

# UC Davis

## UC Davis Previously Published Works

### Title

Biochemical characterization of Fsa16295Glu from “Fervidibacter sacchari,” the first hyperthermophilic GH50 with  $\beta$ -1,3-endoglucanase activity and founding member of the subfamily GH50\_3

### Permalink

<https://escholarship.org/uc/item/5zh738z7>

### Authors

Covington, Jonathan K

Torosian, Nicole

Cook, Allison M

et al.

### Publication Date

2024

### DOI

10.3389/fmicb.2024.1355444

### Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed



## OPEN ACCESS

## EDITED BY

En-Min Zhou,  
Yunnan University, China

## REVIEWED BY

Bhagwan Narayan Rekadwad,  
Yenepoya University, India  
Sara E. Blumer-Schuetz,  
Oakland University, United States

## \*CORRESPONDENCE

Brian P. Hedlund  
✉ brian.hedlund@unlv.edu

RECEIVED 13 December 2023

ACCEPTED 08 April 2024

PUBLISHED 25 April 2024

## CITATION

Covington JK, Torosian N, Cook AM, Palmer M, Bryan SG, Nou NO, Mewalal R, Harmon-Smith M, Blaby IK, Cheng J-F, Hess M, Brumm PJ, Singh NK, Venkateswaran K and Hedlund BP (2024) Biochemical characterization of Fsa16295Glu from *Fervidibacter sacchari*, the first hyperthermophilic GH50 with  $\beta$ -1,3-endoglucanase activity and founding member of the subfamily GH50\_3. *Front. Microbiol.* 15:1355444. doi: 10.3389/fmicb.2024.1355444

## COPYRIGHT

© 2024 Covington, Torosian, Cook, Palmer, Bryan, Nou, Mewalal, Harmon-Smith, Blaby, Cheng, Hess, Brumm, Singh, Venkateswaran and Hedlund. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Biochemical characterization of Fsa16295Glu from *Fervidibacter sacchari*, the first hyperthermophilic GH50 with $\beta$ -1,3-endoglucanase activity and founding member of the subfamily GH50\_3

Jonathan K. Covington<sup>1</sup>, Nicole Torosian<sup>1</sup>, Allison M. Cook<sup>1</sup>, Marike Palmer<sup>1,2</sup>, Scott G. Bryan<sup>1</sup>, Nancy O. Nou<sup>1</sup>, Ritesh Mewalal<sup>3</sup>, Miranda Harmon-Smith<sup>3</sup>, Ian K. Blaby<sup>3</sup>, Jan-Fang Cheng<sup>3</sup>, Matthias Hess<sup>4</sup>, Phillip J. Brumm<sup>5</sup>, Nitin K. Singh<sup>6</sup>, Kasthuri Venkateswaran<sup>6</sup> and Brian P. Hedlund<sup>1,7\*</sup>

<sup>1</sup>School of Life Sciences, University of Nevada, Las Vegas, NV, United States, <sup>2</sup>Department of Microbiology, University of Manitoba, Winnipeg, MB, Canada, <sup>3</sup>US Department of Energy Joint Genome Institute, Lawrence Berkeley National Laboratory, Berkeley, CA, United States, <sup>4</sup>Department of Animal Science, College of Agricultural and Environmental Sciences, University of California, Davis, CA, United States, <sup>5</sup>C5-6 Technologies LLC, Fitchburg, WI, United States, <sup>6</sup>Biotechnology and Planetary Protection Group, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, United States, <sup>7</sup>Nevada Institute of Personalized Medicine, University of Nevada, Las Vegas, NV, United States

The aerobic hyperthermophile *Fervidibacter sacchari* catabolizes diverse polysaccharides and is the only cultivated member of the class *Fervidibacteria* within the phylum *Armatimonadota*. It encodes 117 putative glycoside hydrolases (GHs), including two from GH family 50 (GH50). In this study, we expressed, purified, and functionally characterized one of these GH50 enzymes, Fsa16295Glu. We show that Fsa16295Glu is a  $\beta$ -1,3-endoglucanase with optimal activity on carboxymethyl curdlan (CM-curdlan) and only weak agarase activity, despite most GH50 enzymes being described as  $\beta$ -agarases. The purified enzyme has a wide temperature range of 4–95°C (optimal 80°C), making it the first characterized hyperthermophilic representative of GH50. The enzyme is also active at a broad pH range of at least 5.5–11 (optimal 6.5–10). Fsa16295Glu possesses a relatively high  $k_{cat}/K_M$  of  $1.82 \times 10^7 \text{ s}^{-1} \text{ M}^{-1}$  with CM-curdlan and degrades CM-curdlan nearly completely to sugar monomers, indicating preferential hydrolysis of glucans containing  $\beta$ -1,3 linkages. Finally, a phylogenetic analysis of Fsa16295Glu and all other GH50 enzymes revealed that Fsa16295Glu is distant from other characterized enzymes but phylogenetically related to enzymes from thermophilic archaea that were likely acquired horizontally from *Fervidibacteria*. Given its functional and phylogenetic novelty, we propose that Fsa16295Glu represents a new enzyme subfamily, GH50\_3.

## KEYWORDS

hyperthermophile, carboxymethyl curdlan,  $\beta$ -Glucan,  $\beta$ -1,3-endoglucanase, glycoside hydrolase, GH50, CAZyme, *Armatimonadota*

## 1 Introduction

Thermoactive enzymes are an essential component of hyperthermophiles, which thrive in geothermal environments and have optimal growth at temperatures above 80°C (Adams and Kelly, 2001; Hedlund et al., 2016). These environments, which include marine hydrothermal vents, geysers, fumaroles, and hot springs, host bacteria and archaea that use polysaccharides as carbon and energy sources (Adams and Kelly, 2001; Blumer-Schuette et al., 2008; Strazzulli et al., 2017). A variety of hyperthermophilic bacteria, such as *Thermotoga maritima* and *Caldicellulosiruptor bescii*, can ferment diverse polysaccharides, including lignocellulose, hemicellulose, and glucans, such as  $\beta$ -glucan (Bhalla et al., 2013; Chen et al., 2013; Meng et al., 2016; Crosby et al., 2022).

$\beta$ -glucans are polysaccharides found in the cell walls of cereal grains and some bacteria and yeast and are composed of glucose residues connected by glycosidic linkages (Wang and Fisher, 2015). The structure and aqueous solubility of  $\beta$ -glucan depend on its origin; for example,  $\beta$ -glucans from plants are branched, relatively soluble, and have alternating  $\beta$ -1,3/1,4 linkages, while those from bacteria are unbranched, insoluble, and have only  $\beta$ -1,3 linkages (Mudgil, 2017; Paudel et al., 2021). Although the specific sources of  $\beta$ -glucans in hot springs have not been investigated, we can better understand how certain hyperthermophilic organisms use  $\beta$ -glucans by studying the enzymes responsible for depolymerizing them.

Glycoside hydrolases (GHs) are a class of enzymes that hydrolyze the glycosidic linkages between sugar residues in carbohydrates, breaking them down into oligosaccharides or even their base monomers (Adams and Kelly, 2001). GHs that degrade  $\beta$ -glucans do so by hydrolyzing either their  $\beta$ -1,3- or  $\beta$ -1,4-glycosidic bonds. These linkages can be cleaved linearly by either exo- or endo-acting enzymes (Santos et al., 2020). The activity of a GH can be predicted to some degree from its amino acid sequence through the presence of conserved active sites and catalytic residues, or by phylogenetic analysis. GHs have been categorized into GH families based on phylogenetic relatedness. These GH families often share one or more identifying features, like conserved catalytic residues, mechanisms, or similar substrates (Henrissat and Bairoch, 1996; Summers et al., 2016). GHs can be classified more broadly into GH clans, comprising GH families that share structural characteristics, and more finely into subfamilies (Henrissat and Bairoch, 1996; Viborg et al., 2019).

One example of a catalytically homogenous GH family is GH family 50 (GH50), of which nearly all known members to date have been classified as  $\beta$ -agarases. These enzymes hydrolyze  $\beta$ -1,4-glycosidic linkages found in agarose, a component of agar alongside agaropectin, to release its degradation products: neoagarohexaose, neoagarotetraose, and neoagarobiose. There are two exceptions to this trend: a  $\beta$ -galactosidase from *Victivallis vadensis* and a  $\beta$ -1,3-glucanase from *Pseudomonas aeruginosa* (Temuujin et al., 2012a; Yi et al., 2018). The earliest studied GH50 was AgaA from *Vibrio* sp. JT0107, which was described as the first endo- $\beta$ -agarase capable of hydrolyzing both agarose and neoagarotetraose (Sugano et al., 1993). To date, four structures are known from GH50, one forming monomers, two forming homodimers, and one forming homotetramers (Pluvinage et al., 2013; Giles et al., 2017; Zhang et al., 2019; Pluvinage et al., 2020). These structures have C-termini with  $(\beta/\alpha)_8$ -barrel folds, which are characteristic of the family's clan, GH clan A (GH-A). Although the

catalytic residues of GH50 have not been identified, its designation as a member of GH-A infers the use of two conserved glutamic acid residues as the acid/base catalyst and nucleophile (Kumar et al., 2019). Current research surrounding GH50s has focused on their potential for the production of neoagarobiose, which is used as a skin moisturizer and melanoma whitener (Lee et al., 2008). However, GH50 is an understudied family. Of the 1,631 GH50s that are currently deposited in the carbohydrate-active enzyme (CAZy) database (cazy.org), only 30 have been functionally characterized (Drula et al., 2022). The  $\beta$ -galactosidase (VadG925) and  $\beta$ -1,3-glucanase (PaBglu50A) set a precedent that GH50 may be more catalytically diverse than currently known. The study of GH50s from novel or understudied microbes broadens our understanding of the true diversity of this GH family.

