

Complete Genome Sequence of the Type Strain of the Acetogenic Bacterium *Moorella thermoacetica* DSM 521^T

Anja Poehlein,^a Frank R. Bengelsdorf,^b Carola Esser,^b Bettina Schiel-Bengelsdorf,^b Rolf Daniel,^a Peter Dürre^b

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

A.P. and F.R.B. contributed equally to this work.

Here we report the closed genome sequence of the type strain *Moorella thermoacetica* DSM 521^T, an acetogenic bacterium, which is able to grow autotrophically on H₂ + CO₂ and/or CO, using the Wood-Ljungdahl pathway. The genome consists of a circular chromosome (2.53 Mb).

Received 21 August 2015 Accepted 25 August 2015 Published 8 October 2015

Citation Poehlein A, Bengelsdorf FR, Esser C, Schiel-Bengelsdorf B, Daniel R, Dürre P. 2015. Complete genome sequence of the type strain of the acetogenic bacterium *Moorella thermoacetica* DSM 521^T. *Genome Announc* 3(5):e01159-15. doi:10.1128/genomeA.01159-15.

Copyright © 2015 Poehlein et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 3.0 Unported license](#).

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

Homoacetogenic organisms are able to grow autotrophically on H₂ + CO₂ by using the Wood-Ljungdahl pathway. *Moorella thermoacetica* DSM 521^T, formerly *Clostridium thermoaceticum*, which was isolated from horse feces in 1942 (1), was used as a model acetogen for elucidating the Wood-Ljungdahl pathway. Mainly two strains, *M. thermoacetica* DSM 521^T and *M. thermoacetica* strain ATCC 39073, were used. For example, strain *M. thermoacetica* DSM 521^T was used to study the metabolism of carbohydrates (2), features of acetate kinase (3), the presence of cytochrome and menaquinone (4), and use of CO as electron donor (5). The biochemistry for reduction of H₂ + CO₂ and/or CO using the Wood-Ljungdahl pathway is described in detail in a number of reviews (6, 7). In some studies, cells or chromosomal DNA of *M. thermoacetica* DSM 521^T were used, but the genome sequence of the non-type strain *M. thermoacetica* ATCC 39073 (8) was cited (for example see references 9 and 10). Therefore, we decided to sequence and publish the genome of the type strain *M. thermoacetica* DSM 521^T.

A MasterPure complete DNA purification kit (Epicentre, Madison, WI, USA) was used to isolate chromosomal DNA of *M. thermoacetica* DSM 521^T. Isolated DNA was used to generate Illumina shotgun sequencing libraries. Sequencing was performed with a MiSeq system using a MiSeq reagent kit v3 (600 cycles) as recommended by the manufacturer (Illumina, San Diego, CA, USA). Sequencing resulted in 4,553,522 paired-end reads (2 × 300 bp) that were trimmed using Trimmomatic 0.32 (11). *De novo* assembly performed with the SPAdes genome assembler software version 3.5.0 (12) yielded in 40 contigs (>500 bp) and an average coverage of 382-fold. For scaffolding, we used the Move Contigs tool of the Mauve genome alignment software (13) and the genome of *M. thermoacetica* ATCC 39073 (CP000232) as reference. Remaining gaps were closed by PCR-based techniques, primer walking, and Sanger sequencing of products using BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technologies GmbH, Darmstadt, Germany). For

this purpose, the Gap4 (v4.11) software of the Staden package (14) was employed.

The genome of *M. thermoacetica* DSM 521^T consists of a circular 2.53-Mb chromosome with an overall G+C content of 55.95%. Automatic gene prediction was performed by using Prodigal software tool (15). Genes coding for rRNA and tRNA were identified using RNAmmer (16) and tRNAscan (17), respectively. An Integrated Microbial Genomes–Expert Review (IMG-ER) system (18) was used for automatic annotation, which was subsequently manually curated by using the Swiss-Prot, TrEMBL, and InterPro databases (19). The genome harbored 3 rRNA genes, 52 tRNA genes, and 2,553 protein-encoding genes. Genes encoding proteins of the carbonyl branch of the Wood-Ljungdahl pathway showed the same arrangement as in *Moorella thermoacetica* ATCC 39073. Comparison revealed that the genome of strain DSM 521^T (2,527,564 bp) is smaller than that of ATCC 39073 (2,628,784 bp) and that of the recently published *Moorella thermoacetica* DSM 2955^T (2,623,349 bp) (20). We detected 992 single nucleotide polymorphisms in the final genome compared to *M. thermoacetica* ATCC 39073.

