Complete Genome Sequence of Rnf- and Cytochrome-Containing Autotrophic Acetogen *Clostridium aceticum* DSM 1496

Anja Poehlein,[a] Frank R. Bengelsdorf,[b] Bettina Schiel-Bengelsdorf,[b] Gerhard Gottschalk,[a] Rolf Daniel,[a] Peter Dürre[b]

Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Georg-August University Göttingen, Göttingen, Germany[a]; Institut für Mikrobiologie und Biotechnologie, Universität Ulm, Ulm, Germany[b]

Here, we report the closed genome sequence of *Clostridium aceticum*, an Rnf- and cytochrome-containing autotrophic acetogen that is able to convert CO_2 and H_2 to acetate using the Wood-Ljungdahl pathway. The genome consists of a circular chromosome (4.2 Mbp) and a small circular plasmid (5.7 kbp).

Address correspondence to Peter Dürre, peter.duerre@uni-ulm.de.

*C*lostridium aceticum, an anaerobic and endospore-forming organism, was the first bacterial isolate described to be able to form acetate autotrophically from a $CO_2 + H_2$ gas mixture. This organism is also able to grow heterotrophically on sugars, organic acids, and alcohols. *C. aceticum* was originally isolated from sludge from a town canal in Wageningen in 1936 (1–3), thought to be lost during World War II, and was reisolated from a spore preparation by one of us (G.G.) in 1979 (4). Chromosomal DNA was isolated using the MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA). 454 shotgun and paired-end libraries and Illumina shotgun libraries were generated from the extracted DNA according to the protocol of the manufacturer. Sequencing was performed using a 454 GS-FLX system (Titanium GS70 chemistry; Roche Life Science, Mannheim, Germany) and a Genome Analyzer II (Illumina, San Diego, CA) resulting in 374,679 total 454 shotgun reads containing 182,232 paired reads with an average pair distance of 5.7 kb and a pair distance deviation of 1.4 kb and 2,280,716 Illumina 112-bp paired-end reads. Mira 3.4 (5) software and Roche Newbler Assembly 2.3 software were used to perform the hybrid *de novo* assembly resulting in 77 contigs within 24 scaffolds with an average coverage of 76.97-fold. The remaining gaps were closed by PCR-based techniques and primer walking with Sanger sequencing of the products using BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technology GmbH, Darmstadt, Germany). The closed genome of *C. aceticum* consists of a circular chromosome (4.2 Mbp) and a small circular plasmid (5.7 kbp) with an overall G+C content of 35.3%. Automatic gene prediction was performed by using the software tool Prodigal (6). Genes coding for rRNA and tRNA were identified using RNAmmer (7) and tRNAscan (8), respectively. The Integrated Microbial Genomes-Expert Review (IMG-ER) system (9) was used for automatic annotation, which was subsequently manually curated by using the Swiss-Prot, TrEMBL, and InterPro databases (10). The genome harbored 6 rRNA cluster, 74 tRNA genes, 3,181 protein-coding genes with predicted functions, and 743 genes coding for hypothetical proteins. Genes coding for enzymes involved in the methyl and carbonyl branch of the Wood-Ljungdahl pathway are organized in one large gene cluster together with P2 and P3 proteins of the glycine decarboxylase (11). This cluster showed the same arrangement as identified in other autotrophic clostridia such as *C. ljungdahlii* (12) and *C. autoethanogenum* (13). Genome analysis also revealed the presence of genes encoding the Rnf complex and all genes necessary for cytochrome synthesis, but no genes coding for quinone biosynthesis are present. *C. aceticum* is also able to grow heterotrophically on fructose, fumarate or malate, but might also be able to use glycine or betaine as energy and carbon source, as gene clusters coding for a glycine and betaine reductase comparable to those identified in *Eubacterium acidaminophilum* (14) and *Sporomusa ovata* (15) are encoded.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession numbers CP009687 and CP009688.

REFERENCES

1. **Wieringa KT**. 1936. Over het verdwijnen van waterstof en koolzuur onder anaerobe voorwaarden. Ant Leeuwenhoek **3**:263–273. http://dx.doi.org/10.1007/BF02059556.
2. **Wieringa KT**. 1939. The formation of acetic acid from carbon dioxide and hydrogen by anaerobic spore-forming bacteria. Ant Leeuwenhoek **6**:251–262. http://dx.doi.org/10.1007/BF02146190.
3. **Lux MF, Drake HL**. 1992. Re-examination of the metabolic potentials of the acetogens *Clostridium aceticum* and *Clostridium formicoaceticum*: chemolithoautotrophic and aromatic-dependent growth. FEMS Microbiol Lett **95**:49–56. http://dx.doi.org/10.1111/j.1574-6968.1992.tb05341.x.
4. **Braun M, Mayer F, Gottschalk G**. 1981. *Clostridium aceticum* (Wieringa), a microorganism producing acetic acid from molecular hydrogen and carbon dioxide. Arch Microbiol **128**:288–293. http://dx.doi.org/10.1007/BF00422532.
5. **Chevreux B, Wetter T, Suhai S**. 1999. Genome sequence assembly using

trace signals and additional sequence information, p 45–56. *In* Wingender E (ed), Computer science and biology: proceedings of the German Conference on Bioinformatics (GCB) 1999 Hannover, Germany. GBF-Braunschweig. Department of Bioinformatics, Braunschweig, Germany.

6. **Hyatt D, Chen GL, Locascio PF, Land ML, Larimer FW, Hauser LJ.** 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics **11:**119. http://dx.doi.org/10.1186/1471-2105-11-119.

7. **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.

8. **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.

9. **Markowitz VM, Chen IM, Palaniappan K, Chu K, Szeto E, Pillay M, Ratner A, Huang J, Woyke T, Huntemann M, Anderson I, Billis K, Varghese N, Mavromatis K, Pati A, Ivanova NN, Kyrpides NC.** 2014. IMG 4 version of the integrated microbial genomes comparative analysis system. Nucleic Acids Res **42:**D560–D567. http://dx.doi.org/10.1093/nar/gkt963.

10. **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.

11. **Andreesen JR.** 1994. Glycine metabolism in anaerobes. Ant Leeuwenhoek **66:**223–237. http://dx.doi.org/10.1007/BF00871641.

12. **Köpke M, Held C, Hujer S, Liesegang H, Wiezer A, Wollherr A, Ehrenreich A, Liebl W, Gottschalk G, Dürre P.** 2010. *Clostridium ljungdahlii* represents a microbial production platform based on syngas. Proc Natl Acad Sci U S A **107:**13087–13092. http://dx.doi.org/10.1073/pnas.1004716107.

13. **Brown SD, Nagaraju S, Utturkar S, De Tissera S, Segovia S, Mitchell W, Land ML, Dassanayake A, Köpke M.** 2014. Comparison of single-molecule sequencing and hybrid approaches for finishing the genome of *Clostridium autoethanogenum* and analysis of CRISPR systems in industrial relevant *Clostridia*. Biotechnol Biofuels **7:**40. http://dx.doi.org/10.1186/1754-6834-7-40.

14. **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.

15. **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.