We recently isolated "*Fervidibacter sacchari*," an aerobic hyperthermophile ( $T_{opt}$  80°C) belonging to the phylum *Armatimonadota* (Nou et al., 2024). "*F. sacchari*" uses at least 16 diverse polysaccharides as sole carbon sources and electron donors, including glucans, hemicelluloses, and plant biomass. The genome of "*F. sacchari*" has 117 genes encoding GH domains across 49 GH families, including two that were annotated as members of GH50, designated Fsa16295Glu and Fsa11540. Since the optimal growth temperature of "*F. sacchari*" exceeds the optimal temperature of the most thermophilic GH50 enzyme characterized to date, AgrA from *Agarivorans* sp. AG17 ( $T_{opt}$  65°C) (Nikapitiya et al., 2010), we probed the temperature optimum and thermostability of Fsa16295Glu and compared its properties to other GH50 enzymes. To do this, we codon-optimized the gene encoding Fsa16295Glu for expression in *Escherichia coli*, synthesized it, ligated it into the pET21b-GB1 vector, and expressed it in *E. coli*. The enzyme was then purified and used to determine its substrates, optimal conditions, kinetic parameters, and biochemical characteristics. Finally, we examined the evolutionary history of Fsa16295Glu within the GH50 family by conducting a comprehensive phylogenetic analysis of GH50 enzymes. We show that Fsa16295Glu is indeed hyperthermophilic ( $T_{opt}$  80°C) and is a  $\beta$ -1,3-endoglucanase with only weak activity on agarose. Because of this unusual activity and its unique phylogenetic position, we propose that Fsa16295Glu represents a novel GH50 subfamily herein designated as GH50\_3.

## 2 Materials and methods

### 2.1 Strains, gene synthesis, heterologous expression, enzyme purification, and computed structure models

"*F. sacchari*" inhabits ~80°C sediments in Great Boiling Spring, NV, USA, (Costa et al., 2009; Cole et al., 2013) and has been enriched *in situ* following lignocellulose addition (Peacock et al., 2013). A single strain of "*F. sacchari*," strain PD1, was isolated following *in situ* and laboratory enrichment by using optical tweezers followed by aerobic growth at 80°C with polysaccharides as the sole carbon source and electron donor, and its whole genome was assembled (Nou et al., 2024) (GCA\_030520105.1). Since the name *Fervidibacter sacchari* has not yet been validly published, we use the name in quotations in this manuscript. The amino acid sequence of Fsa16295Glu was downloaded and submitted to the dbCAN2 meta

server to determine the GH family of Fsa16295Glu (Zhang et al., 2018). The amino acid sequence was then submitted for annotation using SignalP (Teufel et al., 2022) to determine the presence of a signal sequence. A MAFFT-DASH (Katoh et al., 2019) multiple sequence alignment of seven GH50 amino acid sequences (Fsa16295Glu, Fsa11540, Aga50D, AgaA, AgrA, PaBglu50A, and VadG925) was generated and visualized in Jalview v2.11.2.6 (Waterhouse et al., 2009) to confirm the presence of two conserved glutamic acid residues. The gene encoding Fsa16295Glu was codon-optimized using an *E. coli* codon frequency table (Oberortner et al., 2017). This refactored sequence was synthesized (Twist BioScience, CA) inclusive of 30mer flanking linkers enabling assembly into the BamHI site of the pET21b vector with NEBuilder Hi-Fi Assembly (Novagen, Madison, WI; New England Biolabs, Ipswich, MA). The final construct, which contained a 6× His tag, a GB1 solubility tag, and a TEV protease cleavage site at the protein N-terminus, was sequence verified on the Pacific Biosciences Sequel IIe platform and transformed into *E. coli* T7 Express cells (New England Biolabs, Ipswich, MA) for expression.

*E. coli* cells containing the gene encoding Fsa16295Glu were grown on Luria-Bertani (LB) agar plates (10 g tryptone, 5 g yeast extract, 5 g NaCl per L) containing 100 µg/mL ampicillin. For expression, LB broth with 100 µg/mL ampicillin was inoculated and incubated at 37°C with shaking at 225 rpm until reaching an OD<sub>600</sub> of 0.6–0.8. Isopropyl-β-thiogalactoside (IPTG) was added to a concentration of 0.5 mM and the culture was incubated overnight at 37°C with shaking at 150 rpm. The cells were centrifuged at 16,100 × g, resuspended in lysis buffer (50 mM Tris-HCl, 100 mM NaCl, pH 7.0), and lysed by sonication for 40 s with 40% cycle and level 4 output using a Branson 450 Sonifier (Branson Ultrasonics, Brookfield, CT, United States). The crude lysate was clarified by centrifugation at 16,100 × g to isolate the soluble fraction.

Fsa16295Glu was purified by heating the soluble fraction to 80°C in a heat block for 30 min, followed by centrifugation at 16,100 × g to remove denatured host proteins in the pellet. Fsa16295Glu expression and purity were assessed using 7.5% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE). The molecular weight was estimated using a BriteRuler Pre-stained Protein Ladder (Abcam, Cambridge, United Kingdom). The ability of Fsa16295Glu to oligomerize was assessed using 7.5% native PAGE alongside a PageRuler™ Plus Pre-stained Protein Ladder (Thermo Scientific, Vilnius, Lithuania). Protein concentration was assessed by a Qubit® Protein and Protein Broad Range Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) using a Qubit® 3.0 Fluorometer (Thermo Fisher Scientific, Waltham, MA, USA).

Computed structure models (CSMs) of Fsa16295Glu were generated using AlphaFold 2 via the ColabFold platform integrated with ChimeraX v 1.5 using default settings (Pettersen et al., 2021; Mirdita et al., 2022). A total of five Fsa16295Glu CSMs were generated: (i) native and mature Fsa16295Glu (signal sequence removed); (ii) non-native Fsa16295Glu containing an N-terminal 6× His tag, GB1 solubility tag, and a TEV protease site; (iii) homodimeric native, and mature Fsa16295Glu; (iv) homotrimeric native, and mature Fsa16295Glu; and (v) homotetrameric native, and mature Fsa16295Glu. The resultant structures were visualized and edited in ChimeraX v 1.5, and the predicted aligned error (PAE) plots of the multimeric CSMs were analyzed to gain insight into the potential of Fsa16295Glu to homooligomerize.

## 2.2 Enzyme activity assays

To determine the enzyme activity of Fsa16295Glu, the following potential substrates were tested initially: chondroitin sulfate (Alfa Aesar), colloidal chitin as prepared according to Hsu and Lockwood (1975) (Beantown Chemicals), brown algae fucoïdan (BestVite), lupin galactan (Megazyme), gellan gum (Serva), karaya gum (Sigma), locust bean gum (Spectrum), xanthan gum (Sigma), birch wood xylan (Sigma), tamarind xyloglucan (CarboMer), oat β-glucan (Megazyme), oyster glycogen (TCI), potato starch (J.T. Baker), and ammonia fiber expansion (AFEX)-pretreated corn stover, *Miscanthus*, and sugarcane bagasse (DuPont). These 16 substrates were chosen based on their use by “*F. sacchari*” as sole carbon sources and electron donors. Agarose (VWR Chemicals) was included in the screen to test for β-agarase activity that would be consistent with GH50; carboxymethyl curdlan (CM-curdlan) (Megazyme) was included to test for β-1,3-glucanase activity; laminarin from *Eisenia bicyclis* (TCI America) was included to test for β-1,3/1,6-glucanase activity; and carboxymethyl cellulose (CMC) (Spectrum) was included to test for β-1,4-glucanase activity. The substrates (0.5% w/v) and enzyme (141 µg/mL) or empty vector control were mixed 1:1 in microplate wells (final volume 40 µL) and then sealed and incubated overnight at 80°C with a beaker of water to prevent evaporation. To measure the quantity (µmol) of reducing sugars released by the reactions, 3,5-dinitrosalicylic acid (DNS) solution (0.25 g DNS, 75 g sodium potassium tartrate, 50 mL 2 M sodium hydroxide solution, brought to 250 mL using ultrapure water) was added 4:1 to each reaction mixture (final volume 200 µL). The plate was wrapped in foil and incubated at 100°C for 20 min (Kim et al., 2010). Absorbance was measured using a SpectraMax Plus Plate Reader at 570 nm. A second screen with robust quantitation was conducted with Fsa16295Glu (141 µg/mL) in triplicate and compared with the empty vector control using an unpaired t-test. Additionally, a parallel screen was conducted using Fsa16295Glu pretreated with TEV protease (New England Biolabs, Ipswich, MA) on chondroitin sulfate (Alfa Aesar), colloidal chitin (Beantown Chemicals), lupin galactan (Megazyme), gellan gum (Serva), karaya gum (Sigma), locust bean gum (Spectrum), xanthan gum (Sigma), birch wood xylan (Sigma), beech wood xylan (Megazyme), tamarind xyloglucan (CarboMer), oat β-glucan (Megazyme), and potato starch (J.T. Baker). A DNS standard curve was generated using glucose at 0, 5, 10, 15, 20, and 25 mM.

