PLOS ONE

RESEARCH ARTICLE

Identification and characterization of a hyperthermophilic GH9 cellulase from the Arctic Mid-Ocean Ridge vent field

Anton A. Stepnov[1], **Lasse Fredriksen**[1], **Ida H. Steen**[2], **Runar Stokke**[2], **Vincent G. H. Eijsink**[1]*

1 Faculty of Chemistry, Biotechnology and Food Science, NMBU—Norwegian University of Life Sciences, Ås, Norway, **2** Department of Biological Sciences and KG Jebsen Centre for Deep Sea Research, University of Bergen, Bergen, Norway

* vincent.eijsink@nmbu.no

Citation: Stepnov AA, Fredriksen L, Steen IH, Stokke R, Eijsink VGH (2019) Identification and characterization of a hyperthermophilic GH9 cellulase from the Arctic Mid-Ocean Ridge vent field. PLoS ONE 14(9): e0222216. https://doi.org/10.1371/journal.pone.0222216

Editor: Jean-Guy Berrin, Institut National de la Recherche Agronomique, FRANCE

Received: May 27, 2019

Accepted: August 23, 2019

Published: September 6, 2019

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: This work was funded by the Research Council of Norway through the NorZymeD project, project number 221568 (VGHE). Infrastructure was in part supported by NorBioLab grants 226247 and 270038 provided by the Research Council of Norway (VGHE). The research cruise to the Arctic Mid-Ocean Ridge vent fields was funded by the

Abstract

A novel GH9 cellulase (AMOR_GH9A) was discovered by sequence-based mining of a unique metagenomic dataset collected at the Jan Mayen hydrothermal vent field. AMOR_GH9A comprises a signal peptide, a catalytic domain and a CBM3 cellulose-binding module. AMOR_GH9A is an exceptionally stable enzyme with a temperature optimum around 100˚C and an apparent melting temperature of 105˚C. The novel cellulase retains 64% of its activity after 4 hours of incubation at 95˚C. The closest characterized homolog of AMOR_GH9A is *Tf*Cel9A, a processive endocellulase from the model thermophilic bacterium *Thermobifida fusca* (64.2% sequence identity). Direct comparison of AMOR_GH9A and *Tf*Cel9A revealed that AMOR_GH9A possesses higher activity on soluble and amorphous substrates (phosphoric acid swollen cellulose, konjac glucomannan) and has an ability to hydrolyse xylan that is lacking in *Tf*Cel9A.

Introduction

Cellulose is a main structural component of plant biomass and the most abundant carbohydrate on Earth. It is composed of repeating d-anhydroglucose units linked by β(1→4) glycoside bonds [1,2]. Individual cellulose chains are arranged into crystalline microfibrils that are stabilized by an extensive network of intra- and intermolecular hydrogen bonds [1]. The renewability of cellulose makes it an attractive source of green energy, but its exploitation is complicated by its resistance to depolymerization [3].

In Nature, degradation of cellulose is carried out by the synergistic action of endo-acting and exo-acting enzymes that include glycosyl hydrolases (GHs) and lytic polysaccharide monooxygenases (LPMOs) [4,5]. However, despite decades of research, industrial enzymatic processing of cellulosic plant biomass is still hampered by enzyme costs [6]. Thus, there is a clear incentive for discovering better cellulases.

The temperature optima of cellulases in currently available commercial enzyme cocktails are typically around 50˚C [7], i.e., not particularly high and likely not optimal, for example

Research Council of Norway, project number 179560 (IHS).

Competing interests: The authors have declared that no competing interests exist.

considering the risk of microbial contamination. The introduction of thermostable enzymes could be beneficial since this would allow the use of higher temperatures, resulting in increased substrate solubility, lower viscosity and reduced microbial growth [8]. Furthermore, the use of thermostable cellulases can simplify process design by minimizing or eliminating cooling periods between stages that require different temperatures (e.g. between heat pre-treatment and enzymatic conversion) [8, 9].

Metagenomics has proven to be a powerful tool for the discovery of thermostable enzymes from microbial sources. The crucial advantage of this approach is the ability to access extremophile genomes in a culture-independent manner [10]. Metagenomics has been successfully used to mine for novel enzymes in various high temperature environments such as compost, hot springs, deserts and deep sea vent fields [11]. Deep sea vents are promising niches for the search of extremozymes because they accommodate an impressive variety of microorganisms some of which can grow at temperatures as high as 121˚C [12]. Although deep-sea hydrothermal vents are characterized by lack of plant biomass [13] cellulolytic activity is not uncommon in the bacterial communities in these environments [14, 15]. Microbial biofilms are thought to be the most likely source of complex polysaccharide substrates around hydrothermal vents [16].

In recent years, we have been exploring the biodiversity of the Jan Mayen hydrothermal vent field at the Arctic Mid-Ocean Ridge, where temperatures can rise up to 260˚C [17]. In this paper, we report on a novel hyperthermophilic GH9 cellulase, AMOR_GH9A, which was discovered by *in silico* mining of a metagenome from the Jan Mayen hydrothermal vent field. The closest characterized homolog of AMOR_GH9A is *Tf*Cel9A, a processive endocellulase from the moderately thermophilic bacterium *Thermobifida fusca* [18]. We have also expressed and purified *Tf*Cel9A, which was then used as a reference enzyme in a comparative assessment of AMOR_GH9A functional properties. The results show that AMOR_GH9A has higher thermal stability and broader substrate specificity than its homolog from *T. fusca*.

Materials and methods

Sample collection, sequencing and identification of genes

A sample of unbleached Norway spruce (*Picea abies*) that had been pretreated by sulfite-pulping using the BALI process [19, 20], at Borregaard AS (Sarpsborg, Norway), was incubated for one year in ~70˚C hot sediments at the Arctic Mid-Ocean Ridge (AMOR), 570 m below sea level, and then recovered by a remotely operated vehicle. In short, 1 g of spruce material was mixed with approximately 16 mL of sediment sampled at the site and placed in the bottom chamber of a titanium incubator (2.5 cm chamber length, 16 mL chamber volume, 1 mm pores). The sampling was performed in a responsible way in accordance with the Norwegian Marine Resource Act and did not involve endangered or protected species. No permits were required to access the sampling site. DNA was extracted from 6.9 g of material and 1.1 μg of DNA was submitted for sequencing. The sampling procedure and the methods used for DNA extraction and sequencing have been described in detail elsewhere [21, 22]. Filtering and assembly of the raw Illumina MiSeq 300 paired-end reads were performed using the CLC genomics workbench utility (Qiagen, v.9.5.3) as previously described in [21, 22]. Open reading frames were predicted using Prodigal software [23, 24]. The resulting metagenomic dataset was mined for putative glycosyl hydrolases using the dbCAN service (csbl.bmb.uga.edu/dbCAN) [25]. The signal peptides of the candidate genes were annotated using SignalP [26]. The full characteristics of the metagenomic dataset will be published elsewhere.

The sequence-based mining led to the identification of a 2065 bp gene encoding a putative GH9 cellulase (AMOR_GH9A). The NCBI BLAST server (https://blast.ncbi.nlm.nih.gov/Blast.

cgi) was used to identify homologues of AMOR_GH9A. The sequence of AMOR_GH9A has been submitted to Genbank under accession number MK869727, and the DNA sequence of *Tf*Cel9A was obtained from GenBank (accession number L20093.1).

Gene synthesis and subcloning

The AMOR_GH9A and *Tf*Cel9A genes were codon optimized for expression in *E. coli* and synthesised by GenScript (Piscataway, NJ, USA). The genes were then amplified by PCR using Q5 high-fidelity DNA polymerase (New England Biolabs, Ipswich, MA, USA). The forward and reverse PCR primers incorporated plasmid-specific regions for ligation-independent cloning [27] to the pNIC-CH expression vector (AddGene, Cambridge, MA, USA) (see S1 Table for details). The PCR products were purified from 1% agarose gels using a Nucleospin Gel Clean-Up kit (Macherey-Nagel, Düren, Germany). After ligation-independent cloning, the reaction mixture was used for heat-shock transformation of OneShot TOP10 *E. coli* competent cells (Invitrogen, Carlsbad, USA) as recommended by the supplier. The transformed cells were incubated in SOC medium for 60 minutes at 37˚C prior to plating on LB agar medium supplied with 50 μg/ml kanamycin and 5% (w/v) sucrose. The clones from overnight incubation at 37˚C were screened for the target inserts by colony PCR using RedTaq polymerase (VWR International, Radnor, PA, USA) and pNIC-CH forward and reverse sequencing primers (see S1 Table). Positive clones were transferred to liquid LB medium with 50 μg/ml kanamycin for overnight cultivation at 37˚C, 200 rpm. The pNIC-CH plasmids harbouring target genes were purified using a NucleoSpin Plasmid kit (Macherey-Nagel, Düren, Germany) and the sequence of these expression vectors was confirmed by Sanger sequencing (GATC, Konstanz, Germany). The resulting expression plasmids code for AMOR_GH9A or *Tf*Cel9A without a signal peptide, starting with a methionine residue introduced at the N-terminus of the mature protein, and with a C-terminal affinity tag ("-AHHHHHH").

Expression and purification

AMOR_GH9A and *Tf*Cel9A expression strains were established through transformation of the expression plasmids to competent *E. coli* BL-21 Star^{TM} (DE3) cells (Invitrogen, Carlsbad, USA) according to the supplier's protocol. The transformed cells were incubated in LB medium at 37˚C for 1 hour prior to plating on LB agar medium with 50 μg/ml kanamycin, followed by overnight cultivation at the same temperature. The resulting clones were transferred to 500 ml of Terrific Broth (TB) medium with 50 μg/ml kanamycin and cultivated for 24 hours in a Harbinger system (Harbinger Biotechnology & Engineering, Markham, Canada) at 23˚C. The cultures were then induced by adding ITPG to a final concentration of 1 mM, and incubated for another 24h at 23 ^{o}C. The cells were harvested by centrifugation at 5000 x g for 15 minutes at 4˚C, using a Beckman Coulter centrifuge (Brea, CA, USA) and resuspended in 50 ml 50 mM Tris-HCl buffer pH 8.0 containing 500 mM NaCl and 5 mM imidazole. The cell suspensions were subjected to sonication on ice using a Vibracell sonicator (Sonics & Materials Inc., Newtown, Connecticut, USA) with 5 seconds on/off pulses for 3 minutes at 30% amplitude. The debris was removed by centrifugation at 15,000 x *g* for 15 minutes at 4˚C and the supernatant was filtered through a 0.22 μm syringe filter (Sarstedt, Nümbrecht, Germany), yielding sterile cell-free extracts, which were stored at 4˚C prior to enzyme purification.

