

# Bioinformatics and functional selection of GH77 4-α-glucanotransferases for potato starch modification

**Christensen, Stefan Jarl; Madsen, Michael Schmidt; Zinck, Signe Schram; Hedberg, Christinne; Sørensen, Ole Bandsholm; Svensson, Birte; Meyer, Anne S.**

*Published in:*
New Biotechnology

*Link to article, DOI:*
10.1016/j.nbt.2023.12.002

*Publication date:*
2024

*Document Version*
Publisher's PDF, also known as Version of record

New BIOTECHNOLOGY 79 (2024) 39–49

Contents lists available at ScienceDirect

New BIOTECHNOLOGY

journal homepage: www.elsevier.com/locate/nbt

# Bioinformatics and functional selection of GH77 4-α-glucanotransferases for potato starch modification

Stefan Jarl Christensen [a,b], Michael Schmidt Madsen [a], Signe Schram Zinck [a,b], Christinne Hedberg [b], Ole Bandsholm Sørensen [b], Birte Svensson [a], Anne S. Meyer [a,*]

[a] *Protein Chemistry and Enzyme Technology, DTU Bioengineering, Department of Biotechnology and Biomedicine, Technical University of Denmark, Kgs. Lyngby, Denmark*

[b] *KMC, Brande, Denmark*

ARTICLE INFO

*Keywords:*
4-α-glucanotransferase
Potato starch
Enzyme discovery
Conserved unique peptide patterns (CUPP)

ABSTRACT

4-α-glucanotransferases (4αGTs, EC 2.4.1.25) from glycoside hydrolase family 77 (GH77) catalyze chain elongation of starch amylopectin chains and can be utilized to structurally modify starch to tailor its gelation properties. The potential relationship between the structural design of 4αGTs and functional starch modification is unknown. Here, family GH77 was mined in silico for enzyme candidates based on sub-grouping guided by Conserved Unique Peptide Patterns (CUPP) bioinformatics categorization. From + 12,000 protein sequences a representative set of 27 4αGTs, representing four different domain architectures, different bacterial origins and diverse CUPP groups, was selected for heterologous expression and further study. Most of the enzymes catalyzed starch modification, but their efficacies varied substantially. Five of the 4αGTs were characterized in detail, and their action was compared to that of the industrial benchmark enzyme, Tt4αGT (CUPP 77_1.2), from *Thermus thermophilus*. Reaction optima of the five 4αGTs ranged from ~40–60 °C and pH 7.3–9.0. Several were stable for a minimum 4 h at 70 °C. Domain architecture type A proteins, consisting only of a catalytic domain, had high thermal stability and high starch modification ability. All five novel 4αGTs (and Tt4αGT) induced enhanced gelling of potato starch. One, At4αGT from *Azospirillum thermophilum* (CUPP 77_2.4), displayed distinct starch modifying abilities, whereas T24αGT from *Thermus* sp. *2.9* (CUPP 77_1.2) modified the starch similarly to Tt4αGT, but slightly more effectively. T24αGT and At4αGT are thus interesting candidates for industrial starch modification. A model is proposed to explain the link between the 4αGT induced molecular modifications and macroscopic starch gelation.

## Introduction

4-α-glucanotransferases (4αGTs, EC. 2.4.1.25) are multifunctional enzymes that are categorized in three different glycoside hydrolase (GH) families, namely GH13, GH57, and GH77 in the CAZy database (http://www.cazy.org/) [1]. The main reaction catalyzed by 4αGTs is a disproportionation reaction, which involves cleavage of an α-1,4-bond in a (starch-type) glucan polysaccharide, followed by transfer of the reducing end product to an acceptor glucan to establish a new α-1, 4-bond [2,3]. In starch, this chain transfer reaction results in chain elongation of the amylopectin side chains while simultaneously depleting amylose [3,4]. When the glucan transfer occurs intramolecularly within a single glucan chain, α-1,4-glucan cyclization results, but this is considered a side reaction of the 4αGTs disproportionation reaction. Hydrolysis is another secondary activity alongside the predominant disproportionation reaction.