To determine how much Fsa16295Glu depolymerized its substrates, acid hydrolysates of CM-curdlan, oat β-glucan, laminarin, and agarose were prepared by heating 1% solutions of each substrate in 2 M sulfuric acid at 121°C for 60 min followed by neutralization with 5 M sodium hydroxide. The acid hydrolysates were incubated with water instead of Fsa16295Glu, and the amount of reducing sugars released from exhaustive overnight degradation by Fsa16295Glu was compared with the reducing sugar content of the acid hydrolysates normalized to enzyme assay concentrations.

The β-glucosidase activity of Fsa16295Glu was assessed using the *para*-nitrophenyl (pNP)-linked substrate pNP-β-D-glucopyranoside (pNPG) (Megazyme). pNPG (20 µg/mL) was mixed with Fsa16295Glu (141 µg/mL) and incubated for 30 min at 80°C, and the reaction was terminated with 2% w/v disodium phosphate solution (pH 12.0). Bond cleavage was assessed as an increase in absorbance at 400 nm, indicating substrate hydrolysis. A standard curve was generated using a range of *para*-nitrophenol concentrations (0, 6, 12, 18, 24, and 30 µg/

mL). In addition to laminarin,  $\beta$ -1,3/1,6-glucan cleavage by Fsa16295Glu was assessed using the Enzymatic Yeast Beta-Glucan test kit (Megazyme) with active dry yeast (Ward's Science) by substituting Fsa16295Glu for Glucamix.

## 2.3 Biochemical characterization and kinetic parameters

To determine the temperature range and optimum of Fsa16295Glu, the enzyme was incubated with 0.5% w/v oat  $\beta$ -glucan at a range of temperatures (4, 20, 30, 50, 60, 70, 80, 90, and 95°C). The resulting solution was analyzed as described before using a DNS solution. To determine the thermostability of the enzyme, Fsa16295Glu was incubated for 1 h at either 70, 80, 90, or 100°C, then incubated at 80°C with oat  $\beta$ -glucan and tested using the DNS assay. To determine the pH range and optimum of Fsa16295Glu, the enzyme was expressed and purified in 50 mM buffers at different pHs in 0.5 intervals: 2-morpholin-4-ylethanesulfonic acid (MES) (5.5–6.5), 2-amino-2-(hydroxymethyl)propane-1,3-diol (Tris) (7.0–9.0), 2-(cyclohexylamino)ethane-1-sulfonic acid (CHES) (9.5–10.0), and 3-(cyclohexylamino)propane-1-sulfonic acid (CAPS) (10.5–11.0), all with 100 mM NaCl. Oat  $\beta$ -glucan was dissolved in the same buffers to capture the whole pH range. Assays were conducted with 0.5% w/v oat  $\beta$ -glucan in triplicate and then analyzed as previously described using the DNS method. The temperature and pH ranges and optima and thermostability were determined using one-way analysis of variance (ANOVA) with *post-hoc* Tukey's honestly significant difference (HSD) tests ( $p \leq 0.05$ ).

To determine the Michaelis–Menten kinetic parameters  $k_{cat}$  and  $k_{cat}/K_M$  of Fsa16295Glu, CM-curdlan or oat  $\beta$ -glucan was mixed with purified Fsa16295Glu (141  $\mu\text{g}/\text{mL}$ ) at a range of concentrations (156, 313, 625, 1,250, 2,500, and 5,000  $\mu\text{g}/\text{mL}$ ) and incubated for 30 min. Reducing sugars released by the reactions were quantified using the DNS assay as before. Lineweaver–Burk plots were generated from these data to calculate  $V_{max}$ ,  $K_M$ ,  $k_{cat}$ , and  $k_{cat}/K_M$  (Ochs, 2022).

## 2.4 Phylogenetic analysis of GH50

To perform a phylogenetic analysis of all GH50s, all sequences belonging to GH50 in the CAZy database (Drula et al., 2022) were extracted from the National Center for Biotechnology Information (NCBI; <https://www.ncbi.nlm.nih.gov>) using the Batch Entrez site (<https://www.ncbi.nlm.nih.gov/sites/batchentrez>). BioEdit v. 7.2.5 (Hall, 1999) was used for visualizing and editing all data matrices. After the removal of all duplicate sequences, protein sequences of less than 50 amino acids were also removed. Fsa16295Glu along with all other GH50 homologs from the class *Fervidibacteria* (133 in total) was added to the dataset, resulting in 1,638 protein sequences. A multiple sequence alignment was generated using the MAFFT online server v. 7 (Katoh et al., 2019) with the MAFFT FFT-NS-2 progressive alignment algorithm (Rozewicki et al., 2019). A smaller alignment was similarly generated to more easily visualize the relationship between Fsa16295Glu and its closest relatives. For this smaller phylogeny, the MAFFT-DASH structural alignment algorithm (Katoh et al., 2019) was used instead to bring structural information into account. The larger alignment was then subjected to an approximate maximum

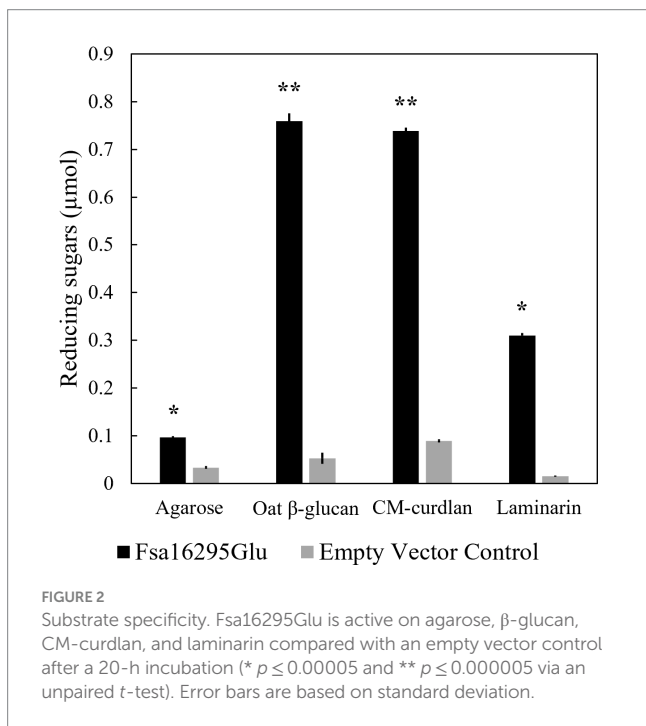
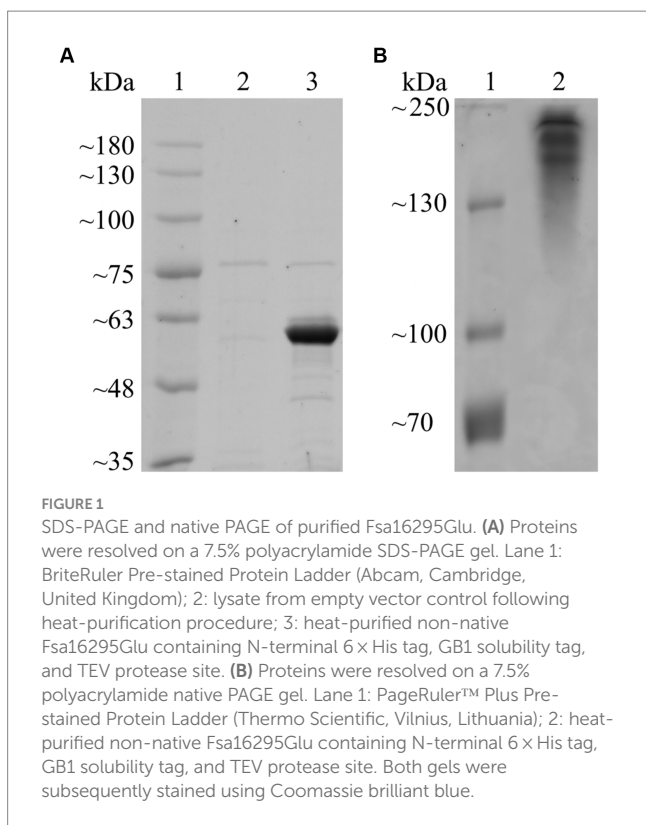
likelihood analysis with FastTree v. 2.1.11 (Price et al., 2010) using Shimodaira–Hasegawa approximate likelihood ratio test (SH-aLRT) branch support values (Guindon et al., 2010) from 1,000 replicates. The smaller dataset was subjected to a true maximum likelihood analysis with IQTree v. 2.2.0 (Minh et al., 2020) using SH-aLRT branch support values from 1,000 replicates and ultrafast bootstrapping (UFBoot) branch support from 1,000 replicates. The obtained phylogenies were visualized using the Interactive Tree of Life v. 6.7 website (<https://itol.embl.de>) and edited in Inkscape v. 0.92.4 (<https://www.inkscape.org>).