AMOR_GH9A and *Tf*Cel9A proteins were isolated from cell-free extracts using metal affinity chromatography on a Ni^{2+} affinity HisTrap^{TM} HP 5 ml column (GE Healthcare, Chicago, USA). The enzymes were eluted with a linear gradient of imidazole (5–500 mM) in 50 mM Tris-HCl buffer, pH 8.0, containing 500 mM NaCl. Chromatography fractions were analyzed by SDS-PAGE (Bio-Rad, Hercules, California, USA). Fractions containing purified enzymes

were pooled and concentrated using 10,000 MWCO Vivaspin ultrafiltration tubes (Sartorius, Göttingen, Germany), with concomitant buffer exchange to 50 mM Tris-HCl, pH 8.0, containing 200 mM NaCl. The enzyme concentrations were determined by measuring optical absorbance at 280 nm with a Biophotometer UV-VIS spectrophotometer (Eppendorf, Hamburg, Germany), using theoretical extinction coefficients (web.expasy.org/protparam). The protein stock solutions were stored at 4˚C.

Optimal operating conditions

The temperature optima of the enzymes were assessed by incubation of 1μM AMOR_GH9A or 2 μM *Tf*Cel9A with 1% (w/v) carboxymethyl cellulose (CMC) for 6 minutes at temperatures ranging from 20˚C to 100˚C in citrate-phosphate buffer pH 5.7 (AMOR_GH9A) or pH 6.2 (*Tf*Cel9A). The pH optima were determined by carrying out the same reactions in various citrate-phosphate (pH 3.0–7.6) and glycine-NaOH buffers (pH 9.3–10.8) at 98˚C (AMOR_GH9A) or 65˚C (*Tf*Cel9A). The pH of the buffer solutions was set at room temperature. The experiments were conducted in a thermomixer (Eppendorf, Hamburg, Germany) at 600 rpm. The cellulase activity was determined by measuring the release of reducing sugars using the 3,5-dinitrosalicylic acid (DNS) reagent [28] and glucose as a standard.

The effect of salt on the performance of AMOR_GH9A was determined by incubation of 1 μM enzyme and 1% (w/v) CMC in citrate-phosphate buffer pH 5.7 at 98˚C with 0, 100, 500, 1000 or 2000 mM NaCl. Product formation was analyzed using the DNS assay, as described above.

Thermal stability

The thermal stability of AMOR_GH9A and *Tf*Cel9A was assessed by measuring the residual activity of the enzymes on 1% (w/v) CMC after up to 24 hours of pre-incubation in citrate-phosphate buffer pH 5.7 (AMOR_GH9A) or pH 6.2 (*Tf*Cel9A). The pre-incubation was performed at 98, 90, 85, and 80˚C (AMOR_GH9A) or at 65, 60, 55 and 50˚C (*Tf*Cel9A). The reactions with CMC were carried out for 6 minutes at 98˚C (AMOR_GH9A) or 65˚C (*Tf*Cel9A) and product formation was quantified with the DNS assay as described above.

Apparent melting temperature

Differential scanning calorimetry (Nano-Differential Scanning Calorimeter III, Calorimetry Sciences Corporation, Lindon, USA) was used to determine the apparent melting temperatures of AMOR_GH9A and *Tf*Cel9A. The protein solutions were desalted using MiniTrap^{tm} G-25 gel filtration columns and citrate-phosphate running buffers with pH 5.7 (AMOR_GH9A) or pH 6.2 (*Tf*Cel9A). These running buffers were utilized as reference samples in the subsequent calorimetry experiments. The protein solutions (final protein concentration approximately 0.5 mg/ml) and the reference solutions were filtered through a 0.22 μm syringe filter (Sarstedt, Nümbrecht, Germany) and degassed for 5 minutes using a ThermoVac system (GE Healthcare, Chicago, IL, USA) prior to data collection. The calorimetry was carried out in a pressurized chamber (4 atm) at 20–130˚C temperature range with 1˚C/min scan rate. The data were processed using the NanoAnalyze software provided by TA Instruments (New Castle, DE, USA). The buffer baselines were subtracted from the enzyme melting curves.

Substrate specificity

Avicel PH-101 (Sigma-Aldrich, St. Louis, MO, USA) was selected as a model crystalline substrate in this study. Phosphoric-acid swollen cellulose (PASC) was prepared from Avicel as

described in [29]. Beechwood xylan and konjac glucomannan were purchased from Megazyme (Wicklow, Ireland) and prepared according to the supplier protocol. The reactions were carried out in citrate-phosphate buffer pH 5.7 at 85˚C (AMOR_GH9A) or in citrate-phosphate buffer pH 6.2 at 55˚C (*Tf*Cel9A) with 1 μM enzyme. The substrate concentrations were 1% (w/v) for Avicel and 0.5% (w/v) for the other substrates. Aliquots were taken at various time points and the reactions were stopped by addition of NaOH to 100 mM final concentration. The products were quantified using the DNS assay and glucose standards.

Product analysis by HPAEC-PAD

Degradation products from cellulose and xylan were analyzed using high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD). The cellooligosaccharides were separated using a Dionex ICS3000 system (Thermo Scientific, San Jose, CA, USA) equipped with a CarboPac PA1 2 × 250 mm analytical column. A stepwise gradient with an increasing amount of eluent B (eluent B is 0.1 M NaOH and 1 M NaOAc; eluent A is 0.1 M NaOH) was applied starting right after sample injection, as follows: 0–10% B over 10 min, 10–30% B over 25 min, 30–100% B over 5 min, 100–0% B over 1 min, 0% B over 9 min. Data analysis was performed using Chromeleon 7.0 software. Cellooligosaccharide standards with a degree of polymerization of one to five (DP1—DP5) and xylo-oligosaccharide standards with a degree of polymerization of one to six (DP1 –DP6) were purchased from Megazyme (Wicklow, Ireland) and used to identify the products.

Product analysis by MALDI-TOF MS

The products of xylan degradation were identified using a matrix-assisted laser desorption/ionization time-of-flight (MALDI-TOF) UltrafleXtreme mass spectrometer (Bruker Daltonics GmbH, Bremen, Germany) equipped with a Nitrogen 337-nm laser. 1 μl of reaction mixture was added to 2 μl of 9 mg/ml 2,5-dihydrooxybenzoic acid (DHB) solution on a MTP 384 ground steel target plate (Bruker Daltonics). After air-drying, spectral data was acquired and processed using Bruker flexControl and flexAnalysis software.

Results and discussion

Metagenomic data analysis

After sequencing and assembly of metagenomic data, a 2065 bp gene encoding a putative GH9 cellulase was identified using the dbCAN annotation tool. The candidate enzyme was named AMOR_GH9A. According to the Pfam domain classification server [30], AMOR_GH9A is a 688 residue protein (Fig 1) comprising a signal peptide, a catalytic GH9 domain and a CBM3 cellulose binding module.

BLAST searches identified a hypothetical endoglucanase from the thermophilic marine bacterium *Ardenticatena maritima* as having the highest degree of sequence similarity to AMOR_GH9A (77.8% identity between catalytic domains, sequence ID: WP_060687350.1). *Ardenticatena maritima* was isolated from a hydrothermal field sediment and can grow at temperatures as high as 75˚ C [31]. The closest characterized homolog of AMOR_GH9A is *Tf*Cel9A from the cellulolytic model actinomycete *Thermobifida fusca (*67.7% identity between catalytic domains). *Tf*Cel9A is a well-known thermostable GH9 cellulase with a complex domain architecture and an endo-processive mode of action [32, 33]. *Tf*Cel9A consists of a signal peptide, an N-terminal catalytic domain, two cellulose binding modules and a fibronectin type III domain (Fig 1). In this study, *Tf*Cel9A was selected as the reference enzyme to assess AMOR_GH9A thermostability and substrate specificity via direct comparison.

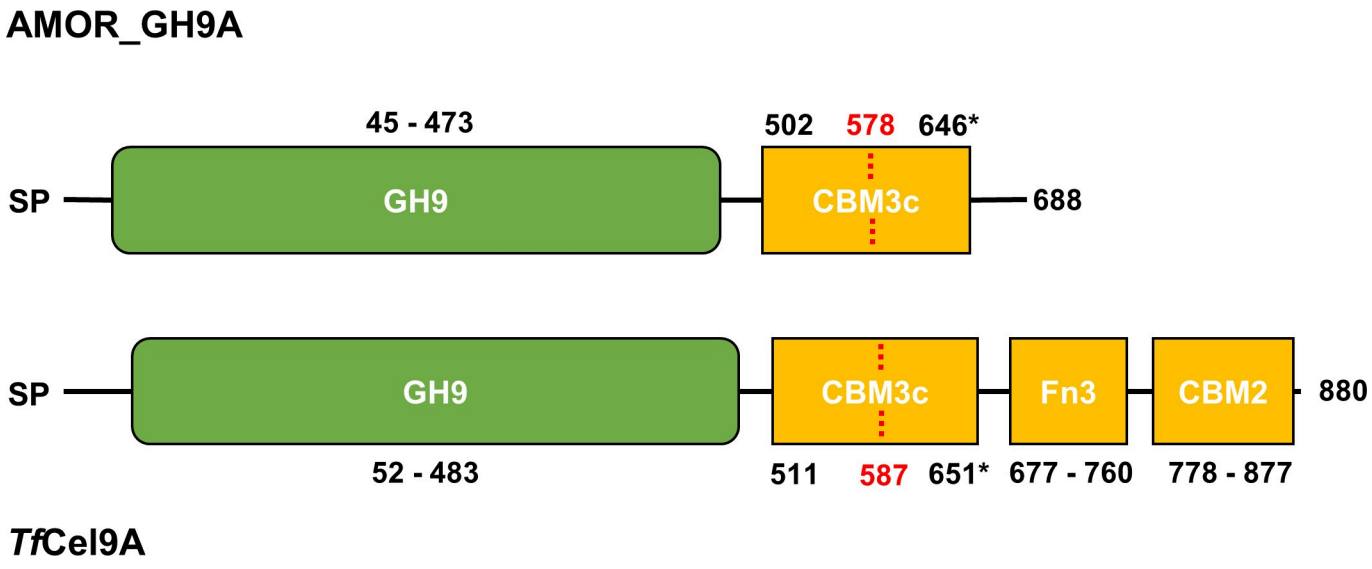

Fig 1. Domain architecture of AMOR_GH9A and *Tf*Cel9A cellulases. SP, signal peptide; GH9, catalytic domain; CBM2/CBM3c, family 2/family 3c cellulose-binding module; Fn3, type III fibronectin domain. The dotted line and number in red colour indicate the boundaries of the CBM3c domain according to the erroneous Pfam prediction. The domain coordinates marked with "*" were derived from sequence comparisons and the crystal structure of *Tf*Cel9A. See text for details.

https://doi.org/10.1371/journal.pone.0222216.g001

Of note, both Pfam and dbCAN did not predict the domain boundaries of the CBM3 in AMOR_GH9A correctly, recognizing only the N-terminal half of this domain (Fig 1). Sequence alignment with *Tf*Cel9A (Fig 2A) and the X-ray structure of *Tf*Cel9A (Fig 2B; [34]) clearly show that the CBM3 comprises approximately 140 residues, as one would expect [35]. Domain annotation of *Tf*Cel9A with Pfam gave a similarly incorrect result. The CBM3 domain of AMOR_GH9A lacks the so-called "planar strip" (a conserved array of mostly aromatic amino acids involved in binding to crystalline substrates) and belongs to subfamily CBM3c [32, 35].