The enzymatic reaction catalyzed by 4αGTs immediately results in decreased viscosity of starch suspensions. Nonetheless, the reaction enables production of firm thermo-reversible starch gels upon cooling and retrogradation [4–6]. The ability to generate molecular changes that induce such changes in the macromolecular properties of starch can be used to create enhanced gelling, improved mouth-feel, and decreased

*Abbreviations:* 4αGTs, 4-α-glucanotransferases; BCA, bicinchoninic acid; CBM, carbohydrate binding module, CD, catalytic domain; CUPP, conserved unique peptide patterns; DP, degree of polymerization; GH, glycoside hydrolase; HPAEC-PAD, high-performance anion-exchange chromatography with pulsed amperometric detection; HPSEC, high-performance size-exclusion chromatography; TTT, tube turn test.

* Corresponding author.
*E-mail address:* asme@dtu.dk (A.S. Meyer).

https://doi.org/10.1016/j.nbt.2023.12.002
Received 29 July 2023; Received in revised form 26 November 2023; Accepted 9 December 2023
Available online 12 December 2023

digestibility of starch in food applications [7,8]. As a result, microbially derived 4αGTs are attracting interest in the production of clean label texturizing ingredients to replace, for example, gelatin and fats in foods such as yoghurts and spreads [3,5,9–11]. However, the relationship between the enzyme characteristics, *e.g.*, the modular architecture of the protein, the enzyme-catalyzed molecular changes in starch, and the physicochemical properties of the modified starch is not fully understood [12].

The enzymes of the monospecific family GH77 are currently the most widely used 4αGTs in industry [6,13]. Based on sequence, the GH77 proteins can roughly be divided into four domain architectures, denoted A through D [14]. Architecture A consists solely of a catalytic domain, while B and C have an additional N-terminal domain and two N-terminal carbohydrate binding modules (CBMs), respectively. Architectures C and D have a domain of unknown function interrupting the catalytic domain. The GH77 family is large, consisting of almost 16000 sequences in the CAZy database [1]. Although in silico analysis of GH77 sequences has been reported [15], the family has not been categorized into subgroups depending on functionality.

The objective of the present study was to examine the functional diversity of the GH77 family and use Conserved Unique Peptide Patterns (CUPP) [16] as a tool for enzyme mining to discover 4αGTs with efficient potato starch modification abilities for textural enhancement. A secondary objective was to elucidate whether there was a relationship between 4αGT domain architecture, functional molecular starch modification, and gelling tendency of 4αGT modified potato starch.

## Materials and methods

### *Materials*

Native potato starch (moisture content 15.2% w/w, dry basis; purity>99%) was obtained from KMC (Brande, Denmark). Potato amylopectin was purchased from Carbosynth (Compton, Berkshire, England). Potato amylose, isoamylase, and glucose oxidase-peroxidase (GOPOD) reagent were purchased from Megazyme (Wicklow, Ireland). Chemicals were obtained from Merck (Darmstadt, Germany).

