Permanent Draft Genome Sequence of *Desulfurococcus amylolyticus* Strain Z-533ᵀ, a Peptide and Starch Degrader Isolated from Thermal Springs in the Kamchatka Peninsula and Kunashir Island, Russia

Dwi Susanti,[a] Eric F. Johnson,[a] Alla Lapidus,[b] James Han,[c] T. B. K. Reddy,[c] Supratim Mukherjee,[c] Manoj Pillay,[d] Anna A. Perevalova,[e,f] Natalia N. Ivanova,[c] Tanja Woyke,[c] Nikos C. Kyrpides,[c,d,g] Biswarup Mukhopadhyay[a,e,h]

Department of Biochemistry, Virginia Tech, Blacksburg, Virginia, USA[a]; Centre for Algorithmic Biotechnology, ITBM, St. Petersburg State University, St. Petersburg, Russia[b]; U.S. Department of Energy Joint Genome Institute, Walnut Creek, California, USA[c]; Biological Data Management and Technology Center, Lawrence Berkeley National Laboratory, Berkeley, California, USA[d]; Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, Virginia, USA[e]; Winogradsky Institute of Microbiology, Research Center of Biotechnology Russian Academy of Sciences, Moscow, Russia[f]; Department of Biology, Faculty of Science, King Abdulaziz University, Jeddah, Saudi Arabia[g]; Department of Biological Sciences, Virginia Tech, Blacksburg, Virginia, USA[h]

ABSTRACT *Desulfurococcus amylolyticus* Z-533ᵀ, a hyperthermophilic crenarcheon, ferments peptide and starch, generating acetate, isobutyrate, isovalerate, CO_2, and hydrogen. Unlike *D. amylolyticus* Z-1312, it cannot use cellulose and is inhibited by hydrogen. The reported draft genome sequence of *D. amylolyticus* Z-533ᵀ will help to understand the molecular basis for these differences.

Desulfurococcus amylolyticus Z-533ᵀ (DSM 3822ᵀ), an inhabitant of thermal springs in the Kamchatka Peninsula and Kunashir Island, Russia, is a hyperthermophilic, anaerobic, sulfur-reducing crenarcheon (1). It is a nonmotile regular coccus of 0.7 to 1.5 μm in diameter (1). In laboratory cultures, *D. amylolyticus* Z-533ᵀ uses peptides such as casein hydrolysates, peptone, and yeast extract and simple carbohydrates that include starch, pectin, and glycogen as energy substrates. Sulfur stimulates growth, and, when it is employed, H_2S is produced. The absence of sulfur results in poor growth and H_2 production.

The 16S rRNA gene sequences of *D. amylolyticus* Z-533ᵀ differ by 0.1 to 0.3% from those of other *D. amylolyticus* strains, namely, Z-1312 and 1221n (2–4), which were formerly known as *D. fermentans* Z-1312 and *D. kamchatkensis* 1221n, respectively (5). Of these, only *D. amylolyticus* Z-1312 degrades cellulose (2) and is not inhibited by the presence of hydrogen (6). However, it lacks known cellulose genes and might employ novel mechanisms for cellulose degradation. Thus, a genomic analysis of three desulfurococci will give insight into the mechanisms by which *D. amylolyticus* Z-1312 degrades cellulose and by which other strains are inhibited by hydrogen.

The permanent draft genome of *D. amylolyticus* strain Z-533ᵀ (DSM 3282ᵀ) was generated at the DOE Joint Genome Institute (JGI) (7). The term "permanent" indicates that this genome sequence has been completed at a draft level and submitted to GenBank (8). The sequencing of a standard shotgun library on the Illumina platform generated 28,000,000 reads of 150 bp. These raw sequences were passed through DUK (9), and filtered reads were assembled with Velvet (10) and ALLPATHS-LG (11, 12). The final assembly contained one scaffold and two contigs. Structural and functional

Address correspondence to Dwi Susanti, dsusanti@vt.edu.

annotations were performed using the JGI's microbial genome annotation pipeline (13). The predicted coding sequences were translated and used to search the National Center for Biotechnology Information (NCBI) nonredundant, UniProt, TIGR-Fam, Pfam, PRIAM, KEGG, COG, and InterPro databases. Ribosomal and tRNA genes were identified with HMMER version 3.0rc1 (14) and tRNAscan-SE version 1.23 (15), respectively. Noncoding genes were predicted using Infernal version 1.0.2 (16). Additional annotation was performed within the Integrated Microbial Genomes—Expert Review platform (17). Clustered regularly interspaced short palindromic repeats (CRISPR) elements were detected using CRT (18) and PILER-CR (19).

The draft assembly of *D. amylolyticus* Z-533^T resulted in one scaffold and two contigs with a total sequence size of 1,309,099 bp and a 45.2% G+C content. It encoded 1,394 polypeptides, 59 RNAs (one 5S-, one 16S-, one 23S rRNA, 47 tRNAs, and nine other RNAs), and one putative CRISPR. The analysis predicted a peptide and starch degradation system for *D. amylolyticus* Z-533^T that is present in *D. amylolyticus* 1221n and Z-1312.

Accession number(s). This whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession number AZUU00000000. The version described in this paper is the first version, AZUU01000000.

REFERENCES

1. Bonch-Osmolovskaya EA, Slesarev A, Miroshnichenko ML, Svetlichnaya YP, Alekseev VA. 1988. Characteristics of *Desulfurococcus amylolyticus* n. sp.—a new extremely thermophilic archaebacterium isolated from thermal springs of Kamchatka and Kunashir island. Microbiology 57:78–85.