Protein production

Genes encoding AMOR_GH9A and *Tf*Cel9A were codon optimized for expression in *E. coli*, synthesized and then cloned into pNIC-CH vectors using ligation-independent cloning. The enzymes were produced (S1 Fig) in *E. coli* BL-21 Star^{TM} (DE3), without signal peptides and with a C-terminal affinity tag for purification by metal affinity chromatography. AMOR_-GH9A and *Tf*Cel9A were produced in soluble form and the final yield was approximately 90 mg of purified protein per liter of *E. coli* culture for both enzymes.

Optimal operating conditions

The optimal operating conditions of AMOR_GH9A and *Tf*Cel9A were assessed using carboxymethyl cellulose (CMC) as a model substrate (Fig 3). Of note, despite the large amount of data published on *Tf*Cel9A, the pH and temperature dependency of the full-length enzyme have not been addressed in detail before.

The results indicate that *Tf*Cel9A performs best at pH 6.2, 65˚C. The pH optimum of AMOR_GH9A is approximately 5.7, while the temperature optimum is 100˚C or higher. During the experiment, the highest activity was observed when the boiling point was reached and a further increase in incubation temperature was not possible for practical reasons. After the initial temperature optima assays, all the subsequent AMOR_GH9A reactions with CMC were

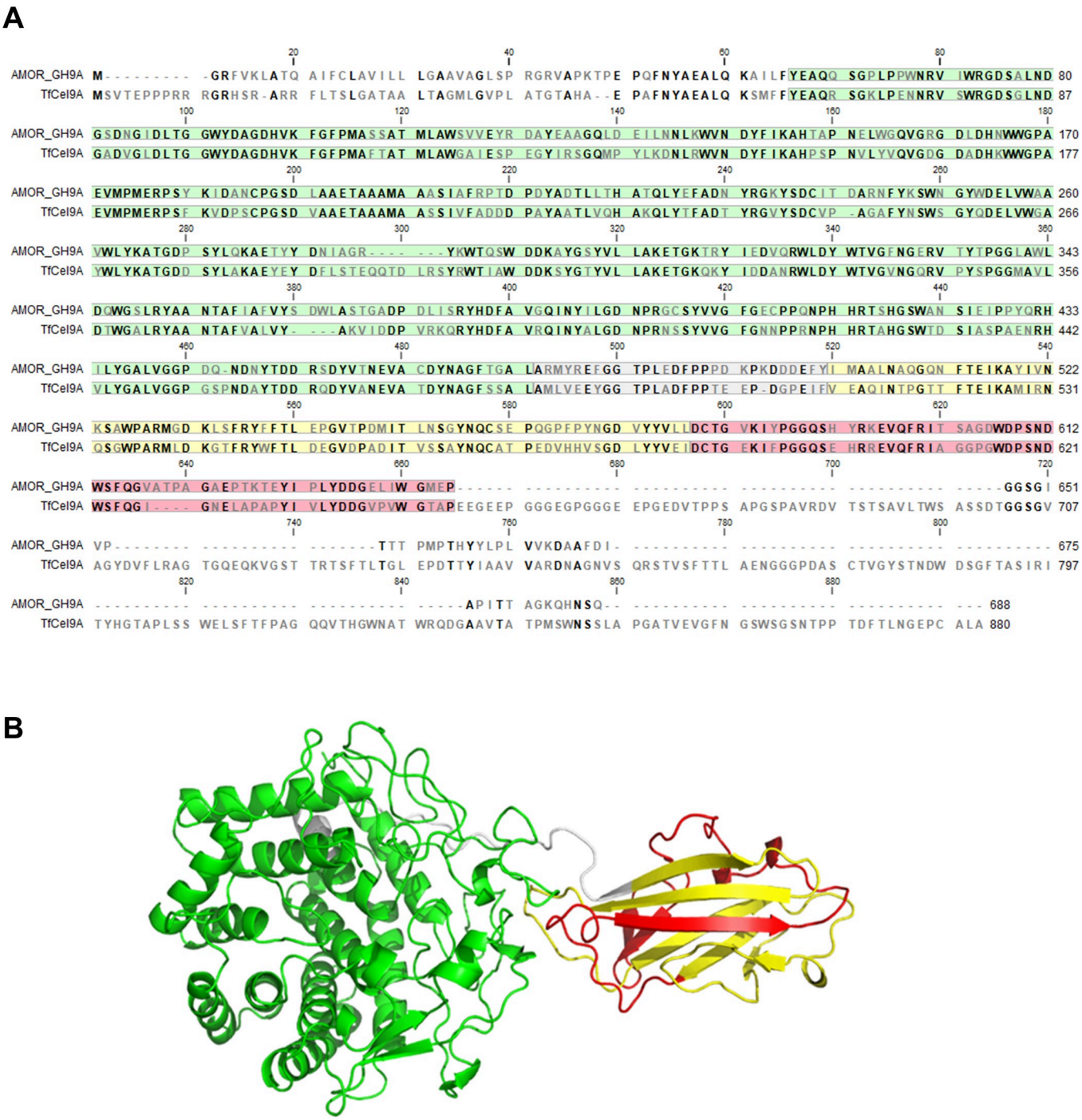

Fig 2. AMOR_GH9A compared to *Tf*Cel9A. Panel (A) shows a sequence alignment of AMOR_GH9A and *Tf*Cel9A, whereas panel B shows the crystal structure of a fragment of *Tf*Cel9A comprising the catalytic domain and the CBM3c domain (residues 47–651) that was obtained by limited proteolysis [34]. The protein regions are marked with colour as follows: green, catalytic domain; grey, linker; yellow, N-terminal part of the CBM3c that is recognized by Pfam; red, C-terminal part of the CBM3c that is not recognized by Pfam (see text for details). The conserved amino acid residues are indicated by bold font. Note that AMOR_GH9A is shorter than *Tf*Cel9A and that the alignment of the C-terminal "tail" of AMOR_GH9A (residues 665–688) with the much longer C-terminal part of *Tf*Cel9A is inaccurate and does not necessarily indicate structural or functional similarities.