### *Sequence acquisition and phylogenetic analysis*

A sequence database was created based on all 12192 GH77 protein sequences available in the CUPP database [17]. Additionally, an NCBI BLASTp search (https://blast.ncbi.nlm.nih.gov/Blast.cgi) of the commercially used Tt4αGT (**BAA33728.1)** was performed. From this search an additional GH77 sequence (**WP_039459293.1**) from the bacterium *Thermus* sp. 2.9, denoted T24αGT, was identified and added to the dataset. Surprisingly, the dataset contained a high number of duplicate sequences (5402 had duplicate sequences); all those duplicate sequences were excluded. Furthermore, partial sequences, defined as having less than 400 amino acid residues, were excluded (a total of 271 sequences). Moreover, to strengthen the sequence alignment for the phylogenetic analysis, sequences having more than 1000 amino acid residues were excluded (coincidentally, also here 271 sequences were removed from the dataset). The remaining 6229 sequences were aligned using MUSCLE [18]. In order to obtain a sequence space comprising sequences of high diversity, redundancy was removed with a 90% sequence identity threshold using the Jalview software (version 2.11.2.6) [19]. Despite the relatively high identity threshold, this step resulted in removal of 3264 sequences, mainly due to the presence of large groups of sequences with very high similarity (>90%). Alignment optimization was done using MaxAlign [20]. Despite having only a few members, the CUPP groups 8.1, 12.1, 13.1, 16.1, and 19.1 were preserved to avoid deletion of these groups from the analysis. Sequences introducing long gaps in the alignment were removed manually. The resulting 2190 sequences were realigned, and a maximum likelihood tree was interfered in CIPRES [21] using IQ-tree [22] with 1000 bootstrap replicates using the Ultrafast algorithm [23]. Tree inference was performed using the Le and Gascuel substitution model [24] with invariable sites allowed, four gamma categories, and empirical amino acid frequencies. The tree was visualized using iTOL (https://itol.embl.de/).

### *Production and purification of enzymes*

A total of 27 enzymes were selected for production and purification (Table 1). Synthetic genes were codon optimized for *E. coli* expression and cloned into the pET28a-TEV vector using *Nco*I and *Xho*I (C-terminal His-tag) or *Nde*I and *Bam*HI (N-terminal His-tag) restriction sites (Genscript, Piscataway, NJ, USA). The His-tag was present throughout all the described stages. The enzymes were produced in *E. coli* BL21(DE3) and purified by $Ni^{2+}$ affinity column chromatography using a Histrap™ FF crude column (GE Healthcare, UK) [14]. To customize the expression and achieve maximum yields, different fermentation temperatures, inducers and inducer levels were used (indicated in Suppl. Table S1). Protein purity was verified by SDS-PAGE and protein concentrations were determined spectrophotometrically at 280 nm and converted using theoretical extinction coefficients for each enzyme calculated from the amino acid sequence and protomer molecular weight [25] (Suppl.

**Table 1**
GH77 4αGT enzymes selected for study towards potato starch modification.

| Enzyme name | Organism | CUPP group | Domain architecture | Accession number |
|---|---|---|---|---|
| Tt4αGT | *Thermus thermophilus* | 1.2 | A | BAA33728.1 |
| Ai4αGT | *Allomeiothermus silvanus* | 1.2 | A | ADH63360.1 |
| Ca4αGT | *Caldilinea aerophila* | 1.2 | A | BAL98727.1 |
| Ch4αGT | *Chloracidobacterium thermophilum* | 1.2 | A | AEP13688.1 |
| Ci4αGT | *Calidithermus timidus* | 1.2 | A | WP_018465025.1 |
| Ht4αGT | *Hydrogenovibrio thermophilus* | 1.2 | A | QAB14742.1 |
| Mr4αGT | *Meiothermus ruber* | 1.2 | A | ADD27885.1 |
| Op4αGT | *Oceanithermus profundus* | 1.2 | A | ADR36759.1 |
| T24αGT | *Thermus sp. 2.9* | 1.2 | A | WP_039459293.1 |
| T94αGT | *Thioalkalivibrio sp. K90mix* | 1.2 | A | ADC72031.1 |
| Tl4αGT | *Thermostichus lividus* | 1.2 | A | ATS18307.1 |
| Ty4αGT | *Thermus thermamylovorans* | 1.2 | A | WP_130839446.1 |
| As4αGT | *Acidilobus saccharovorans* | 1.3 | A | ADL19304.1 |
| Tj4αGT | *Thermocrinis jamiesonii* | 1.3 | A | WP_029552136.1 |
| Mg4αGT | *Mixta gaviniae* | 2.1 | B | AUX91938.1 |
| At4αGT | *Azospirillum thermophilum* | 2.4 | B | AWK89952.1 |
| Cg4αGT | *Corynebacterium glutamicum* | 3.1 | B | BAB99690.1 |
| Tu4αGT | *Thermomonospora curvata* | 3.1 | B | ACY96929.1 |
| Hl4αGT | *Herbinix luporum* | 4.1 | A | CUH92217.1 |
| Tm4αGT | *Thermotalea metallivorans* | 4.1 | A | WP_068554365.1 |
| Bt4αGT | *Bacteroides thetaiotaomicron* | 7.1 | C | AAO77253.1 |
| Ct4αGT | *Chryseobacterium taklimakanense* | 7.2 | C | AZI19670.1 |
| Co4αGT | *Cohaesibacter sp. ES.047* | 9.1 | B | SNY93242.1 |
| Ac4αGT | *Azorhizobium caulinodans* | 10.1 | B | BAF90115.1 |
| Sa4αGT | *Sandaracinus amylolyticus* | 13.1 | A | AKF08577.1 |
| Rp4αGT | *Rhodopseudomonas palustris* | 18.1 | B | ABJ07644.1 |
| Ag4αGT | *Akkermansia glycaniphila* | 0.0 | A | SEH74989.1 |

