JOURNAL OF BACTERIOLOGY, Aug. 2011, p. 4294–4295 Vol. 193, No. 16
0021-9193/11/$12.00 doi:10.1128/JB.05302-11

Genome Sequence of the Curdlan-Producing *Agrobacterium* sp. Strain ATCC 31749[▽]

Anne M. Ruffing,[1] Marlene Castro-Melchor,[2] Wei-Shou Hu,[2] and Rachel R. Chen[1]*

School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, 311 Ferst Drive, Atlanta, Georgia 30332-0100,[1] *and University of Minnesota, Department of Chemical Engineering and Materials Science, 151 Amundson Hall, 421 Washington Avenue SE, Minneapolis, Minnesota 55455-0132*[2]

Received 14 May 2011/Accepted 6 June 2011

***Agrobacterium* sp. ATCC 31749 is an industrial strain for the commercial production of curdlan, an important exopolysaccharide with food and medical applications. Here we report the genome sequence of the curdlan-producing strain ATCC 31749. Genome sequencing is the first step toward the understanding of regulation of curdlan biosynthesis.**

Agrobacterium sp. ATCC 31749 is an alphaproteobacterium of the *Rhizobiaceae* family (6). But unlike other species of the rhizobium family, ATCC 31749 is not plant associated. This agrobacterium is unique in its ability to synthesize, upon nitrogen exhaustion, a linear β-1,3-glucan exopolysaccharide (EPS) known as curdlan. Curdlan has been commercialized for applications in the food, construction, and pharmaceutical industries, and numerous other applications, such as antiviral and anticancer treatments, are being pursued (3, 5, 14). While extensive studies have led to optimized conditions for curdlan synthesis (8–12), very little is known about the genetics outside the curdlan biosynthesis operon (*crdASC*) (16), and as with many other microbial EPSs, its regulation is largely unexplored.

Agrobacterium sp. ATCC 31749 was sequenced using the Genome Sequencer FLX system from 454 Life Sciences. A total of 92,994,272 bp were sequenced in 399,219 reads, with an average read length of 233 bp. Using the GS de novo assembler, the reads were assembled into 95 contigs of 500 bp or longer, yielding an approximate genome size of 5.5 Mbp with 17× coverage and 59% GC content. The contigs were arranged in random order with stop and start codons inserted in each reading frame between the contigs to generate a pseudochromosome. From this pseudochromosome, 5,585 genes were predicted using the GeneMarkS (1) and Glimmer v3.02 (2) software programs. Annotation with BLAST identified 3,466 genes with an assigned function and 1,635 conserved hypothetical proteins, leaving the remaining 484 predicted genes as hypothetical proteins. Analysis with the tRNAscan-SE 1.21 server (13) identified 44 tRNAs.

Besides the curdlan biosynthesis operon (*crdASC*), the genome sequence of ATCC 31749 identifies many potential regulatory genes associated with environmental conditions previously shown to influence curdlan biosynthesis, including nitrogen, oxygen, phosphate, and pH (8–11). While no evidence of a nitrogenase was found, the ATCC 31749 genome includes the sigma factor associated with nitrogen-limited metabolism (*rpoN*) and genes involved in the nitrogen signaling cascade (*ntrBC*, *ntrXY*, and *glnBDK*) and nitrogen fixation (*fixK*, *nifR*, and *ptsN*). Genes encoding the oxygen-responsive global regulators *fnrN* and *nolR* and the genes involved in phosphate (*phoB*) and pH (*chvG* and *chvI*) regulation are also found. Additionally, the ATCC 31749 genome contains 31 predicted genes with GGDEF/EAL domains, known to regulate the production of many bacterial polysaccharides (7, 18). Like that of *Agrobacterium tumefaciens* (15), the ATCC 31749 genome contains genes associated with the production of acidocalcisomes, organelles that contain large quantities of polyphosphate (polyP) (4).

The genome sequence of *Agrobacterium* sp. ATCC 31749 provides the basis for subsequent experimentation to decipher the regulation mechanisms. Synthesis of EPSs is common among many bacterial species, some of which are involved in biofilms and have important medical implications (17). While chemically diverse, their syntheses share common environmental triggers, suggesting that EPS regulatory elements may be more universal than their chemical structures suggest. Continued study of curdlan synthesis, aided by the genome sequence, may uncover these common mechanisms.

Nucleotide sequence accession numbers. This whole-genome shotgun project has been deposited at DDBJ/EMBL/GenBank under the accession no. AECL00000000. The version described here is the first version, AECL01000000.