## 3 Results

### 3.1 Computed structure modeling, heterologous expression, and substrate screening

Following annotation by HMMER (E-value =  $3.90 \times 10^{-102}$ ; cutoff =  $1 \times 10^{-15}$ ) and dbCAN-sub (E-value =  $5.00 \times 10^{-134}$ ; cutoff =  $1 \times 10^{-15}$ ), Fsa16295Glu was determined to belong to GH50 with no carbohydrate-binding modules or other annotated domains. Fsa16295Glu had an N-terminal signal sequence (MSWTRREFVKLVGFATTFAGASCRSEG) that was predicted using SignalP to belong to the twin-arginine translocation (TAT) pathway, suggesting that Fsa16295Glu is secreted to the extracellular space as a folded protein (Ren et al., 2016). A multiple sequence alignment of seven GH50 proteins (Fsa16295Glu, Fsa11540, Aga50D, AgaA, AgrA, PaBglu50A, and VadG925) revealed that Fsa16295Glu has two residues, Glu192 and Glu354, that align with known catalytic residues of other clan GH-A enzymes and are highly conserved among the clan (Kumar et al., 2019). Similarly, Fsa11540 has the conserved residues at Glu409 and Glu575 (Supplementary Figure S1). After expression and heat purification, the purity of Fsa16295Glu was estimated by SDS-PAGE to be >90%, and the molecular weight was estimated at 60 kDa (Figure 1). The size of the recombinant protein was affected by the removal of the TAT signal sequence (3 kDa) and the addition of a GB1 solubility tag (6 kDa), a 6  $\times$  His tag (1 kDa), and a TEV protease site (1 kDa). These modifications made the recombinant Fsa16295Glu slightly larger than the predicted native size of 52 kDa (Supplementary Figure S2). The resolved native PAGE gel revealed three bands that are apparently larger than 130 kDa, suggesting the formation of multimeric structures (Figure 1).

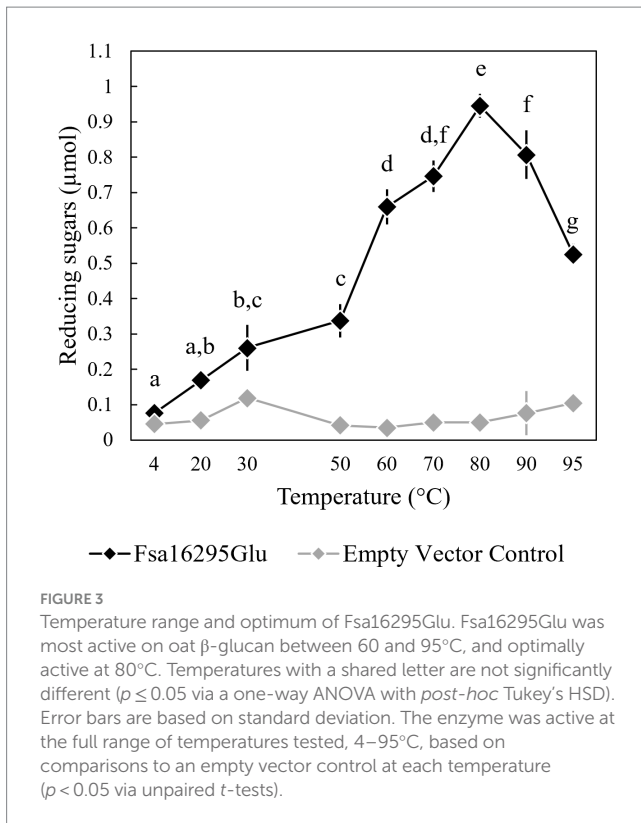
The CSM of native, mature Fsa16295Glu had low predicted aligned error (PAE) scores (0–5) indicative of high confidence in the accuracy of the CSM (Supplementary Figures S3A,B). Glu192 and Glu354 (Glu165 and Glu327 in the mature form) are found in a tunnel through the structure's center, marking this as the candidate active site (Supplementary Figure S3C). Furthermore, the residues surrounding the active site tunnel had higher PAE scores (2.5–5), suggesting greater structural flexibility, supporting this tunnel as the active site. The CSM of non-native Fsa16295Glu containing an N-terminal 6  $\times$  His tag, GB1 solubility tag, and a TEV protease site (Supplementary Figure S3D) also had low PAE scores (0–5) in the catalytic domain, indicative of normal predicted folding, and was the form of Fsa16295Glu used for biochemical characterization. CSMs of the homodimeric and homotrimeric forms of Fsa16295Glu



predicted asymmetric multimers (Supplementary Figure S3E), supporting multimerization of Fsa16295Glu observed by native PAGE. However, the PAE plot of the homodimer had low PAE scores within each monomer (0–5), and low-to-moderate PAE scores between monomers (5–15), while the PAE plots of the homotrimer and homotetramer showed high PAE scores between monomers (25–30). These results suggest low confidence in a homotrimer or

homotetramer structure and moderate confidence in a homodimer structure of Fsa16295Glu (Supplementary Figure S3F). The CSM of Fsa16295Glu was similar to the CSMs of three uncharacterized β-agarases in the AlphaFold Protein Structure Database (Jumper et al., 2021; Varadi et al., 2022), all from the bacterial phylum *Armatimonadota*. These three predicted enzymes also had high Protein Basic Local Alignment Search Tool [NCBI BLASTp] scores with Fsa16295Glu: A0A7C3N829 [UniProt accession number], (E-value=0, percent identity=92.68%; A0A7C2VJI7, E-value=0, percent identity=78.17%; and A0A2H5XB18, E-value=0, percent identity=77.90%) (Kato et al., 2018; Zhou et al., 2020). Previously experimentally validated GH50 structures most similar to Fsa16295Glu were Aga50D, a homotetrameric β-agarase from *Saccharophagus degradans* 2–40 (E-value=6 × 10<sup>-54</sup>, percent identity=33.42%) (Kim et al., 2010; Pluvinage et al., 2013) and AgWH50C, a homodimeric β-agarase from *Agarivorans gilvus* WH0801 (E-value=3 × 10<sup>-53</sup>, percent identity=33.24%) (Liu et al., 2012; Zhang et al., 2019). Aga50D homotetramers possess similar tunnel-shaped active sites on each molecule that contain conserved GH-A catalytic residues Glu695 and Glu534, which align with Glu354 and Glu192 from Fsa16295Glu, respectively. Additionally, Fsa16295Glu possesses other GH50 active site residues present in both Aga50D and AgWH50C: Glu417, Arg413, Asn191, and Phe403 in Fsa16295Glu; Glu757, Arg752, Asn533, and Phe742 in Aga50D; Glu705, Arg700, Asn483, and Phe690 in AgWH50C. Unlike Aga50D and AgWH50C, which have conserved N-terminal carbohydrate-binding module (CBM) domains, Fsa16295Glu does not have an annotated N-terminal CBM domain. Overall, these shared active site residues provide additional support for the putative active site tunnel of Fsa16295Glu. The catalytic mechanisms of Aga50D and AgWH50C were described as retaining and putatively retaining, respectively, based on the role of the conserved Aga50D active site glutamic acid residues as a nucleophile (Glu695) and proton donor (Glu534), which is consistent with the retaining mechanism of GH-A enzymes (Kumar et al., 2019). Thus, we putatively identified the catalytic mechanism of Fsa16295Glu to be retaining, wherein Glu354 acts as a nucleophile and Glu192 serves as a proton donor.

After screening with 20 potential substrates using the DNS method, Fsa16295Glu was highly active on CM-curdlan and oat β-glucan, and slightly active on laminarin and agarose (Figure 2). Removal of N-terminal 6 × His and GB1 solubility tags via TEV protease site cleavage had no effect on activity with oat β-glucan or any of the 11 other substrates that were tested. Complete acid hydrolysis of CM-curdlan, oat β-glucan, laminarin, and agarose, yielded 0.93 ± 0.01, 1.16 ± 0.08, 0.89 ± 0.01, and 0.66 ± 0.01, μmol of reducing sugars, suggesting that Fsa16295Glu incubated overnight at 80°C hydrolyzed 79.2% of the glycosidic bonds in CM-curdlan, 65.6% of those in oat β-glucan, 35.0% of those in laminarin, and 14.6% of those in agarose. Fsa16295Glu released 0.37 μmol of pNP from pNPG after 30 min, confirming that it had β-glucosidase activity (Supplementary Figure S4). Fsa16295Glu was not active on yeast β-glucan, CMC, chondroitin sulfate, colloidal chitin, brown algae fucoidan, lupin galactan, gellan gum, karaya gum, locust bean gum, xanthan gum, birch wood xylan, tamarind xyloglucan, oyster glycogen, potato starch, or the AFEX-pretreated substrates corn stover, *Miscanthus*, or sugarcane bagasse. The absence of activity on the insoluble AFEX-pretreated substrates is consistent with