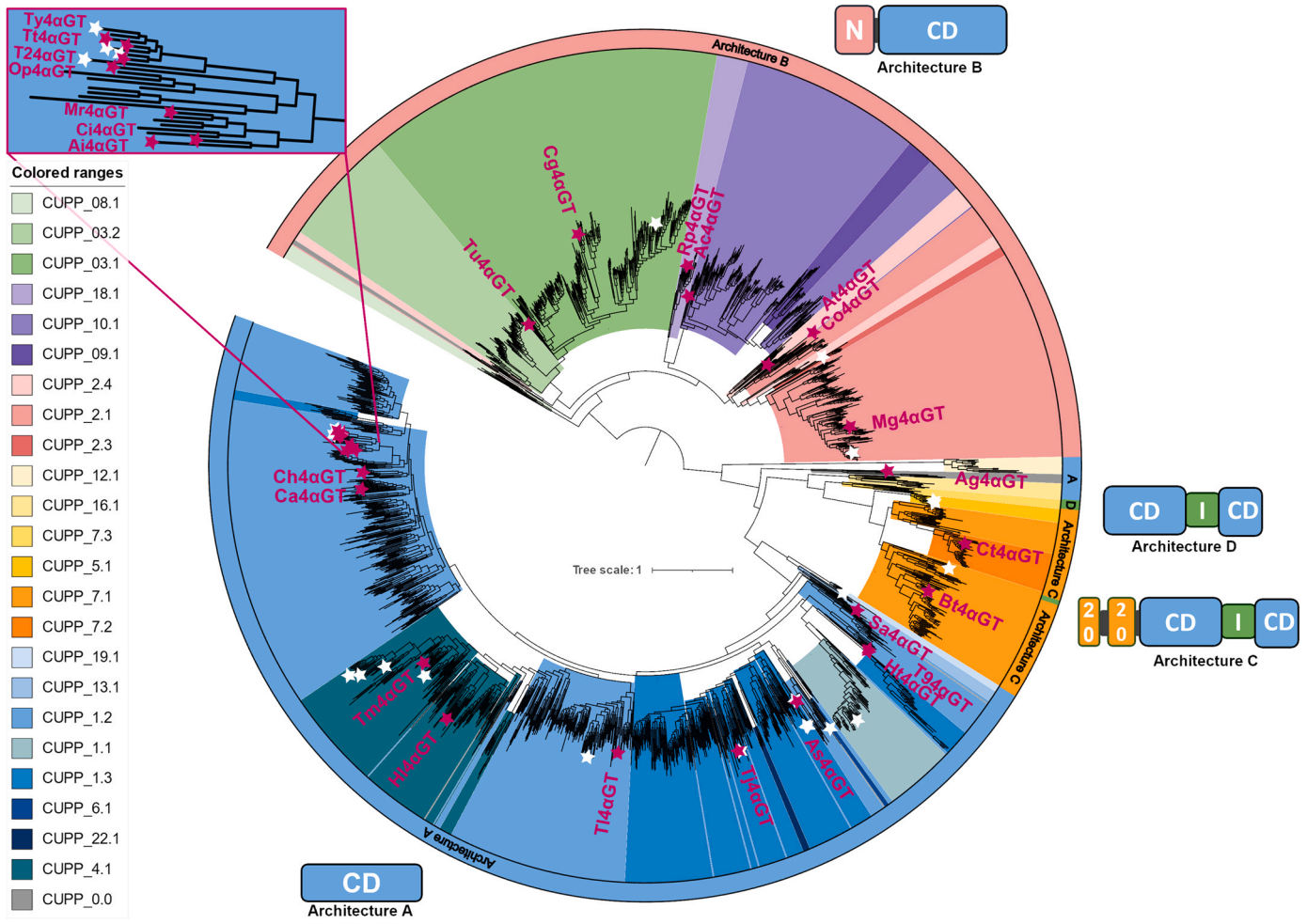


**Fig. 1.** Maximal likelihood tree based on 2190 GH77 protein sequences. Color ranges indicate CUPP groups. The legend is ordered according to the tree-order. The different colors around the circle's edge represent different enzyme domain architectures. Catalytic domains (CD), interrupting domains of unknown function (I), N-terminal domains (N), and CBM20s (20) are indicated in domain architecture figures. White stars represent previously characterized GH77 4αGTs (Suppl. Table S2). Pink stars represent GH77 members selected for screening in this study, with enzyme names indicated in pink text. For enzyme abbreviations, please see Table 1.

Table S1).

*Starch modification efficacy assessment*

Potato starch was gelatinized by boiling during dispersion using an Ultra Turrax T25 (Ika, Germany) as described previously [14]. Starch suspensions were pH adjusted to pH 7.0 using NaOH. Enzymatic treatment of triplicate reactions containing 3.5% (w/w) gelatinized potato starch and 75 nM enzyme were run at 30 or 50 °C and stirred at 1500 rpm for 30 min. Enzymatic reactions were terminated by heating to 100 °C for 15 min. Control samples were prepared without enzyme.

*Amylose and amylopectin chain-length analysis*

Amylose quantification of modified starch samples (0.1 g/L) was performed using iodine staining as described previously [14]. Standard curves were prepared using potato amylose and amylopectin, respectively. Chain-length of amylopectin chains, including 4αGT catalyzed chain elongations, were investigated by chain-length distribution analysis [14,25] by high-performance anion-exchange chromatography with pulsed amperometric detection (HPAEC-PAD) on a Dionex™ DC 5000 system equipped with a CarboPac PA1 column as described earlier [14]. Briefly, gelatinized starch samples (0.4 g/L) were debranched using isoamylase (1.5 U/mg starch), inactivated (100 °C for 10 min), and filtered (0.22 μm syringe filters) prior to being injected on the HPAEC system. Glucose, maltose, and linear α-1,4-glucans with a degree of polymerization (DP) of 3–50 were analyzed on samples of 10 μg of starch. Amylose was washed off the column with 1 M sodium acetate without NaOH for 5 min after each run [26].