This work was funded by the National Science Foundation (BES 0455193) and the American Cancer Society. A. M. Ruffing acknowledges support from an NSF Graduate Research Fellowship.

REFERENCES

1. **Besemer, J., A. Lomsadze, and M. Borodovsky.** 2001. GeneMarkS: a self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res. **29:**2607–2618.
2. **Delcher, A. L., K. A. Bratke, E. C. Powers, and S. L. Salzberg.** 2007. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics **23:**673–679.
3. **Demleitner, S., J. Kraus, and G. Franz.** 1992. Synthesis and antitumour activity of sulfoalkyl derivatives of curdlan and lichenan. Carbohydr. Res. **226:**247–252.
4. **Docampo, R., W. de Souza, K. Miranda, P. Rohloff, and S. N. J. Moreno.** 2005. Acidocalcisomes—conserved from bacteria to man. Nat. Rev. Microbiol. **3:**251–261.

* Corresponding author. Mailing address: School of Chemistry and Biomolecular Engineering, Georgia Institute of Technology, 311 Ferst Dr., Atlanta, GA 30332-0100. Phone: (404) 894-1255. Fax: (404) 894-2866. E-mail: rchen@chbe.gatech.edu.

▽ Published ahead of print on 17 June 2011.

5. **Gordon, M., et al.** 1994. A phase I study of curdlan sulfate—an HIV inhibitor. Tolerance, pharmacokinetics and effects on coagulation and on CD4 lymphocytes. J. Med. **25:**163–180.
6. **Harada, T., M. Masada, K. Fujimori, and I. Maeda.** 1966. Production of firm, resilient gel-forming polysaccharide by a mutant of *Alcaligenes faecalis* var. *myxogenes* 10C3. Agric. Biol. Chem. **30:**196–198.
7. **Hickman, J. W., and C. S. Harwood.** 2008. Identification of FleQ from *Pseudomonas aeruginosa* as a c-di-GMP-responsive transcription factor. Mol. Microbiol. **69:**376–389.
8. **Kim, M. K., I. Y. Lee, J. H. Ko, Y. H. Rhee, and Y. H. Park.** 1999. Higher intracellular levels of uridinemonophosphate under nitrogen-limited conditions enhance metabolic flux of curdlan synthesis in *Agrobacterium* species. Biotechnol. Bioeng. **62:**317–323.
9. **Kim, M. K., et al.** 2000. Residual phosphate concentration under nitrogen-limiting conditions regulates curdlan production in *Agrobacterium* species. J. Ind. Microbiol. Biotechnol. **25:**180–183.
10. **Lee, I. Y., et al.** 1999. Influence of agitation speed on production of curdlan by *Agrobacterium* species. Bioprocess. Eng. **20:**283–287.
11. **Lee, J.-H., I.-Y. Lee, M.-K. Kim, and Y.-H. Park.** 1999. Optimal pH control of batch processes for production of curdlan by *Agrobacterium* species. J. Ind. Microbiol. Biotechnol. **23:**143–148.
12. **Lee, J. H., and I. Y. Lee.** 2001. Optimization of uracil addition for curdlan (β-1→3-glucan) production by *Agrobacterium* sp. Biotechnol. Lett. **23:**1131–1134.
13. **Lowe, T. M., and S. R. Eddy.** 1997. tRNAscan-SE: a program for improved detection of tRNA genes in genomic sequence. Nucleic Acids Res. **25:**955–964.
14. **McIntosh, M., B. A. Stone, and V. A. Stanisich.** 2005. Curdlan and other bacterial (1→3)-β-D-glucans. Appl. Microbiol. Biotechnol. **68:**163–173.
15. **Seufferheld, M., et al.** 2003. Identification of organelles in bacteria similar to acidocalcisomes of unicellular eukaryotes. J. Biol. Chem. **278:**29971–29978.
16. **Stasinopoulos, S. J., P. R. Fisher, B. A. Stone, and V. A. Stanisich.** 1999. Detection of two loci involved in (1→3)-β-glucan (curdlan) biosynthesis by *Agrobacterium* sp. ATCC31749, and comparative sequence analysis of the putative curdlan synthase gene. Glycobiology **9:**31–41.
17. **Vuong, C., et al.** 2004. A crucial role for exopolysaccharide modification in bacterial biofilm formation, immune evasion, and virulence. J. Biol. Chem. **279:**54881–54886.
18. **Weinhouse, H., et al.** 1997. c-di-GMP-binding protein, a new factor regulating cellulose synthesis in *Acetobacter xylinum*. FEBS Lett. **416:**207–211.