https://doi.org/10.1371/journal.pone.0222216.g002

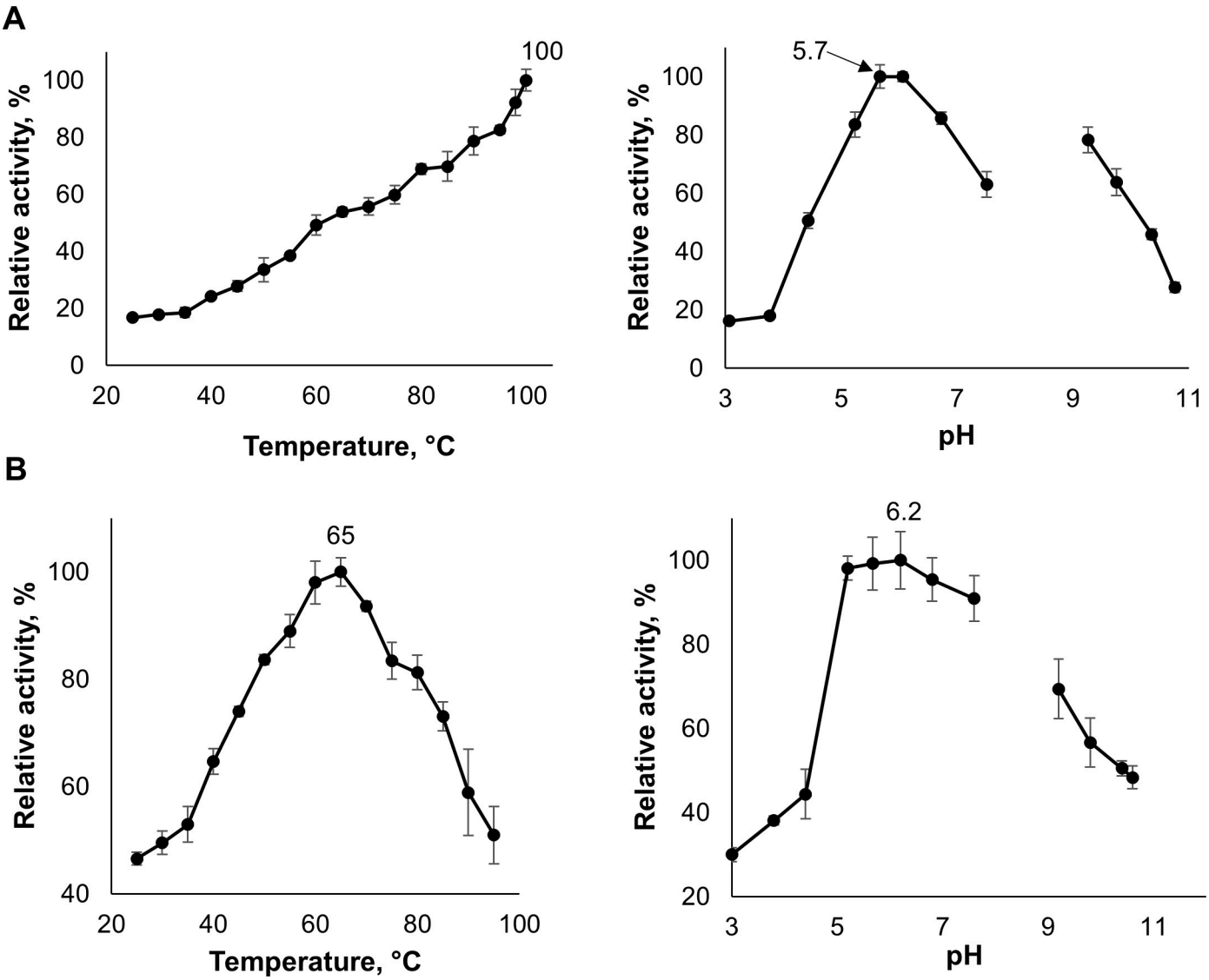

Fig 3. Optimal conditions for AMOR_GH9A and *Tf*Cel9A. Panels (A) and (B) show the temperature and pH optima of AMOR_GH9A (A) and *Tf*Cel9A (B). Temperature optima were determined at pH 5.7 and 6.2 and pH optima were determined at 98 °C and 65 °C, for AMOR_GH9A and *Tf*Cel9A, respectively. Enzyme activities were assessed by measuring product formation from CMC after 6 min reactions. The maximum level of product formation was set to 100% and the temperatures or pH values at which this maximum level was obtained are shown in the graph. Note that two different buffers were used in the determination of the pH optimum, citrate-phosphate, covering pH 3.0–7.6 and glycine-NaOH, covering pH 9.2–10.7. The pH values displayed in the figure were measured at room temperature. While the temperature dependency of the pH of the citrate-phosphate buffer is close to negligible [36], the temperature dependency of the pH of the glycine-NaOH buffer is considerable ($dpKa_2/dt$ = -0.025, [37]). Thus, considering the assay temperatures of 98 °C and 65 °C, for this buffer, the actual pH values were about 1.8 and 1 units lower than shown in the Figure for AMOR_GH9A and *Tf*Cel9A, respectively. Accordingly, the apparent gaps in pH-dependency curves are in fact nonexistent. Error bars indicate standard deviations between triplicates.

https://doi.org/10.1371/journal.pone.0222216.g003

carried out at 98˚C. Note that while the pH of the assay buffers was set at the room temperature, the reported pH optima for both enzymes are hardly affected by the experimental conditions, since the temperature-dependency of the citrate-phosphate buffer is extremely low [36].

The pronounced difference between the temperature optima of AMOR_GH9A and *Tf*Cel9A makes sense when considering the origin of the enzymes. *Thermobifida fusca* is a soil bacterium typically found in decomposing organic matter (e.g. compost or rotting hay), which can heat up to approximately 70˚C due to exothermic reactions [38]. In comparison, the temperatures at the Jan Mayen vent field can rise up to 260˚C [17] with steep thermal gradients.

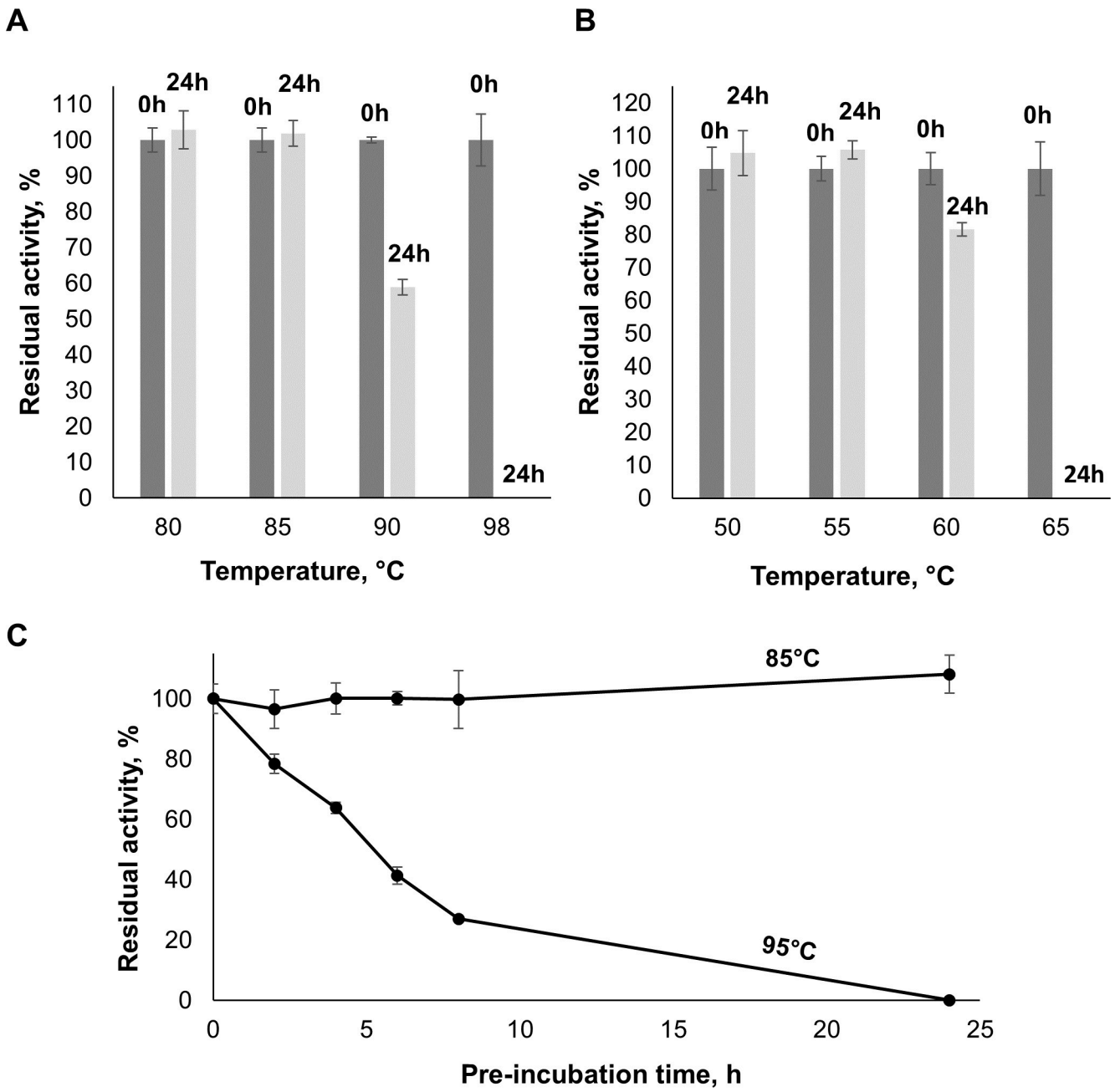

Fig 4. Thermal stability of AMOR_GH9A. Panels (A), (B) and (C) show the residual activity of AMOR_GH9A (A, C) and *Tf*Cel9A (B) in 6 min reactions with CMC after pre-incubation at various temperatures for various time periods (0 or 24 h in panels A and B; multiple time points in panel C). The pre-incubations were done in citrate-phosphate buffer, pH 5.7, or in citrate-phosphate buffer, pH 6.2, for AMOR_GH9A and *Tf*Cel9A, respectively. The activity assays were done at 98 °C or 65 °C and at pH 5.7 or 6.2 for AMOR_GH9A and *Tf*Cel9A, respectively. Error bars indicate standard deviations between triplicates.

https://doi.org/10.1371/journal.pone.0222216.g004

In a recent review, Escuder-Rodríguez et al. [8] summarized temperature optima of 185 thermophilic cellulases (64 endoglucanases, 121 exoglucanases) of bacterial and fungal origin. Only six of the listed thermophilic cellulases (3.2%) across all the GH families possess a temperature optimum similar to the optimum of AMOR_GH9A (i.e., ≥100˚C). There are six GH9 enzymes in the dataset and AMOR_GH9A has a higher optimal temperature than all of

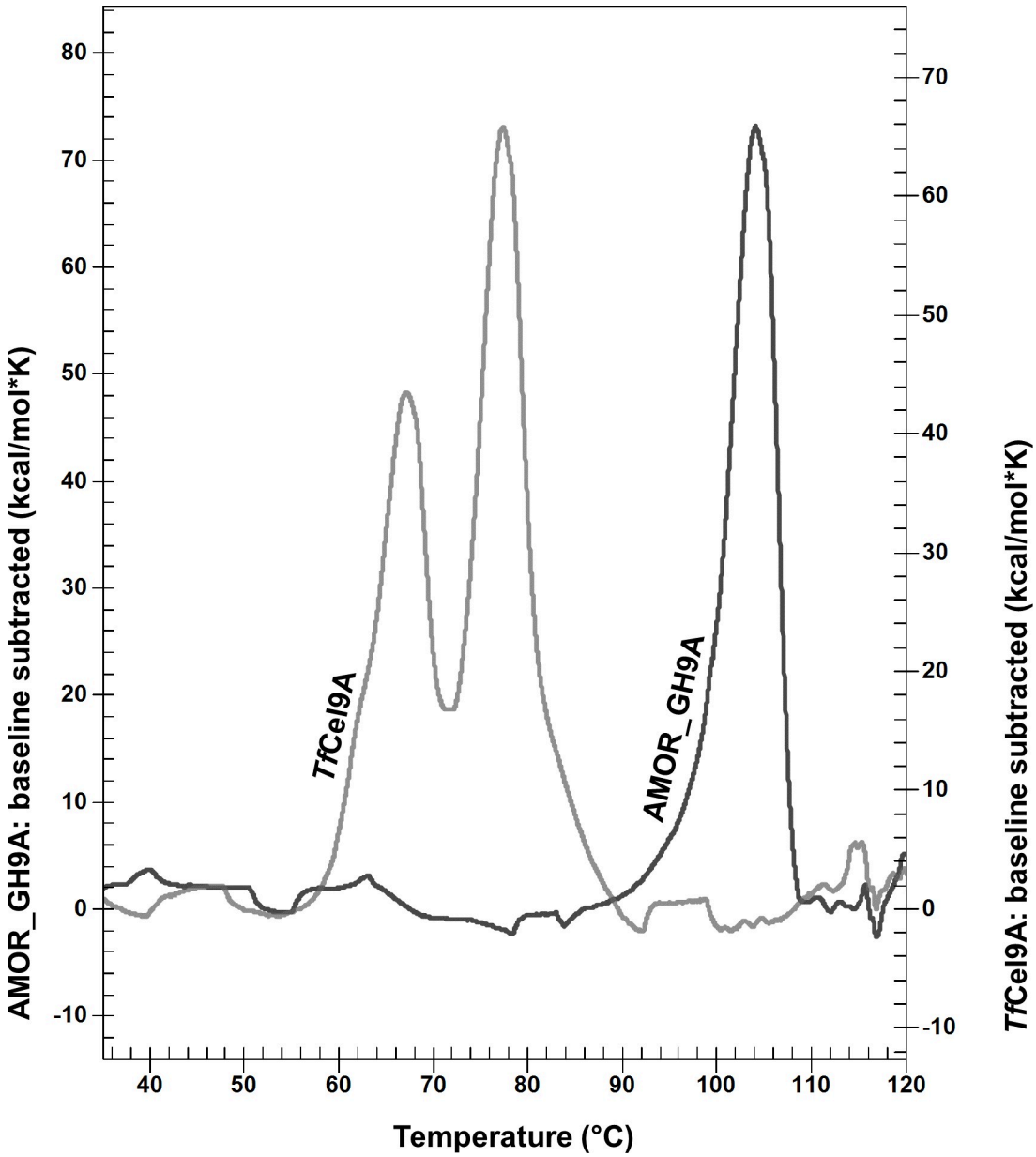

Fig 5. Differential scanning calorimetry (DSC) melting curves for AMOR_GH9A and *Tf*Cel9A. The molar heat capacity of the enzyme solutions is plotted as a function of the temperature. The enzymes were dissolved at 0.5 mg/ml concentration in citrate-phosphate buffer pH 5.7 or citrate-phosphate buffer pH 6.2 for AMOR_GH9A and *Tf*Cel9A, respectively. Before plotting, baseline curves (i.e., buffer only) were subtracted from the protein curves. The heating rate was 1˚C per minute. In both cases, the unfolding was irreversible.