*pH-temperature reaction optima*

The combined temperature-pH reaction optimum was estimated for each enzyme by fitting the data of three independent pH optimum curves (using gelatinized potato starch (1% w/w) as substrate), performed at three temperatures, to a multiple regression model. JMP® Pro 15.0.0 (SAS, Cary, NC, USA) was used as aid for the multivariate analyses, and graphical display of the data. The tested pH and temperature ranges varied slightly for the different enzymes (based on preliminary experiments), but generally ranged from pH 5–10 and 30–80 °C, respectively. The pH was adjusted using universal buffer UB5 (20 mM sodium acetate, bis-tris, tris-HCl, and glycine) - a modification of the previously described UB2 buffer system [27] augmented with glycine to extend the pH-range. The sodium ion concentration was kept constant at 40 mM by NaCl addition. Reactions were run for 10 min during stirring at 1500 rpm and terminated by addition of NaOH to a final concentration of 0.25 M. Amylose quantification to assess amylose depletion was done as described above; the amylose depletion rate was defined as the total mass of amylose depleted within a 10 min reaction and calculated per μM of enzyme in the reaction mixture.

*Thermal stability of enzymes*

The enzyme stability in presence of starch was determined by incubating 0.5 μM enzyme with 1% (w/w) potato starch at three different temperatures for each enzyme. Prior to addition of the enzyme, the gelatinized starch was pH adjusted with 20 mM UB5 and heated to the desired temperature. Incubation was done for up to 4 h, with different sampling times using transfer to a thermo-block at 4 °C. Enzymes were subsequently incubated with 20 mM maltotriose at 30 °C for 30 min, to determine residual activity. The glucose concentration was quantified using the GOPOD glucose oxidase assay as described elsewhere [14]. This method was suitable for measuring residual activity since 4αGT activity only produced glucose in the presence of maltose and not during the incubation with starch. All experiments were done in triplicate.

*Starch hydrolysis quantification*

8% (w/w) gelatinized starch were treated with 5 μM enzyme, stirred at 1500 rpm, and run (triplicate reactions) for 1, 2, and 4 h at the optimum pH-temperature for each enzyme. Reaction pH adjustment and reaction termination were done as described above for pH-temperature optima determinations. Starch hydrolysis was determined by quantifying reducing sugar formation in the enzyme-treated starch samples by the bicinchoninic acid (BCA) method (samples were diluted to 1% w/w starch) [14]. Initial rates were derived by linear regression of the progress of the reducing sugar formation, quantified based on a glucose standard series ranging from 0 to 400 μM.

*4αGT catalyzed starch modification rates and starch gelation evaluation*

Starch modification rates were determined by incubating each 4αGT enzyme (triplicate reactions) on 8% (w/w) gelatinized starch for 4, 7, and 10 min at the enzyme's optimum reaction pH and temperature while stirring at 1500 rpm. Reaction termination and amylose quantification were as described above. Initial rates were determined using linear regression on the progress of amylose depletion. The enzyme-induced starch modification and gel formation were also evaluated in 30 mL batch reactions on 8% (w/w) gelatinized starch. Based on activity measurements, the enzymes were dosed as follows: 150 nM Tt4αGT, 126 nM T24αGT, 195 nM Op4αGT, 285 nM Ca4αGT, 450 nM Sa4αGT, 450 nM At4αG. Reactions were run for 4 h with defined sampling times (inactivation of 0.5 mL aliquots at 100 °C for 15 min). Samples were analyzed for amylose depletion, substrate chain-length modification, and molecular weight distribution as described above and the gelling tendency was evaluated on 4% (w/w) modified starch (inactivated at 100 °C for 15 min) using the tube turn test (TTT) assay [14].

*Molecular weight distribution of enzymatically modified starch*

The molecular weight distribution of enzymatically modified starch samples was analyzed by High-Performance Size-Exclusion Chromatography (HPSEC) using an Ultimate system (a WPS-3000 sampler, an iso-3100 SD pump, and an ERC RefractoMax 520 refractive index detector) (Thermo Scientific, Waltham, MA, USA). The system was equipped with a Shodex SB-806 HQ column (300 × 8 mm), and a Shodex SB-G guard column (50 mm × 6 mm) (Showa Denko K.K., Tokyo, Japan) and samples were run as described previously [28,29]. External pullulan standards in the range of 342–805,000 Da (PSS Polymer Standards Service GmbH, Mainz, Germany) were applied to establish a polynomial relationship between the logarithmic molecular weight and the corresponding retention time in order to convert the retention times of the samples to molecular weights [29].