2. Perevalova AA, Svetlichny VA, Kublanov IV, Chernyh NA, Kostrikina NA, Tourova TP, Kuznetsov BB, Bonch-Osmolovskaya EA. 2005. *Desulfurococcus fermentans* sp. nov., a novel hyperthermophilic archaeon from a Kamchatka hot spring, and emended description of the genus *Desulfurococcus*. Int J Syst Evol Microbiol 55:995–999. https://doi.org/10.1099/ijs.0.63378-0.

3. Susanti D, Johnson EF, Rodriguez JR, Anderson I, Perevalova AA, Kyrpides N, Lucas S, Han J, Lapidus A, Cheng JF, Goodwin L, Pitluck S, Mavrommatis K, Peters L, Land ML, Hauser L, Gopalan V, Chan PP, Lowe TM, Atomi H, Bonch-Osmolovskaya EA, Woyke T, Mukhopadhyay B. 2012. Complete genome sequence of *Desulfurococcus fermentans*, a hyperthermophilic cellulolytic crenarchaeon isolated from a freshwater hot spring in Kamchatka, Russia. J Bacteriol 194:5703–5704. https://doi.org/10.1128/JB.01314-12.

4. Ravin NV, Mardanov AV, Beletsky AV, Kublanov IV, Kolganova TV, Lebedinsky AV, Chernyh NA, Bonch-Osmolovskaya EA, Skryabin KG. 2009. Complete genome sequence of the anaerobic, protein-degrading hyperthermophilic crenarchaeon *Desulfurococcus kamchatkensis*. J Bacteriol 191:2371–2379. https://doi.org/10.1128/JB.01525-08.

5. Perevalova AA, Kublanov IV, Bidzhieva SKh, Mukhopadhyay B, Bonch-Osmolovskaya EA, Lebedinsky AV. 2016. Reclassification of *Desulfurococcus mobilis* as a synonym of *Desulfurococcus mucosus*, *Desulfurococcus fermentans* and *Desulfurococcus kamchatkensis* as synonyms of *Desulfurococcus amylolyticus*, and emendation of the *D. mucosus* and *D. amylolyticus* species descriptions. Int J Syst Evol Microbiol 66:514–517. https://doi.org/10.1099/ijsem.0.000747.

6. Slobodin AI, Bonch-Osmolovskaya EA. 1994. Growth and formation of metabolic products by extremely thermophilic archaea of the genus *Desulfurococcus* in the presence and absence of elemental sulfur. Microbiology 63:552–554.

7. Chen IA, Markowitz VM, Chu K, Palaniappan K, Szeto E, Pillay M, Ratner A, Huang J, Andersen E, Huntemann M, Varghese N, Hadjithomas M, Tennessen K, Nielsen T, Ivanova NN, Kyrpides NC. 2017. IMG/M: integrated genome and metagenome comparative data analysis system. Nucleic Acids Res 45:D507–D516. https://doi.org/10.1093/nar/gkw929.

8. Pagani I, Liolios K, Jansson J, Chen IMA, Smirnova T, Nosrat B, Markowitz VM, Kyrpides NC. 2012. The genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. Nucleic Acids Res 40:D571–D579. https://doi.org/10.1093/nar/gkr1100.

9. Mingkun L, Han J. 2011. DUK—a fast and efficient kmer matching tool. U.S. Department of Energy Joint Genome Institute, Walnut Creek, CA. https://publications.lbl.gov/islandora/object/ir%3A155199/datastream/PDF/view.

10. Zerbino DR, Birney E. 2008. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Res 18:821–829. https://doi.org/10.1101/gr.074492.107.

11. Butler J, MacCallum I, Kleber M, Shlyakhter IA, Belmonte MK, Lander ES, Nusbaum C, Jaffe DB. 2008 ALLPATHS: de novo assembly of whole-genome shotgun microreads. Genome Res 18:810–820. https://doi.org/10.1101/gr.7337908.

12. MacCallum I, Przybylski D, Gnerre S, Burton J, Shlyakhter I, Gnirke A, Malek J, McKernan K, Ranade S, Shea TP, Williams L, Young S, Nusbaum C, Jaffe DB. 2009. ALLPATHS 2: small genomes assembled accurately and with high continuity from short paired reads. Genome Biol 10:R103. https://doi.org/10.1186/gb-2009-10-10-r103.

13. Huntemann M, Ivanova NN, Mavromatis K, Tripp HJ, Paez-Espino D, Palaniappan K, Szeto E, Pillay M, Chen IMA, Pati A, Nielsen T, Markowitz VM, Kyrpides NC. 2015. The standard operating procedure of the DOE-JGI microbial genome annotation pipeline (MGAP v.4). Stand Genomic Sci 10:86. https://doi.org/10.1186/s40793-015-0077-y.

14. Finn RD, Clements J, Eddy SR. 2011. HMMER web server: interactive sequence similarity searching. Nucleic Acids Res 39:W29–W37. https://doi.org/10.1093/nar/gkr367.

15. 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.

16. Nawrocki EP, Kolbe DL, Eddy SR. 2009. Infernal 1.0: inference of RNA alignments. Bioinformatics 25:1335–1337. https://doi.org/10.1093/bioinformatics/btp157.

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

18. Bland C, Ramsey TL, Sabree F, Lowe M, Brown K, Kyrpides NC, Hugenholtz P. 2007. CRISPR Recognition Tool (CRT): a tool for automatic detection of clustered regularly interspaced palindromic repeats. BMC Bioinformatics 8:209. https://doi.org/10.1186/1471-2105-8-209.

19. Edgar RC. 2007. PILER-CR: fast and accurate identification of CRISPR repeats. BMC Bioinformatics 8:18. https://doi.org/10.1186/1471-2105-8-18.