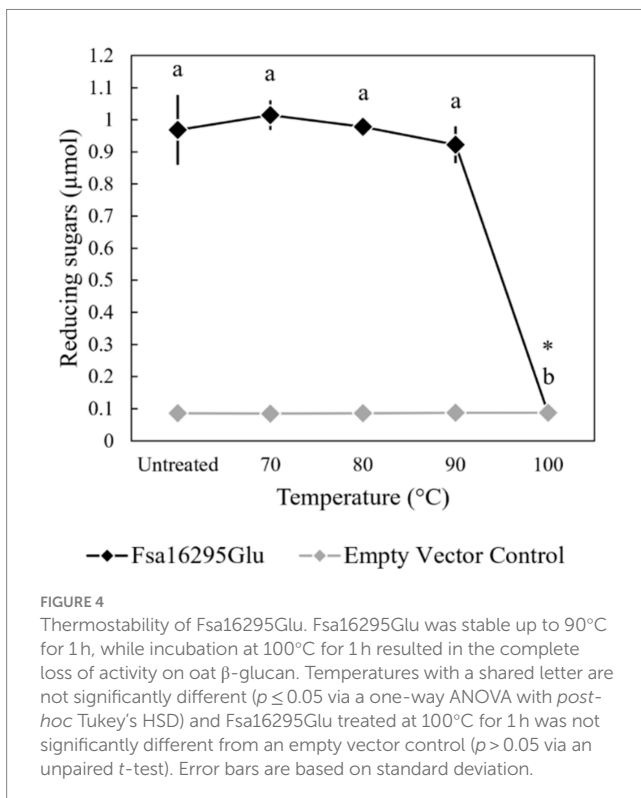


### 3.2 Biochemical characterization and kinetic parameters

When tested with oat  $\beta$ -glucan at a range of temperatures and pH values, Fsa16295Glu was shown to be hyperthermophilic with a broad pH range. Fsa16295Glu was active between 4 and 95°C with an optimum of  $\sim 80^\circ\text{C}$  (Figure 3). Fsa16295Glu was thermostable up to at least 90°C, with similar amounts of reducing sugars released following a 1-h treatment at 70°C (1.02  $\mu\text{mol}$ ), 80°C (0.98  $\mu\text{mol}$ ), and 90°C (0.92  $\mu\text{mol}$ ), while pre-incubation for 1 h at 100°C resulted in the complete loss of activity (Figure 4). Fsa16295Glu was active across the entire tested pH range of 5.5–11.0 and optimally active between 6.5 and 10.0 (Supplementary Figure S5). Using a Lineweaver–Burk plot ( $R^2 = 0.9907$ ) and assuming 100% enzyme activity with Fsa16295Glu at a concentration of 141  $\mu\text{g}/\text{mL}$ ,  $V_{\text{max}}$ ,  $K_M$ , and  $k_{\text{cat}}/K_M$  were calculated as 133  $\mu\text{M}/\text{min}$ , 0.0520  $\mu\text{M}$ , and  $1.82 \times 10^7 \text{ s}^{-1} \text{ M}^{-1}$ , respectively, using CM-curdlan (Supplementary Figure S6).  $V_{\text{max}}$ ,  $K_M$ , and  $k_{\text{cat}}/K_M$  were calculated as 625  $\mu\text{M}/\text{min}$ , 15.7  $\mu\text{M}$ , and  $2.83 \times 10^5 \text{ s}^{-1} \text{ M}^{-1}$ , respectively, using oat  $\beta$ -glucan (Supplementary Figure S7). These  $K_M$  and  $k_{\text{cat}}/K_M$  values reflect greater specificity and catalytic efficiency, respectively, with  $\beta$ -1,3-glycosidic linkages in CM-curdlan over the alternating  $\beta$ -1,3/1,4-glycosidic linkages in oat  $\beta$ -glucan.

### 3.3 Phylogenetic analysis of Fsa16295Glu within family GH50

Based on approximate maximum likelihood phylogenetic analyses of Fsa16295Glu, other GH50 enzymes from class *Fervidibacteria* gathered from related metagenome-assembled genomes (Nou et al., 2024), and all members of family GH50 within the CAZy database, Fsa16295Glu and homologs from other *Fervidibacteria* had a unique phylogenetic placement compared to other characterized GH50s (Figure 5A). *Fervidibacteria* GH50s were placed into five monophyletic clades, containing 18, 19, 22, 30, and 45 sequences, with Fsa16295Glu belonging to the largest clade. The 25 previously characterized GH50s included in this analysis belonged to *Pseudomonadota* (22 from *Agarivorans gilvus* WH0801, *Agarivorans* sp. AG17, *Agarivorans* sp. HZ105, *Agarivorans* sp. JA-1, *Agarivorans* sp. JAMB-A11, *Agarivorans* sp. QM38, *Alteromonas* sp. E-1, *Paraglaciecola hydrolytica* S66, *Pseudoalteromonas* sp. NJ21, *P. aeruginosa*, *Saccharophagus degradans* 2–40, *Thalassotalea agarivorans* BCRC 17492, *Vibrio* sp. CN41, *Vibrio* sp. PO-303); *Verrucomicrobiota* (one from *V. vadensis* ATCC BAA-548); *Bacillota* (one from *Paenibacillus agarexedens* BCRC 16000); and *Actinomycetota* (one from *Streptomyces coelicolor* A3(2)). Most *Pseudomonadota* sequences clustered into two groups of enzymes according to their lowest common taxonomic rank:  $\beta$ -agarases from *Enterobacterales* (genera *Agarivorans* and *Vibrio*) and  $\beta$ -agarases from a broader group of *Gammaproteobacteria* (including the order *Enterobacterales\_A* and the *Pseudomonadales* genera *Alteromonas*, *Paraglaciecola*, *Pseudoalteromonas*, *Pseudomonas*, *Saccharophagus*, and *Thalassotalea*.) Consistent with their distinct activities from the other characterized enzymes, VadG925 (homolog 1 in Figure 5A and Supplementary Table S1) and PaBglu50A (homolog 25 in Figure 5A and Supplementary Table S1) were distant from the well-characterized GH50 groups. The maximum likelihood analysis showed that the *Fervidibacteria* GH50 group that Fsa16295Glu belongs to was most closely related to six archaeal GH50s from two



Fsa16295Glu not having a CBM domain and suggests that Fsa16295Glu may be involved in oligosaccharide degradation following initial hydrolysis of insoluble polysaccharides by GHs containing CBM domains.

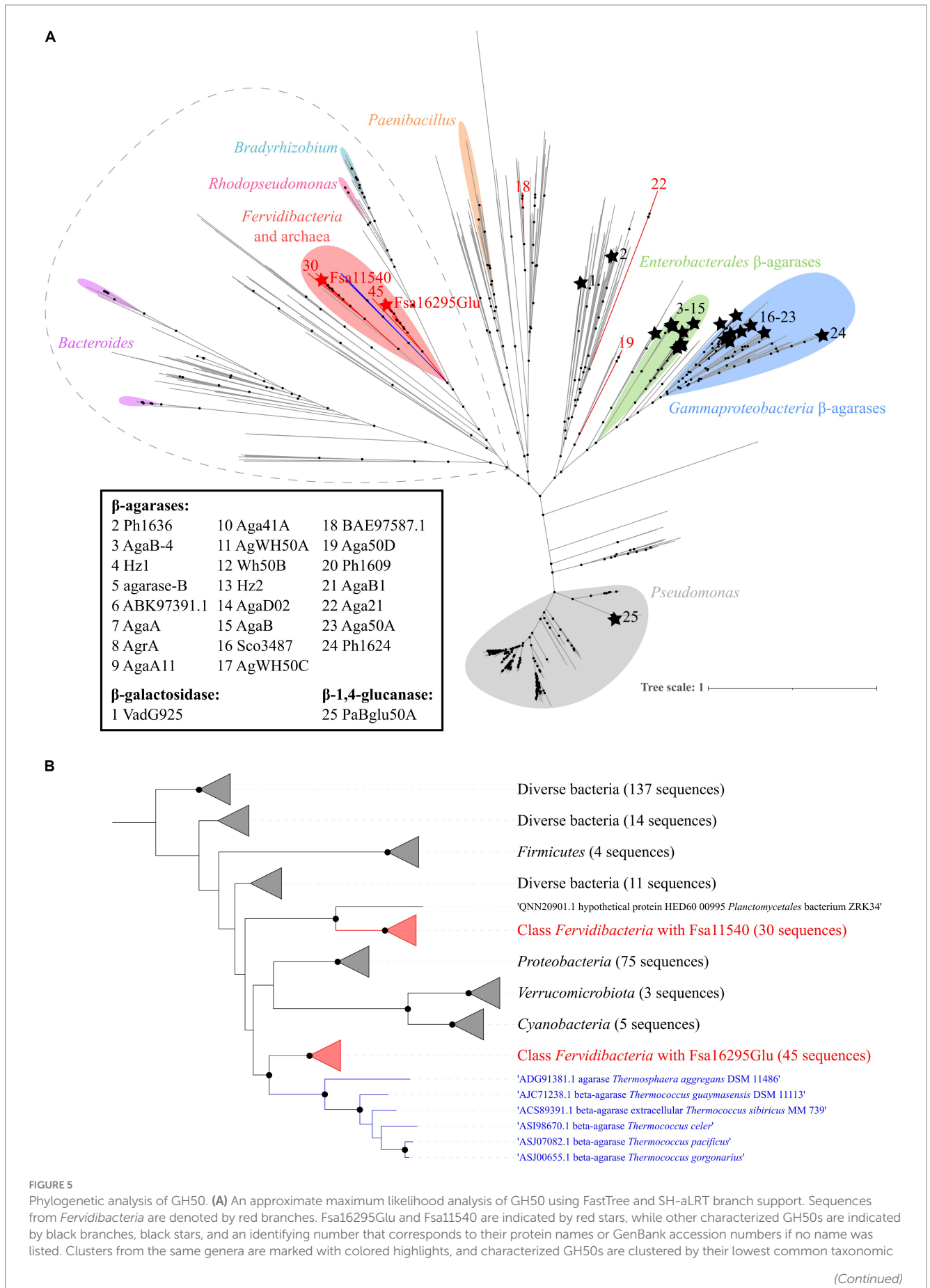




FIGURE 5 (Continued)

rank. The sequences included in panel B are circled by a dashed gray line. Well-supported branches (SH-aLRT  $\geq 0.9$ ) are indicated by a black dot. The scale bar shows the number of amino acid changes per site. (B) A maximum likelihood analysis of the Fsa16295Glu lineage and its closest relatives (circled in dashed grey outline) using SH-aLRT and UFBoot branch support. Sequences are clustered into five *Fervidibacteria* clades (red; the number of sequences is numbered at the tips), the archaeal GH50s (blue), and related GH50s by taxonomic rank. Well-supported branches (SH-aLRT  $\geq 0.9$  and UFBoot  $\geq 95\%$ ) are indicated by a black dot. The scale bar shows the number of amino acid changes per site.