https://doi.org/10.1371/journal.pone.0222216.g005

these. Of note, the majority of GH9 cellulases reported in the review (five out of six proteins) seem to be only moderately thermophilic since their temperature optima do not exceed 70˚C. The most thermophilic of the six GH9s is CelA cellulase from *Caldicellulosiruptor bescii* [39] with a reported temperature optimum of 95˚C (note that this is a multimodular enzyme, also containing a GH48 domain). Although these comparisons have limitations (e.g. due to variation in the conditions used), it is clear that AMOR_GH9A belongs to the most thermophilic cellulases described so far.

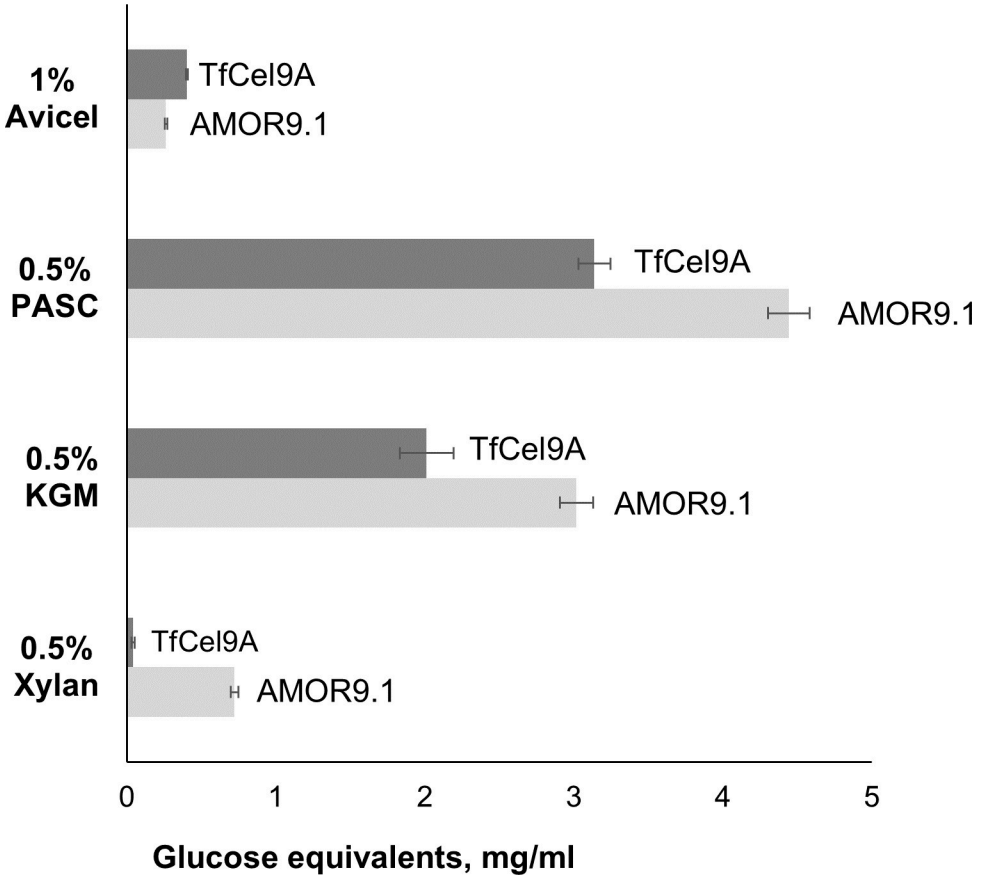

Fig 6. Substrate specificity. The chart shows product release by AMOR_GH9A and *Tf*Cel9A from various substrates after 24h incubation at optimal pH and a temperature not likely to lead to enzyme inactivation. Reaction conditions for AMOR_GH9A: 1 μM enzyme, 85˚C, pH 5.7; for *Tf*Cel9A: 1 μM enzyme, 55˚C, pH 6.2. The substrate concentration was 1% (w/v) in case of Avicel and 0.5% (w/v) in case of all the other substrates. KGM, konjac glucomannan; xylan, beechwood xylan. Error bars indicate standard deviations between triplicates.

https://doi.org/10.1371/journal.pone.0222216.g006

CMC assays at pH 5.7 and 98 °C showed that the activity of AMOR_GH9A was almost insensitive to salt. The highest activity was obtained in reactions without added NaCl. The increase of salt concentration gradually reduced activity but even at 2 M NaCl, the remaining activity was still approximately 85% of the base level (0 M NaCl) (S2 Fig). These findings are in a strong contrast with the results obtained for two other enzymes (a thermostable xylanase, AMOR_GH10A, and a thermostable alginate lyase, AMOR_PL7A) recently discovered using the same metagenomic dataset [21, 22]. Unlike AMOR_GH9A, AMOR_GH10A is a salt-dependent enzyme showing low activity at 0 mM NaCl. AMOR_PL7A is less responsive to salt, but requires the addition of ~430 mM NaCl to the buffer to manifest full activity.

Thermal stability

The thermal stability of AMOR_GH9A and *Tf*Cel9A was assessed and compared by measuring residual activity on CMC after pre-incubation of the enzymes in citrate-phosphate buffer at optimal pH and various temperatures (Fig 4). AMOR_GH9A and *Tf*Cel9A retained 100% activity after 24 hours of pre-incubation at 85˚C and 55˚C, respectively. At higher temperatures, the proteins became unstable. It is worth noting that AMOR_GH9A remains active for quite a long time under extreme conditions. For example, our results indicate that the enzyme

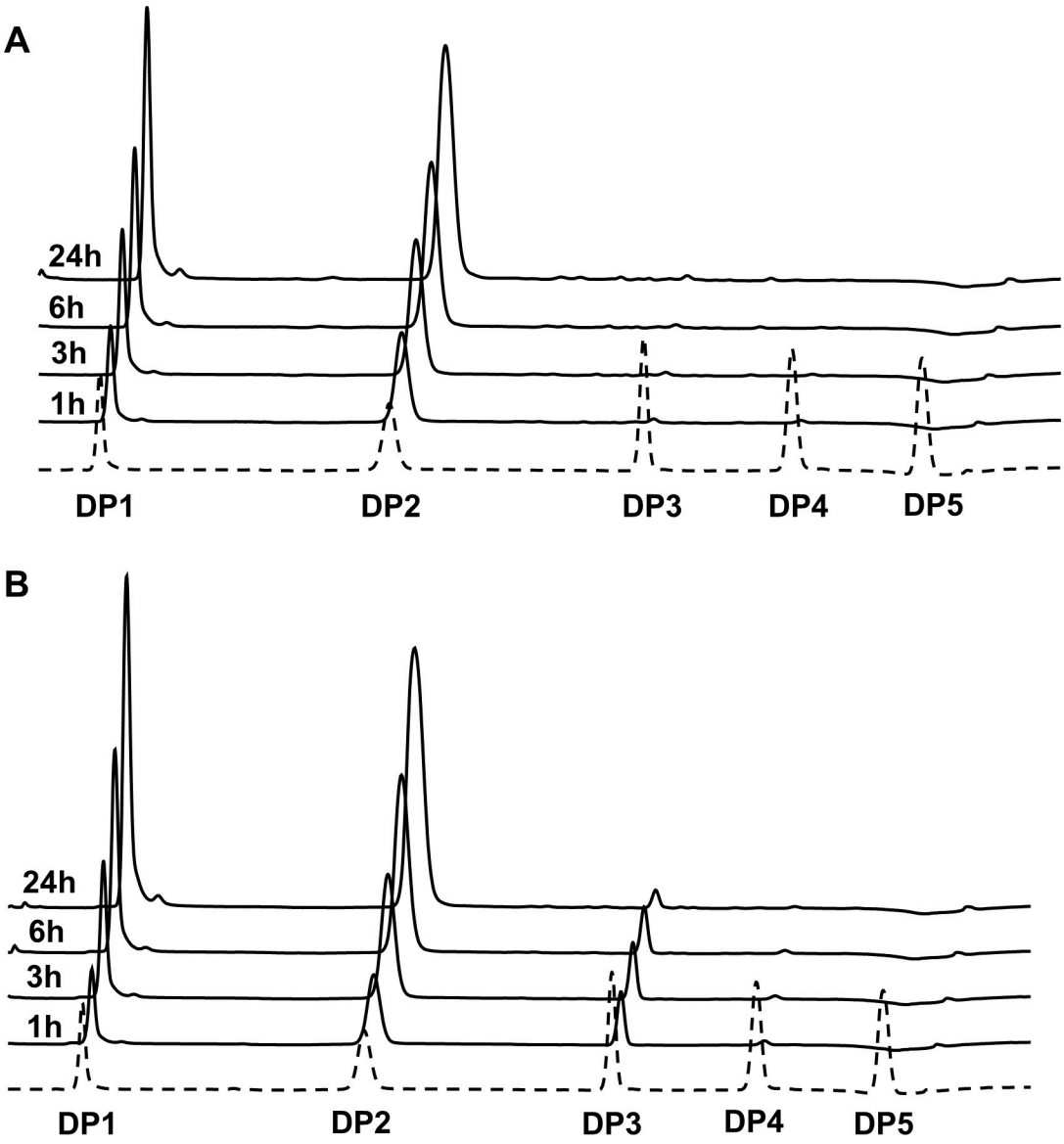

Fig 7. Products generated from Avicel. Panels A and B show cello-oligosaccharides (HPAEC-PAD chromatograms) generated over time from 1% (w/v) Avicel by AMOR_GH9A (A) and *Tf*Cel9A (B). Reaction conditions for AMOR_GH9A: 1 μM enzyme, 85˚C, pH 5.7; for TfCel9A: 1 μM enzyme, 55˚C, pH 6.2. The dotted line is a chromatogram of a standard sample containing cello-oligosaccharides in the DP1 to DP5 range, each at 100 μM concentration.

https://doi.org/10.1371/journal.pone.0222216.g007

retains 64% of its activity after 4 hours of pre-incubation at 95˚C (Fig 4C). To the best of our knowledge, such degree of thermostability is unparalleled among GH9 cellulases reported so far. Melting curves for AMOR_GH9A and *Tf*Cel9A were obtained using differential scanning calorimetry (DSC). Both AMOR_GH9A and *Tf*Cel9A displayed irreversible unfolding. The melting curve for AMOR_GH9A showed a single peak at approximately 105˚C while *Tf*Cel9A demonstrated a two-phase transition with peaks around 65˚C and 78˚C (Fig 5). A bi-phasic nature of *Tf*Cel9A unfolding is not surprising, considering the complex domain structure of the enzyme (Fig 1). Interestingly, the first *Tf*Cel9A unfolding phase happened at the temperature where the enzyme starts losing its activity (~65˚C; Figs 3B & 4B). It is thus conceivable that this first phase corresponds to unfolding of the catalytic domain.

