3108. http://dx.doi.org/10.1093/nar/gkm160.

9. **Lowe TM, Eddy SR.** 1997. tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964. http://dx.doi.org/10.1093/nar/25.5.0955.

10. **Markowitz VM, Mavromatis K, Ivanova NN, Chen IM, Chu K, Kyrpides NC.** 2009. IMG ER: a system for microbial genome annotation expert review and curation. Bioinformatics **25:**2271–2278. http://dx.doi.org/10.1093/bioinformatics/btp393.

11. **Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A, Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN, Kyrpides NC.** 2012. IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res. **40:**D115–D122. http://dx.doi.org/10.1093/nar/gkr1044.

12. **Zdobnov EM, Apweiler R.** 2001. InterProScan—an integration platform for the signature-recognition methods in InterPro. Bioinformatics **17:**847–848. http://dx.doi.org/10.1093/bioinformatics/17.9.847.

13. **Fonknechten N, Chaussonnerie S, Tricot S, Lajus A, Andreesen JR, Perchat N, Pelletier E, Gouyvenoux M, Barbe V, Salanoubat M, Le Paslier D, Weissenbach J, Cohen GN, Kreimeyer A.** 2010. *Clostridium sticklandii*, a specialist in amino acid degradation: revisiting its metabolism through its genome sequence. BMC Genomics **11:**555. http://dx.doi.org/10.1186/1471-2164-11-555.

14. **Poehlein A, Gottschalk G, Daniel R.** 2013. First insights into the genome of the gram-negative, endospore-forming organism *Sporomusa ovata* Strain H1 DSM 2662. Genome Announc. **1**(5):e00734-13. http://dx.doi.org/10.1128/genomeA.00734-13.

15. **Poehlein A, Andreesen JR, Daniel R.** 2014. Complete genome sequence of the amino acid-utilizing *Eubacterium acidaminophilum* al-2 (DSM 3953). Genome Announc. **2**(3):e00573-14. http://dx.doi.org/10.1128/genomeA.00573-14.