unrelated archaea, *Thermosphaera aggregans* in the *Thermoproteota*, and several species of *Thermococcus* in the Methanobacteriota\_B (Figure 5B) (Taxonomic names are listed as in GTDB.) Along with sequences from *Fervidibacteria*, these archaeal enzymes were the only GH50s in this analysis from thermophiles. The monophyly of *Fervidibacteria* homologs in this lineage suggests vertical evolution of the enzymes related to Fsa16295Glu in the class *Fervidibacteria*, with horizontal transfers to different thermophilic archaea, either independently, or horizontal transfer to *T. aggregans*, and then from *T. aggregans* to *Thermococcus*. The second “*F. sacchari*” GH50 enzyme, Fsa11540, belonged to a second major *Fervidibacteria* clade, implying that the two paralogous GH50s of “*F. sacchari*” represent distinct enzymes with potentially different activities or regulation. Organisms encoding GH50s that were related to Fsa16295Glu and Fsa11540 included members of the *Pseudomonadota* (75), *Cyanobacteriota* (5), *Bacillota* (4), and *Verrucomicrobiota* (3).

## 4 Discussion

Prior to this study, characterized GH50s have been almost exclusively classified as  $\beta$ -agarases, with the notable exceptions being a  $\beta$ -galactosidase and a  $\beta$ -1,3-glucanase (Temuujin et al., 2012a; Yi et al., 2018). Our study, however, reveals a GH50  $\beta$ -1,3-endoglucanase that is phylogenetically distant from the previously characterized  $\beta$ -1,3-glucanase, PaBglu50A. Although Fsa16295Glu could degrade agarose, suggesting some  $\beta$ -agarase activity, its hydrolysis of CM-curdlan and oat  $\beta$ -glucan released considerably more reducing sugars. Thus, Fsa16295Glu is the first known GH50 to degrade substrates with both galactose units, found in agarose, and glucose units, found in CM-curdlan, oat  $\beta$ -glucan, and laminarin. Fsa16295Glu furthermore deviates from other GH50s through its high temperature optimum of 80°C, a marked increase over the second-highest temperature optimum in GH50: 65°C of AgrA from *Agarivorans* sp. AG17 (Nikapitiya et al., 2010). Overall, Fsa16295Glu is a functionally distinct GH50 and the first known hyperthermophilic enzyme within the family GH50.

Compared with other characterized GH50s, Fsa16295Glu has a relatively high  $k_{cat}/K_M$  of  $1.82 \times 10^7 \text{ s}^{-1} \text{ M}^{-1}$ . GH50  $\beta$ -agarase  $k_{cat}/K_M$  values range from  $8.1 \times 10^2$  to  $5.7 \times 10^7 \text{ s}^{-1} \text{ M}^{-1}$ , with Fsa16295Glu falling just under that of Sco3487 from *Streptomyces coelicolor* A3(2), which was described as having a high catalytic efficiency ( $5.7 \times 10^7 \text{ s}^{-1} \text{ M}^{-1}$ ) (Fu et al., 2008; Temuujin et al., 2012b; Chen et al., 2018). The lower  $K_M$  of Fsa16295Glu toward CM-curdlan (0.0520  $\mu\text{M}$ ) compared with oat  $\beta$ -glucan (15.7  $\mu\text{M}$ ) indicates preferential hydrolysis of  $\beta$ -1,3-glucans like CM-curdlan. The inability of Fsa16295Glu to degrade CMC suggests it cannot cleave  $\beta$ -1,4-glycosidic linkages connecting glucose, which is supported by the partial (65.6%) hydrolysis of oat  $\beta$ -glucan, indicating hydrolysis of  $\beta$ -1,3-, but not  $\beta$ -1,4-glycosidic linkages. Furthermore, the near-complete (79.2%) hydrolysis of CM-curdlan further supports the catalysis of  $\beta$ -1,3-glycosidic linkages. That CM-curdlan is not more thoroughly degraded

despite the ubiquity of  $\beta$ -1,3-glycosidic linkages may be caused by the presence of carboxymethyl groups on every third glucose, potentially limiting access by Fsa16295Glu to one or both sides of the carboxymethylated residue. Thus, possible degradation products include glucose and CM-glucose, with some Glc ( $\beta$ 1,3)CM-Glc or CM-Glc ( $\beta$ 1,3)Glc. Side chain inhibition is supported by the comparatively poor (35.0%) hydrolysis of laminarin and the inability of Fsa16295Glu to hydrolyze yeast  $\beta$ -glucan. Laminarin from stramenopiles, such as that used in this study, contains a  $\beta$ -1,6-linked glucose side chain on every third residue (Chen et al., 2021), which may further limit access by Fsa16295Glu. Additionally, the inability of Fsa16295Glu to degrade yeast  $\beta$ -glucan suggests that side chain inhibition may be exacerbated by chain length, as yeast  $\beta$ -glucan frequently contains side chains with more than one glucose (Liu et al., 2021). Although degradation of  $\beta$ -1,4-glucan was not observed, activity on agarose suggests some cleavage of  $\beta$ -1,4-glycosidic linkages connecting galactose residues. Due to the microbial origin of some  $\beta$ -1,3-glucans (Mudgil, 2017; Paudel et al., 2021), we speculate that bacteria found *in situ* are the natural source of  $\beta$ -1,3-glucan. Various bacteria produce curdlan, including *Agrobacterium*, *Rhizobium*, and *Cellulomonas* (Gao et al., 2020), but none are known to be thermophilic, obscuring the possible source of curdlan, if it exists, in the geothermal spring environment of “*F. sacchari*” and other *Fervidibacteria*. One possible source of  $\beta$ -1,3-glucan is the photosynthetic mats found at the fringes of the Great Boiling Spring (GBS) that were originally used to enrich “*F. sacchari*” (Nou et al., 2024). Glucose is found in more than 90% of polysaccharides released by *Cyanobacteriota*, which are prominent constituents of photosynthetic mats, and is often the most abundant monosaccharide (De Philippis and Vincenzini, 1998; George et al., 2023). However, photosynthetic mats at GBS are located ~1–2 meters away from higher temperature sediments hosting abundant populations of “*F. sacchari*,” and the polysaccharide composition of these mats has not been explored. Further study is needed to explore the specific  $\beta$ -1,3-glucans depolymerized by Fsa16295Glu and their sources in geothermal springs.

Fsa16295Glu has a unique phylogenetic position relative to other characterized GH50s. Aside from other GH50s from *Fervidibacteria*, Fsa16295Glu is most closely related to a subset of six GH50s encoded by several *Thermococcus* species and *T. aggregans*. Given the presence of homologs of Fsa16295Glu in most available *Fervidibacteria* genomes and the distant relationship between *Fervidibacteria*, *Thermococcus*, and *Thermosphaera*, we infer that this enzyme passed through horizontal gene transfer from *Fervidibacteria* to these archaea, which are common co-inhabitants in geothermal environments. The lower activity of Fsa16295Glu on agarose coincides with substrate availability, as all agarolytic GH50s are from marine environments, where agar is a major constituent of red algae (Yun et al., 2021). The presence of weak agarase activity in Fsa16295Glu suggests that  $\beta$ -agarase and  $\beta$ -1,3-glucanase activities may have been ancestral in GH50, but adaptation to polysaccharides in different environments was

selected for divergent activities. Further characterization of diverse GH50s and ancestral character state reconstructions could offer insight into the validity of this hypothesis.

The assignment of GHs into families based on amino acid sequences is somewhat limited by the breadth of known GH families and the limited knowledge of their activities. However, new GH families are often discovered, as recently as the discovery of GH174 by Liu et al. (2023). Because of this, the functional and phylogenetic novelty of Fsa16295Glu prompted us to consider whether Fsa16295Glu should be assigned to a new GH family. However, the presence of two conserved GH-A glutamic acid residues in Fsa16295Glu and the relatively high amino acid identity to canonical GH50s (29.91–33.26%) suggests that Fsa16295Glu correctly falls within GH50 but should be distinguished from other characterized GH50s. Viborg et al. (2019) previously proposed a roadmap for identifying GH subfamilies based on combinations of phylogenetic coherence and functional similarity. Because of their close phylogenetic relationship and shared agarolysis, we propose GH50s from the clade containing characterized *Enterobacteriales* and *Gammaproteobacteria*  $\beta$ -agarases be designated as subfamily GH50\_1. Similarly, we propose the  $\beta$ -1,3-glucanase PaBglu50A from the *Pseudomonas* clade be designated as the first characterized member of GH50\_2. Finally, because of its  $\beta$ -1,3-glucanase activity, 80°C optimum, and distant relationship to other characterized GH50s, we propose Fsa16295Glu be designated as the first characterized representative of GH50\_3.

## Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

## Author contributions

JC: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Supervision, Visualization, Writing – original draft, Writing – review & editing. NT: Writing – review & editing, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology. AC: Writing – review & editing, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology. MP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Supervision, Visualization, Writing – review & editing. SB: Writing – review & editing, Data curation, Formal analysis, Investigation, Methodology. NN: Conceptualization, Data curation, Funding acquisition, Writing – review & editing. RM: Writing – review & editing, Conceptualization, Data curation, Methodology. MH-S: Writing – review & editing, Data curation, Project administration. IB: Conceptualization, Data curation, Methodology, Project administration, Supervision, Writing – review & editing. J-FC: Writing – review & editing, Conceptualization, Data curation, Methodology, Project administration, Supervision. MH: Conceptualization, Methodology, Resources, Writing – review & editing. PB: Methodology, Writing – review & editing. NS: Writing – review & editing, Funding acquisition, Methodology. KV: Funding acquisition, Methodology, Writing – review & editing. BH: Supervision,

Writing – original draft, Writing – review & editing, Conceptualization, Funding acquisition, Project administration, Resources.

## Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This material is based upon work supported by the National Aeronautics and Space Administration under Grants No. 80NSSC20M0043, 80NSSC17K0548, 80NSSC21M0157, the US National Science Foundation under grants DBI REU 1757316 and HRD-1712523, and the UNLV Office of Undergraduate Research. The work (proposal: <https://doi.org/10.46936/10.25585/60001337>) conducted by the U.S. Department of Energy Joint Genome Institute (<https://ror.org/04xm1d337>), a DOE Office of Science User Facility, is supported by the Office of Science of the U.S. Department of Energy operated under Contract No. DE-AC02-05CH11231. The publication fees for this article were supported by the Graduate & Professional Student Association (GPSA).

## Acknowledgments

We thank Ernesto Abel-Santos, Duane P. Moser, Kurt Regner, and Helen J. Wing for their contributions as the graduate student advisory committee evaluating this study. We thank Casey Hall-Wheeler for SpectraMax Plus Plate Reader training. We thank Monika M. A. Karney for SDS-PAGE training.

## Conflict of interest

PB was employed by C5-6 Technologies LLC.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor E-MZ declared a past collaboration with the author BH.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2024.1355444/full#supplementary-material>

## References

- Adams, M. W. W., and Kelly, R. M. (2001). *Hyperthermophilic enzymes part a*. (Academic Press).
- Bhalla, A., Bansal, N., Kumar, S., Bischoff, K. M., and Sani, R. K. (2013). Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresour. Technol.* 128, 751–759. doi: 10.1016/j.biortech.2012.10.145
- Blumer-Schuette, S. E., Kataeva, I., Westpheling, J., Adams, M. W., and Kelly, R. M. (2008). Extremely thermophilic microorganisms for biomass conversion: status and prospects. *Curr. Opin. Biotechnol.* 19, 210–217. doi: 10.1016/j.copbio.2008.04.007
- Chen, Z. W., Lin, H. J., Huang, W. C., Hsuan, S. L., Lin, J. H., and Wang, J. P. (2018). Molecular cloning, expression, and functional characterization of the  $\beta$ -agarase AgaB-4 from *Paenibacillus agarexedens*. *AMB Express* 8, 49–10. doi: 10.1186/s13568-018-0581-8
- Chen, Z., Pereira, J. H., Liu, H., Tran, H. M., Hsu, N. S., Dibble, D., et al. (2013). Improved activity of a thermophilic cellulase, Cel5A, from *Thermotoga maritima* on ionic liquid pretreated switchgrass. *PLoS One* 8:e79725. doi: 10.1371/journal.pone.0079725
- Chen, J., Yang, J., Du, H., Aslam, M., Wang, W., Chen, W., et al. (2021). Laminarin, a major polysaccharide in stramenopiles. *Mar. Drugs* 19:576. doi: 10.3390/md19100576
- Cole, J. K., Peacock, J. P., Dodsworth, J. A., Williams, A. J., Thompson, D. B., Dong, H., et al. (2013). Sediment microbial communities in great boiling spring are controlled by temperature and distinct from water communities. *ISME J.* 7, 718–729. doi: 10.1038/ismej.2012.157
- Costa, K. C., Navarro, J. B., Shock, E. L., Zhang, C. L., Soukup, D., and Hedlund, B. P. (2009). Microbiology and geochemistry of great boiling and mud hot springs in the United States Great Basin. *Extremophiles* 13, 447–459. doi: 10.1007/s00792-009-0230-x
- Crosby, J. R., Laemthong, T., Bing, R. G., Zhang, K., Tanwee, T. N. N., Lipscomb, G. L., et al. (2022). Biochemical and regulatory analyses of xyranolytic regulons in *Caldicellulosiruptor bescii* reveal genus-wide features of hemicellulose utilization. *Appl. Environ. Microbiol.* 88:e0130222. doi: 10.1128/aem.01302-22
- De Philippis, R., and Vincenzini, M. (1998). Exocellular polysaccharides from cyanobacteria and their possible applications. *FEMS Microbiol. Rev.* 22, 151–175. doi: 10.1016/S0168-6445(98)00012-6
- Drula, E., Garron, M. L., Dogan, S., Lombard, V., Henrissat, B., and Terrapon, N. (2022). The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res.* 50, D571–D577. doi: 10.1093/nar/gkab1045
- Fu, X. T., Lin, H., and Kim, S. M. (2008). Purification and characterization of a novel  $\beta$ -agarase, AgaA34, from *Agarivorans albus* YKW-34. *Appl. Microbiol. Biotechnol.* 78, 265–273. doi: 10.1007/s00253-007-1303-3
- Gao, H., Xie, F., Zhang, W., Tian, J., Zou, C., Jia, C., et al. (2020). Characterization and improvement of curdlan produced by a high-yield mutant of *Agrobacterium* sp. ATCC 31749 based on whole-genome analysis. *Carbohydr. Polym.* 245:116486. doi: 10.1016/j.carbpol.2020.116486
- George, C., Lim, C. X. Q., Tong, Y., and Pointing, S. B. (2023). Community structure of thermophilic photosynthetic microbial mats and flocs at Sembawang hot spring, Singapore. *Front. Microbiol.* 14:1189468. doi: 10.3389/fmicb.2023.1189468
- Giles, K., Pluvinage, B., and Boraston, A. B. (2017). Structure of a glycoside hydrolase family 50 enzyme from a subfamily that is enriched in human gut microbiome Bacteroidetes. *Proteins: Struct. Funct. Genet.* 85, 182–187. doi: 10.1002/prot.25189
- Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010
- Hall, T. A. (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT. *Nucleic Acids Symp. Ser.* 41, 95–98. doi: 10.14601/Phytopathol\_Mediterr-14998u1.29
- Hedlund, B. P., Thomas, S. C., Dodsworth, J. A., and Zhang, C. L. (2016). “Life in high-temperature environments” in *Manual of environmental microbiology*. eds. M. V. Yates, C. H. Nakatsu, R. V. Miller and S. D. Pillai. 4th ed (Washington, DC: ASM Press), 863–877.
- Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316, 695–696. doi: 10.1042/bj3160695
- Hsu, S. C., and Lockwood, J. L. (1975). Powdered chitin agar as a selective medium for enumeration of Actinomycetes in water and soil. *Appl. Microbiol.* 29, 422–426. doi: 10.1128/AM.29.3.422-426.1975
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., et al. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589. doi: 10.1038/s41586-021-03819-2
- Kato, S., Sakai, S., Hirai, M., Tasumi, E., Nishizawa, M., Suzuki, K., et al. (2018). Long-term cultivation and metagenomics reveal ecophysiology of previously uncultivated thermophiles involved in biogeochemical nitrogen cycle. *Microbes Environ.* 33, 107–110. doi: 10.1264/jsm2.ME17165
- Katoh, K., Rozewicki, J., and Yamada, K. D. (2019). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108
- Kim, H. T., Lee, S., Lee, D., Kim, H. S., Bang, W. G., Kim, K. H., et al. (2010). Overexpression and molecular characterization of Aga50D from *Saccharophagus degradans* 2-40: an exo-type beta-agarase producing neoagarobiose. *Appl. Microbiol. Biotechnol.* 86, 227–234. doi: 10.1007/s00253-009-2256-5
- Kumar, R., Henrissat, B., and Coutinho, P. M. (2019). Intrinsic dynamic behavior of enzyme:substrate complexes govern the catalytic action of  $\beta$ -galactosidases across clan GH-A. *Sci. Rep.* 9, 10346–10314. doi: 10.1038/s41598-019-46589-8
- Lee, D. G., Jang, M. K., Lee, O. H., Kim, N. Y., Ju, S. A., and Lee, S. H. (2008). Overproduction of a glycoside hydrolase family 50 beta-agarase from *Agarivorans* sp. JA-1 in *Bacillus subtilis* and the whitening effect of its product. *Biotechnol. Lett.* 30, 911–918. doi: 10.1007/s10529-008-9634-4
- Liu, N., Mao, X., Yang, M., Mu, B., and Wei, D. (2012). Gene cloning, expression and characterisation of a new  $\beta$ -agarase, AgWH50C, producing neoagarobiose from *Agarivorans gilvus* WH0801. *World J. Microbiol. Biotechnol.* 30, 1691–1698. doi: 10.1007/s1274-013-1591-y
- Liu, G., Shen, J., Chang, Y., Mei, X., Chen, G., Zhang, Y., et al. (2023). Characterization of an endo-1,3-fucanase from marine bacterium *Wenyngzhuangia aestuarii*: the first member of a novel glycoside hydrolase family GH174. *Carbohydr. Polym.* 306:120591. doi: 10.1016/j.carbpol.2023.120591
- Liu, Y., Wu, Q., Wu, X., Algharib, S. A., Gong, F., Hu, J., et al. (2021). Structure, preparation, modification, and bioactivities of  $\beta$ -glucan and mannan from yeast cell wall: a review. *Int. J. Biol. Macromol.* 173, 445–456. doi: 10.1016/j.ijbiomac.2021.01.125
- Meng, D.-D., Wang, B., Ma, X. Q., Ji, S. Q., Lu, M., and Li, F. L. (2016). Characterization of a thermostable endo-1,3(4)- $\beta$ -glucanase from *Caldicellulosiruptor* sp. strain F32 and its application for yeast lysis. *Appl. Microbiol. Biotechnol.* 100, 4923–4934. doi: 10.1007/s00253-016-7334-x
- Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37, 1530–1534. doi: 10.1093/molbev/msaa015
- Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022). ColabFold: making protein folding accessible to all. *Nat. Methods* 19, 679–682. doi: 10.1038/s41592-022-01488-1
- Mudgil, D. (2017). “Chapter 3 - the interaction between insoluble and soluble fiber” in *Dietary Fiber for the prevention of cardiovascular disease*. ed. R. A. Samaan (Elsevier Inc), 35–59.
- Nikapitiya, C., Oh, C., Lee, Y.-D., Lee, S.-K., Whang, I.-S., and And Lee, J.-H. (2010). Characterization of a glycoside hydrolase family 50 thermostable  $\beta$ -agarase AgrA from marine bacteria *Agarivorans* sp. AG17. *Fish. Aquat. Sci.* 13, 36–48. doi: 10.5657/FAS.2010.13.1.036
- Nou, N. O., Covington, J. K., Lai, D., Mayali, X., Seymour, C. O., Johnston, J., et al. (2024). Genome-enabled isolation of *Fervidibacter sacchari*, an aerobic, hyperthermophilic polysaccharide-degrading specialist. *Res. Square*. [Preprint]. doi: 10.21203/rs.3.rs-3760178/v1
- Oberortner, E., Cheng, J.-F., Hillson, N. J., and Deutsch, S. (2017). Streamlining the design-to-build transition with build-optimization software tools. *ACS Synth. Biol.* 6, 485–496. doi: 10.1021/acssynbio.6b00200
- Ochs, R. S. (2022). *Biochemistry*. (Second ed). (Boca Raton, FL: CRC Press), 125–126.
- Paudel, D., Dhungana, B., Caffè, M., and Krishnan, P. (2021). A review of health-beneficial properties of oats. *Food Secur.* 10:2591. doi: 10.3390/foods10112591
- Peacock, J. P., Cole, J. K., Murugapiran, S. K., Dodsworth, J. A., Fisher, J. C., Moser, D. P., et al. (2013). Pyrosequencing reveals high-temperature cellulolytic microbial consortia in great boiling spring after in situ lignocellulose enrichment. *PLoS One* 8:e59927. doi: 10.1371/journal.pone.0059927
- Petersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., et al. (2021). UCSF ChimeraX: structure visualization for researchers, educators, and developers. *Protein Sci.* 30, 70–82. doi: 10.1002/pro.3943
- Pluvinage, B., Hehemann, J.-H., and Boraston, A. B. (2013). Substrate recognition and hydrolysis by a family 50 exo- $\beta$ -agarase, Aga50D, from the marine bacterium *Saccharophagus degradans*. *J. Biol. Chem.* 288, 28078–28088. doi: 10.1074/jbc.M113.491068
- Pluvinage, B., Robb, C. S., Jeffries, R., and Boraston, A. B. (2020). The structure of PfGH50B, an agarase from the marine bacterium *Pseudoalteromonas fuliginea* PS47. *Acta Crystallogr.* 76, 422–427. doi: 10.1107/S2053230X20010328
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490
- Ren, G. H., Cao, L. C., Kong, W., Wang, Z. J., and Liu, Y. H. (2016). Efficient secretion of the  $\beta$ -galactosidase Bgal1-3 via both tat-dependent and tat-independent pathways in *Bacillus subtilis*. *J. Agric. Food Chem.* 64, 5708–5716. doi: 10.1021/acs.jafc.6b01735
- Rozewicki, J., Li, S., Amada, K. M., Standley, D. M., and Katoh, K. (2019). MAFFT-DASH: integrated protein sequence and structural alignment. *Nucleic Acids Res.* 47, W5–W10. doi: 10.1093/nar/gkz342