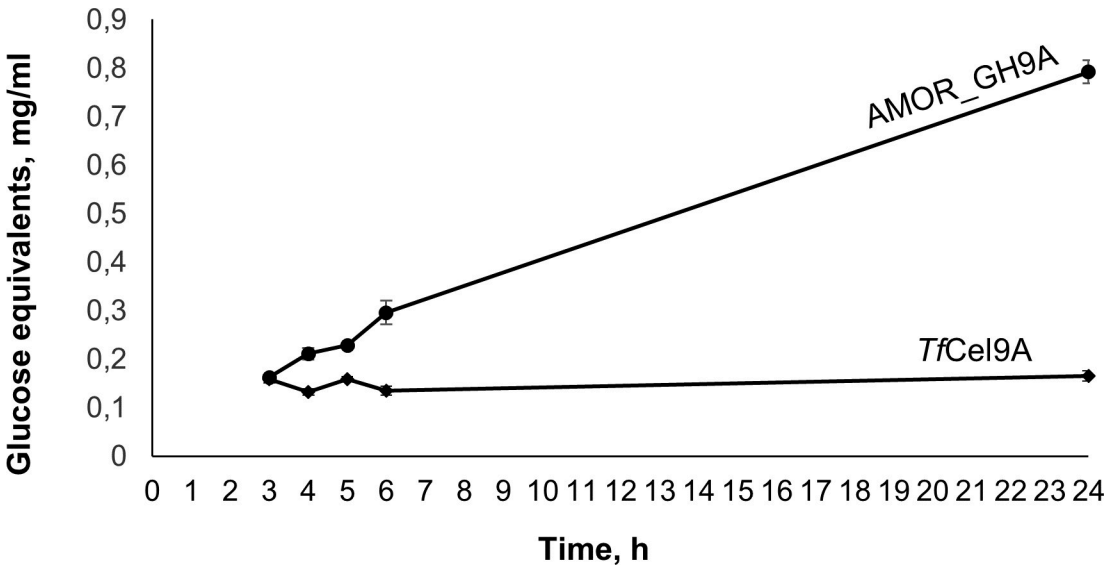

Fig 8. AMOR_GH9A activity on xylan. The chart shows product release by AMOR_GH9A and *Tf*Cel9A from 0.5% (w/v) beechwood xylan at various timepoints. Reaction conditions for AMOR_GH9A: 1 μM enzyme, 85˚C, pH 5.7; for TfCel9A: 1 μM enzyme, 55˚C, pH 6.2. Error bars indicate standard deviations between triplicates.

https://doi.org/10.1371/journal.pone.0222216.g008

Substrate specificity

Studies of substrate specificity showed that both AMOR_GH9A and *Tf*Cel9A hydrolyze PASC and Avicel (Fig 6). AMOR_GH9A outperformed *Tf*Cel9A in reactions with amorphous PASC, releasing approximately 1.4 times more glucose equivalents, whereas *Tf*Cel9A showed the highest activity on Avicel. These differences may in part be due to the different architecture and CBM content of the two enzymes (Fig 1). In particular, the only cellulose binding domain of AMOR_GH9A belongs to the subfamily CBM3c, which has relatively weak affinity towards crystalline substrates [32]. In case of *Tf*Cel9A binding to Avicel is likely to be enhanced by the additional C-terminal CBM2 domain. Of note, the performance of both cellulases on Avicel is relatively poor given the enzyme load of 100 nmol per gram of substrate. The reducing end concentration obtained after 24h incubation with *Tf*Cel9A (Fig 6) corresponds to approximately 5% substrate solubilization. Previous studies of the degradation of Avicel by *Tf*Cel9A gave similar results [40].

The HPAEC-PAD analysis of Avicel depolymerization products revealed some interesting features. Firstly, during the initial phase of the reaction, *Tf*Cel9A generated a significant amount of cellotriose and this trisaccharide was still detectable after 24 hours (Fig 7). The fact that AMOR_GH9A only produced disaccharides and monosaccharides suggests that the two enzymes have different substrate-binding abilities, with *Tf*Cel9A being less capable of cleaving short cello-oligosaccharides such as cellotriose. Indeed, the low ability of *Tf*Cel9A to cleave short oligosaccharides has been observed previously [41]. A second interesting feature is the high level of monosaccharides that are formed. Although disaccharide/monosaccharide ratios need to be used with caution [42], they give an indication of enzyme processivity and the relatively low disaccharide/monosaccharide ratios observed here indicate that the two GH9s are not particularly processive. While some degree of processivity cannot be excluded [33, 41], it may not be a dominating feature of these enzymes.

The vast majority of characterized GH9 enzymes are cellulases [43]. However, some of these cellulose-targeting enzymes are known to display side activities towards hemicellulosic

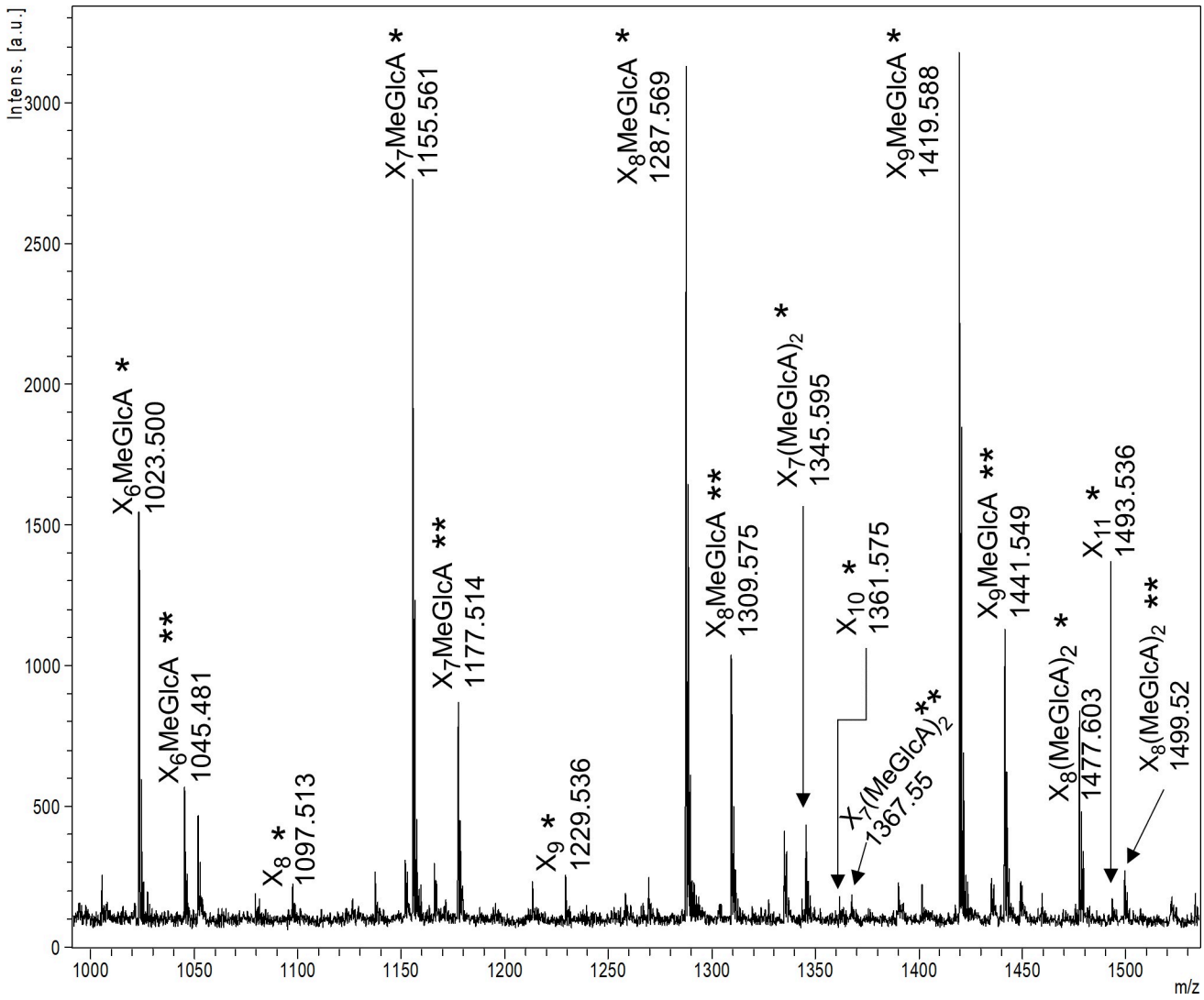

Fig 9. MALDI-TOF MS analysis of products generated from beechwood xylan by AMOR_GH9A. The picture shows a part of a MALDI-TOF spectrum with signals from non-substituted and substituted xylo-oligosaccharides released from 0.5% (w/v) xylan after a 24 hour incubation. Reaction conditions: 1 μM enzyme, 85˚C, pH 5.7. X, xylose; MeGlcA, 4-O-methylglucuronic acid; *, sodium adduct; **, sodium salt of a sodium adduct. None of the labeled peaks were observed in the negative control (i.e. a reaction without added enzyme).

https://doi.org/10.1371/journal.pone.0222216.g009

substrates including glucomannan [44], xylan [45] and xyloglucan [46]. Indeed, we found that both AMOR_GH9A and *Tf*Cel9A are able to hydrolyze konjac glucomannan (Fig 6). Interestingly, AMOR_GH9A showed a clear activity on xylan, in contrast to *Tf*Cel9A (Figs 6 and 8). The ability to hydrolyze xylan is a desirable property considering the high xylan content of several industrially relevant types of plant biomass [47]. MALDI-TOF MS analysis of products released from beechwood xylan showed that AMOR_GH9A generates a mixture of non-substituted xylo-oligosaccharides and xylo-oligosaccharides substituted with methylated glucuronic acid (Fig 9). Chromatographic analysis of products generated from beechwood xylan confirmed that AMOR_GH9A releases a wide variety of substituted and non-substituted xylo-oligosaccharides, including xylobiose and trace amounts of xylose (Fig 10).