Nucleotide sequence accession number. This complete genome project has been deposited at DDBJ/EMBL/GenBank under the accession number [CP012369](#).

ACKNOWLEDGMENTS

Work was supported by grants from the BMBF Gas Fermentation project (FKZ 031A468A), the ERA-IB 5 project CO2CHEM (FKZ 031A566A), and the MWK-BW project Nachhaltige und effiziente Biosynthesen (AZ 33-7533-6-195/7/9).

We thank Kathleen Gollnow and Frauke-Dorothee Meyer for technical support.

REFERENCES

1. Fontaine FE, Peterson WH, McCoy E, Johnson MJ, Ritter GJ. 1942. A new type of glucose fermentation by *Clostridium thermoaceticum*. *J Bacteriol* 43:701–715.
2. Andreesen JR, Schaupp A, Neurauter C, Brown A, Ljungdahl LG. 1973.

- Fermentation of glucose, fructose, and xylose by *Clostridium thermoaceticum*: effect of metals on growth yield, enzymes and the synthesis of acetate from CO₂. *J Bacteriol* 114:743–751.
3. Schaupp A, Ljungdahl LG. 1974. Purification and properties of acetate kinase from *Clostridium thermoaceticum*. *Arch Mikrobiol* 100:121–129. <http://dx.doi.org/10.1007/BF00446312>.
 4. Gottwald M, Andreesen JR, Le Gall J, Ljungdahl LG. 1975. Presence of cytochrome and menaquinone in *Clostridium formicoaceticum* and *Clostridium thermoaceticum*. *J Bacteriol* 122:325–328.
 5. Diekert GB, Thauer RK. 1978. Carbon monoxide oxidation by *Clostridium thermoaceticum* and *Clostridium formicoaceticum*. *J Bacteriol* 136:597–606.
 6. Ragsdale SW. 2004. Life with carbon monoxide. *Crit Rev Biochem Mol Biol* 39:165–195. <http://dx.doi.org/10.1080/10409230490496577>.
 7. Drake H, Küsel K, Matthies C. 2006. Acetogenic prokaryotes, p 354–420. *In* Dworkin M, Falkow S, Rosenberg E, Schleifer K, Stackebrandt E (ed), *Prokaryotes*, vol 2. Springer Verlag, New York, NY.
 8. Pierce E, Xie G, Barabote RD, Saunders E, Han CS, Detter JC, Richardson P, Brettin TS, Das A, Ljungdahl LG, Ragsdale SW. 2008. The complete genome sequence of *Moorella thermoacetica* (f. *Clostridium thermoaceticum*). *Environ Microbiol* 10:2550–2573. <http://dx.doi.org/10.1111/j.1462-2920.2008.01679.x>.
 9. Wang S, Huang H, Kahnt J, Thauer RK. 2013. A reversible electron-bifurcating ferredoxin- and NAD-dependent [FeFe]-hydrogenase (Hyd-ABC) in *Moorella thermoacetica*. *J Bacteriol* 195:1267–1275. <http://dx.doi.org/10.1128/JB.02158-12>.
 10. Mock J, Wang S, Huang H, Kahnt J, Thauer RK. 2014. Evidence for a hexaheteromeric methylenetetrahydrofolate reductase in *Moorella thermoacetica*. *J Bacteriol* 196:3303–3314. <http://dx.doi.org/10.1128/JB.01839-14>.
 11. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
 12. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski 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>.
 13. Darling AE, Mau B, Perna NT. 2015. Progressive mauve: multiple genome alignment with gene gain, loss, and rearrangement. *PLoS One* 5:e11147. <http://dx.doi.org/10.1371/journal.pone.0011147>.
 14. Staden R, Beal KF, Bonfield JK. 2000. The Staden package, 1998. *Methods Mol Biol* 132:115–130.
 15. 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>.
 16. 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>.
 17. 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>.
 18. 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. IMG4 version of the integrated microbial genomes comparative analysis system. *Nucleic Acids Res* 42:D560–D567. <http://dx.doi.org/10.1093/nar/gkt963>.
 19. 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>.
 20. Bengelsdorf FR, Poehlein A, Esser C, Schiel-Bengelsdorf B, Daniel R, Dürre P. Complete genome sequence of the acetogenic bacterium *Moorella thermoacetica* DSM 2955^T. *Genome Announc* 3(5):e01157-15. <http://dx.doi.org/10.1128/genomeA.01157-15>.