- Santos, C. R., Costa, P. A., Vieira, P. S., Gonzalez, S. E., Correa, T. L., Lima, E. A., et al. (2020). Structural insights into  $\beta$ -1,3-glucan cleavage by a glycoside hydrolase family. *Nat. Chem. Biol.* 16, 920–929. doi: 10.1038/s41589-020-0554-5
- Strazzulli, A., Fusco, S., Cobucci-Ponzano, B., Moracci, M., and Contursi, P. (2017). Metagenomics of microbial and viral life in terrestrial geothermal environments. *Rev. Environ. Sci. Biotechnol.* 16, 425–454. doi: 10.1007/s11157-017-9435-0
- Sugano, Y., Matsumoto, T., Kodama, H., and Noma, M. (1993). Cloning and sequencing of agaA, a unique agarase 0107 gene from a marine bacterium, *Vibrio* sp. strain JT0107. *Appl. Environ. Microbiol.* 59, 3750–3756. doi: 10.1128/aem.59.11.3750-3756.1993
- Summers, E. L., Moon, C. D., Atua, R., and Arcus, V. L. (2016). The structure of a glycoside hydrolase 29 family member from a rumen bacterium reveals unique, dual carbohydrate-binding domains. *Acta Crystallogr.* 72, 750–761. doi: 10.1107/S2053230X16014072
- Temuujin, U., Chi, W. J., Park, J. S., Chang, Y. K., Song, J. Y., and Hong, S. K. (2012a). Identification and characterization of a novel  $\beta$ -galactosidase from *Victivallis vadensis* ATCC BAA-548, an anaerobic fecal bacterium. *J. Microbiol.* 50, 1034–1040. doi: 10.1007/s12275-012-2478-6
- Temuujin, U., Won-Jae, C., Yong-Keun, C., and Hong, S. K. (2012b). Identification and biochemical characterization of Sco3487 from *Streptomyces coelicolor* A3(2), an exo- and endo-type  $\beta$ -agarase-producing neogagarbiose. *J. Bacteriol. Res.* 194, 142–149. doi: 10.1128/JB.05978-11
- Teufel, F., Almagro Armenteros, J. J., Johansen, A. R., Gislason, M. H., Pihl, S. I., Tsirigos, K. D., et al. (2022). SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* 40, 1023–1025. doi: 10.1038/s41587-021-01156-3
- Varadi, M., Anyango, S., Deshpande, M., Nair, S., Natassia, C., Yordanova, G., et al. (2022). AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* 50, D439–D444. doi: 10.1093/nar/gkab1061
- Viborg, A. H., Terrapon, N., Lombard, V., Michel, G., Czjzek, M., Henrissat, B., et al. (2019). A subfamily roadmap of the evolutionarily diverse glycoside hydrolase family 16 (GH16). *J. Biol. Chem.* 294, 15973–15986. doi: 10.1074/jbc.RA119.010619
- Wang, X.-Y., and Fisher, P. B. (2015). “Chapter nine: scavenger receptors: emerging roles in cancer biology and immunology,” in *Immunotherapy of Cancer*. Ed. P. B. Fisher. (Cambridge, MA: Academic Press), 1–375.
- Waterhouse, A. M., Procter, J. B., Martin, D. M., Clamp, M., and Barton, G. J. (2009). Jalview version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics* 25, 1189–1191. doi: 10.1093/bioinformatics/btp033
- Yi, P., Yan, Q., Jiang, Z., and Wang, L. (2018). A first glycoside hydrolase family 50 endo- $\beta$ -1,3-d-glucanase from *Pseudomonas aeruginosa*. *Enzym. Microb. Technol.* 108, 34–41. doi: 10.1016/j.enzymtec.2017.09.002
- Yun, E. J., Yu, S., Park, N. J., Cho, Y., Han, N. R., Jin, Y. S., et al. (2021). Metabolic and enzymatic elucidation of cooperative degradation of red seaweed agarose by two human gut bacteria. *Sci. Rep.* 11:13955. doi: 10.1038/s41598-021-92872-y
- Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., et al. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 46, W95–W101. doi: 10.1093/nar/gky418
- Zhang, P., Zhang, J., Zhang, L., Sun, J., Li, Y., Wu, L., et al. (2019). Structure-based design of agarase AgWH50C from *Agarivorans gilvus* WH0801 to enhance thermostability. *Appl. Microbiol. Biotechnol.* 103, 1289–1298. doi: 10.1007/s00253-018-9540-1
- Zhou, Z., Liu, Y., Xu, W., Pan, J., Luo, Z., and Li, M. (2020). Genome- and community-level interaction insights into carbon utilization and element cycling functions of Hydrothermarchaeota in hydrothermal sediment. *mSystems* 5:5. doi: 10.1128/mSystems.00795-19