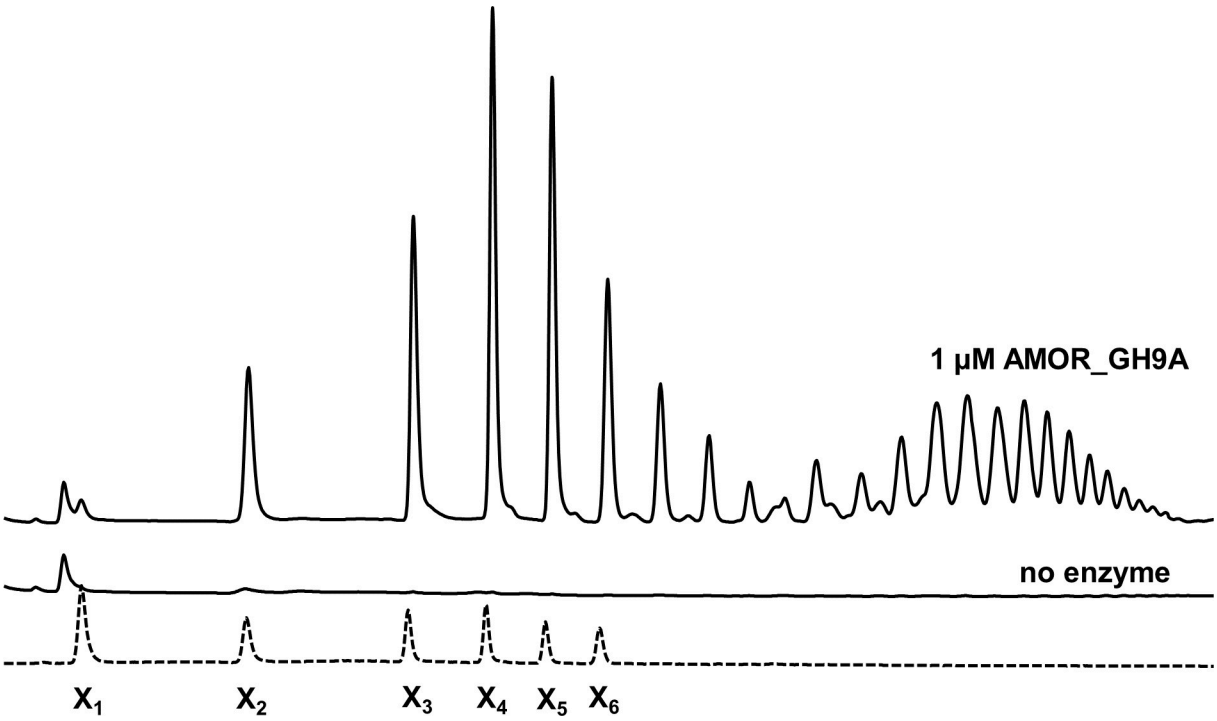

Fig 10. HPAEC-PAD analysis of products generated from beechwood xylan by AMOR_GH9A. The figure shows HPAEC-PAD chromatograms of xylo-oligosaccharides generated after incubation of 0.5% (w/v) beechwood xylan by 1 μM AMOR_GH9A at 85˚C, pH 5.7, for 24h. The dotted line is a chromatogram of a standard sample containing xylo-oligosaccharides in the DP1 to DP6 range, each at 0.01 mg/ml concentration. The peaks in the right half of the chromatogram of the reaction sample represent xylo-oligosaccharides with high degree of polymerization (including substituted xylo-oligosaccharides). The "no enzyme" line is a chromatogram of a negative control sample.

https://doi.org/10.1371/journal.pone.0222216.g010

Concluding remarks

In silico mining of a metagenomic dataset originating from the Jan Mayen hydrothermal vent field led to the identification of the novel GH9 cellulase named AMOR_GH9, which is among the most thermostable and thermoactive cellulases ever described. The enzyme comprises an N-terminal catalytic domain followed by a CBM3 cellulose binding module and is easy to produce in *E. coli*. AMOR_GH9A possesses a remarkably high temperature optimum (≥100˚C) and retains 64% of its activity after 4 hours of incubation at 95˚C. Direct functional comparison with its closest characterized homolog (*Tf*Cel9A from the model thermophilic bacteria *Thermobifida fusca)* revealed that AMOR_GH9A possesses broader substrate specificity and higher activity on soluble and amorphous substrates (PASC, KGM). Thus, the novel GH9 cellulase demonstrates a set of industrially relevant properties and has the potential to become part of the enzymatic toolbox for biomass conversion.

Supporting information

S1 Table. Primers used for amplification and ligation-independent cloning of the genes encoding AMOR_GH9A and *Tf*Cel9A. Vector complimentary sequences are underlined. (PDF)

S1 Fig. SDS-PAGE of AMOR_GH9A and *Tf*Cel9A cellulases after purification. The predicted MW of AMOR_GH9A and TfCel9A is 73.2 kDa and 91.4 kDa respectively. The sample loading was 5μg for each protein. Note that there is some heterogeneity in the AMOR_GH9A band. We were not able to remove this heterogeneity by additional purification steps or

extended boiling of the samples. Note that AMOR_GH9A and *Tf*Cel9A were produced and purified exactly in the same way.
(TIF)

S2 Fig. The effect of salt on the activity of AMOR_GH9A. The graph shows product formation after a 6 min reaction with CMC at 98˚C in citrate-phosphate buffer, pH 5.7, supplied with NaCl at different concentrations. Error bars indicate standard deviations between triplicates.
(TIF)

Acknowledgments

We thank Prof. Rolf-Birger Pedersen at the University of Bergen for leading the Centre for Geobiology research cruise to the Arctic Mid-Ocean Ridge vent fields and Marianne Slang-Jensen at the Norwegian University of Life Sciences for guidance on differential scanning calorimetry.

Author Contributions

Conceptualization: Anton A. Stepnov, Lasse Fredriksen, Ida H. Steen, Runar Stokke, Vincent G. H. Eijsink.

Data curation: Anton A. Stepnov, Lasse Fredriksen, Runar Stokke.

Formal analysis: Anton A. Stepnov, Lasse Fredriksen, Runar Stokke, Vincent G. H. Eijsink.

Funding acquisition: Vincent G. H. Eijsink.

Investigation: Anton A. Stepnov, Lasse Fredriksen.

Methodology: Anton A. Stepnov, Lasse Fredriksen, Ida H. Steen, Runar Stokke, Vincent G. H. Eijsink.

Project administration: Vincent G. H. Eijsink.

Resources: Lasse Fredriksen, Ida H. Steen, Runar Stokke.

Software: Runar Stokke.

Supervision: Vincent G. H. Eijsink.

Validation: Anton A. Stepnov, Vincent G. H. Eijsink.

Visualization: Anton A. Stepnov.

Writing – original draft: Anton A. Stepnov.

Writing – review & editing: Anton A. Stepnov, Runar Stokke, Vincent G. H. Eijsink.

References

1. Mariano M, El Kissi N, Dufresne A. Cellulose nanocrystals and related nanocomposites: review of some properties and challenges. J. Polym. Sci., Part B: Polym. Phys. 2014; 52(12): 791–806
2. Zhang Z, Donaldson AA, Ma X. Advancements and future directions in enzyme technology for biomass conversion. Biotechnol Adv. 2012; 30(4): 913–919 https://doi.org/10.1016/j.biotechadv.2012.01.020 PMID: 22306162
3. Himmel ME, Ding SY, Johnson DK, Adney WS, Nimlos MR, Brady JW. et al. Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science. 2007; 315(5813): 804–807 https://doi.org/10.1126/science.1137016 PMID: 17289988

4. Obeng EM, Adam SNN, Budiman C, Ongkudon CM, Maas R, Jose J. Lignocellulases: a review of emerging and developing enzymes, systems, and practices. Bioresources and Bioprocessing. 2017; 4:16
5. Bissaro B, Várnai A, Røhr AK, Eijsink VGH. Oxidoreductases and reactive oxygen species in conversion of lignocellulosic biomass. Microbiology and Molecular Biology Reviews. 2018; 82(4): e00029–18 https://doi.org/10.1128/MMBR.00029-18 PMID: 30257993
6. Tiwari R, Nain L, Labrou NE, Shukla P. Bioprospecting of functional cellulases from metagenome for second generation biofuel production: a review. Critical Reviews in Microbiology. 2018; 44(2): 244–257 https://doi.org/10.1080/1040841X.2017.1337713 PMID: 28609211
7. Viikari L, Alapuranen M, Puranen T, Vehmaanpera J, Siika-Aho M. Thermostable enzymes in lignocellulose hydrolysis. Adv Biochem Eng Biotechnol. 2007; 108: 121–145 https://doi.org/10.1007/10_2007_065 PMID: 17589813
8. Escuder-Rodríguez JJ, DeCastro ME, Cerdán ME, Rodríguez-Belmonte E, Becerra M, González-Siso MI. Cellulases from thermophiles found by metagenomics. Microorganisms. 2018; 6(3): 66
9. Patel AK, Singhania RR, Sim SJ, Pandey A. Thermostable cellulases: current status and perspectives. Bioresource Technology. 2019; 279: 385–392 https://doi.org/10.1016/j.biortech.2019.01.049 PMID: 30685132
10. Ferrer M, Martínez-Martínez M, Bargiela R, Streit WR, Golyshina OV, Golyshin PN. Estimating the success of enzyme bioprospecting through metagenomics: current status and future trends. Microb Biotechnol. 2016; 9(1): 22–34 https://doi.org/10.1111/1751-7915.12309 PMID: 26275154
11. Guazzaroni ME, Beloqui A, Vieites JM, Al-ramahi Y, Cortés NL, Ghazi A, Golyshin PN et al. Metagenomic mining of enzyme diversity. Handbook of Hydrocarbon and Lipid Microbiology. 2010; 2911–2927
12. Miroshnichenko ML, Bonch-Osmolovskaya EA. Recent developments in the thermophilic microbiology of deep-sea hydrothermal vents. Extremophiles. 2006; 10(2): 85–96 https://doi.org/10.1007/s00792-005-0489-5 PMID: 16418793
13. Tarasov VG, Gebruk AV, Mironov AN, Moskalev LI. Deep-sea and shallow-water hydrothermal vent communities: two different phenomena? Chemical Geology. 2005; 224: 5–39
14. Podosokorskaya OA, Kublanov IV, Reysenbach A.-L, Kolganova TV, Bonch-Osmolovskaya EA. *Thermosipho affectus* sp. nov., a thermophilic, anaerobic, cellulolytic bacterium isolated from a Mid-Atlantic Ridge hydrothermal vent. International Journal of Systematic and Evolutionary Microbiology. 2011; 61: 1160–1164 https://doi.org/10.1099/ijs.0.025197-0 PMID: 20562244
15. Wery N, Lesongeur F, Pignet P, Derennes V, Cambon-Bonavita M-A, Godfroy A et al. *Marinitoga camini* gen. nov., sp. nov., a rod-shaped bacterium belonging to the order Thermotogales, isolated from a deep-sea hydrothermal vent. International Journal of Systematic and Evolutionary Microbiology. 2001; 51: 495–504 https://doi.org/10.1099/00207713-51-2-495 PMID: 11321096
16. Stokke R, Dahle H, Roalkvam I, Wissuwa J, Daae FL, Tooming-Klunderud A et al. Functional interactions among filamentous Epsilonproteobacteria and Bacteroidetes in a deep-sea hydrothermal vent biofilm. Environmental Microbiology. 2015; 17(10): 4063–4077 https://doi.org/10.1111/1462-2920.12970 PMID: 26147346
17. Schander C, Rapp HT, Kongsrud JA, Bakken T, Berge J, Cochrane S et al. The fauna of hydrothermal vents on the Mohn Ridge (North Atlantic). Marine Biology Research. 2010; 6: 155–171
18. Jung ED, Lao G, Irwin D, Barr BK, Benjamin A, Wilson DB. DNA sequences and expression in Streptomyces lividans of an exoglucanase gene and an endoglucanase gene from *Thermomonospora fusca*. Applied and Environmental Microbiology. 1993; 59(9): 3032–3043 PMID: 8215374
19. Sjöde A, Frölander A, Lersch M, Rødsrud G. Lignocellulosic biomass conversion by sulphite pretreatment. 2013; patent EP2376642 B1
20. Rødsrud G, Lersch M, Sjöde A. History and future of world's most advanced biorefinery in operation. Biomass & Bioenergy. 2012; 46: 46–59
21. Fredriksen L, Stokke R, Jensen MS, Westereng B, Jameson J.-K., Steen IH et al. Discovery of a thermostable GH10 xylanase with broad substrate specificity from the Arctic Mid-Ocean Ridge vent system. Applied and Environmental Microbiology. 2019; 85(6); pii: e02970–18 https://doi.org/10.1128/AEM.02970-18 PMID: 30635385
22. Vuoristo KS, Fredriksen L, Oftebro M, Arntzen MO, Aarstad OA, Stokke R et al. Production, characterization, and application of an alginate lyase, AMOR_PL7A, from hot vents in the Arctic Mid-Ocean Ridge. J. Agric. Food Chem. 2019; 67(10): 2936–2945 https://doi.org/10.1021/acs.jafc.8b07190 PMID: 30781951
23. Hyatt D, Chen G-L, LoCascio P, Land M, Larimer F, Hauser L. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010; 11: 119 https://doi.org/10.1186/1471-2105-11-119 PMID: 20211023

24. Hyatt D, LoCascio PF, Hauser LJ, Uberbacher EC. Gene and translation initiation site prediction in metagenomic sequences. Bioinformatics. 2012; 28(17): 2223–2230 https://doi.org/10.1093/bioinformatics/bts429 PMID: 22796954

25. Yin Y, Mao X, Yang J, Chen X, Mao F, Xu Y. dbCAN: a web resource for automated carbohydrate-active enzyme annotation. Nucleic Acids Res. 2012; 40: W445–451 https://doi.org/10.1093/nar/gks479 PMID: 22645317

26. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011; 8(10): 785–786 https://doi.org/10.1038/nmeth.1701 PMID: 21959131

27. Aslanidis C, de Jong PJ. Ligation-independent cloning of PCR products (LIC-PCR). Nucleic Acids Res. 1990; 18(20): 6069–6074 https://doi.org/10.1093/nar/18.20.6069 PMID: 2235490

28. Miller GL. Use of dinitrosalicylic acid reagent for determination of reducing sugar. Anal Chem. 1959; 31 (3): 426–428

29. Wood TM. Preparation of crystalline, amorphous, and dyed cellulase substrates. Methods in Enzymology. 1988; 160: 19–25

30. Sonnhammer ELL, Eddy SR, Durbin R. Pfam: A comprehensive database of protein domain families based on seed alignments. Proteins: Structure, Function, and Genetics. 1997; 28: 405–420

31. Kawaichi S, Ito N, Kamikawa R, Sugawara T, Yoshida T, Sako Y. *Ardenticatena maritima gen. nov., sp. nov.*, a ferric iron- and nitrate-reducing bacterium of the phylum '*Chloroflexi*' isolated from an iron-rich coastal hydrothermal field, and description of *Ardenticatenia* classis nov. International Journal of Systematic and Evolutionary Microbiology. 2013; 63(8): 2992–3002

32. Li Y, Irwin DC, Wilson DB. Processivity, substrate binding, and mechanism of cellulose hydrolysis by *Thermobifida fusca* Cel9A. Appl Environ Microbiol. 2007; 73(10): 3165–72 https://doi.org/10.1128/AEM.02960-06 PMID: 17369336

33. Irwin D, Shin D-H, Zhang S, Barr BK, Sakon J, Karplus PA et al. Roles of the catalytic domain and two cellulose binding domains of *Thermomonospora fusca* E4 in cellulose hydrolysis. J Bacteriol. 1998; 180(7): 1709–1714 PMID: 9537366

34. Sakon J, Irwin D, Wilson DB, Karplus PA. Structure and mechanism of endo/exocellulase E4 from *Thermomonospora fusca*. Nat Struct Biol. 1997; 4(10): 810–818 PMID: 9334746

35. Tormo J, Lamed R, Chirino AJ, Morag E, Bayer EA, Shoham Y et al. Crystal structure of a bacterial family-III cellulose-binding domain: a general mechanism for attachment to cellulose. EMBO J. 1996; 15 (21): 5739–5751 PMID: 8918451

36. Herlet J, Kornberger P, Roessler B, Glanz J, Schwarz WH, Liebl W et al. A new method to evaluate temperature vs. pH activity profiles for biotechnological relevant enzymes. Biotechnol Biofuels. 2017; 10: 234 https://doi.org/10.1186/s13068-017-0923-9 PMID: 29046720

37. Stoll VS, Blanchard JS. Buffers: Principles and practice. Methods in Enzymology. 1990; 24–38

38. Miyatake F, Iwabuchi K. Effect of high compost temperature on enzymatic activity and species diversity of culturable bacteria in cattle manure compost. Bioresource Technology. 2005; 96(16): 1821–1825 https://doi.org/10.1016/j.biortech.2005.01.005 PMID: 16051089

39. Zverlov V, Mahr S, Riedel K, Bronnenmeier K. Properties and gene structure of a bifunctional cellulolytic enzyme (CelA) from the extreme thermophile '*Anaerocellum thermophilum*' with separate glycosyl hydrolase family 9 and 48 catalytic domains. Microbiology. 1998; 144(2): 457–465

40. Watson DL, Wilson DB, Walker LP. Synergism in binary mixtures of *Thermobifida fusca* cellulases Cel6B, Cel9A, and Cel5A on BMCC and Avicel. Applied Biochemistry and Biotechnology. 2002; 101 (2): 97–111 PMID: 12049205

41. Hamre AG, Kaupang A, Payne CM, Väljamäe P, Sørlie M. Thermodynamic signatures of substrate binding for three *Thermobifida fusca* cellulases with different modes of action. Biochemistry. 2019; 58 (12): 1648–1659 https://doi.org/10.1021/acs.biochem.9b00014 PMID: 30785271

42. Horn SJ, Sørlie M, Vårum KM, Väljamäe P, Eijsink VGH. Measuring processivity. Methods in Enzymology. 2012; 510: 69–95 https://doi.org/10.1016/B978-0-12-415931-0.00005-7 PMID: 22608722

43. Nguyen STC, Freund HL, Kasanjian J, Berlemont R. Function, distribution, and annotation of characterized cellulases, xylanases, and chitinases from CAZy. Applied Microbiology and Biotechnology. 2018; 102(4): 1629–1637 https://doi.org/10.1007/s00253-018-8778-y PMID: 29359269

44. Ravachol J, Borne R, Tardif C, Philip P, Fierobe H-P. Characterization of all family-9 glycoside hydrolases synthesized by the cellulosome-producing bacterium *Clostridium cellulolyticum*. The Journal of Biological Chemistry. 2014; 289(11): 7335–7348 https://doi.org/10.1074/jbc.M113.545046 PMID: 24451379

45. Eckert K, Zielinski F, Lo Leggio L, Schneider E. Gene cloning, sequencing, and characterization of a family 9 endoglucanase (CelA) with an unusual pattern of activity from the thermoacidophile

Alicyclobacillus acidocaldarius ATCC27009. Appl. Microbiol. Biotechnol. 2002; 60(4): 428–436 https://doi.org/10.1007/s00253-002-1131-4 PMID: 12466883

46. Hirano N, Hasegawa H, Nihei S, Haruki M. Cell-free protein synthesis and substrate specificity of full-length endoglucanase CelJ (Cel9D-Cel44A), the largest multi-enzyme subunit of the *Clostridium thermocellum* cellulosome. FEMS Microbiol. Lett. 2013; 344(1): 25–30 https://doi.org/10.1111/1574-6968.12149 PMID: 23560999

47. Deutschmann R, Dekker RFH. From plant biomass to bio-based chemicals: latest developments in xylan research. Biotechnology Advances. 2012; 30(6): 1627–1640 https://doi.org/10.1016/j.biotechadv.2012.07.001 PMID: 22776161