## Results

*Enzyme phylogeny, categorization, selection, and recombinant production*

A total of 12192 GH77 protein sequences were obtained from the CUPP database. An additional putative 4αGT, denoted T24αGT, was identified in a BLASTp search of the thermostable Tt4αGT (benchmark enzyme). A phylogenetic tree (Fig. 1) was constructed based on 2190 non-redundant GH77 sequences (Suppl. Fig S1, Table S2), representing

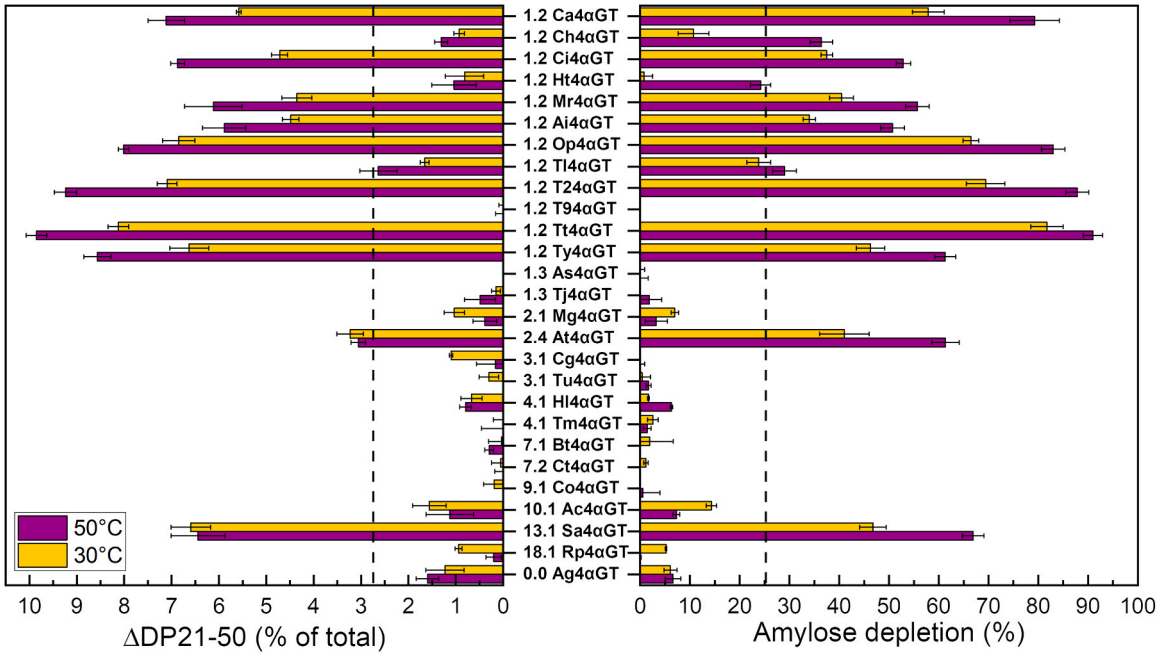


**Fig. 2.** Screening of potato starch modification efficacy of 4αGTs. Growth of the sum of DP21–50 chains, designated ΔDP21–50, (in % of the total mass) is based on chain length distribution analysis (Suppl. Fig. S2) and serves as an indication of the amylopectin chain elongation degree. Amylose depletion is based on a colorimetric iodine staining assay. Dashed lines represent the average. Reactions were performed at 3.5% starch at pH 7, using 75 nM enzyme. Enzyme names and CUPP groups are indicated.

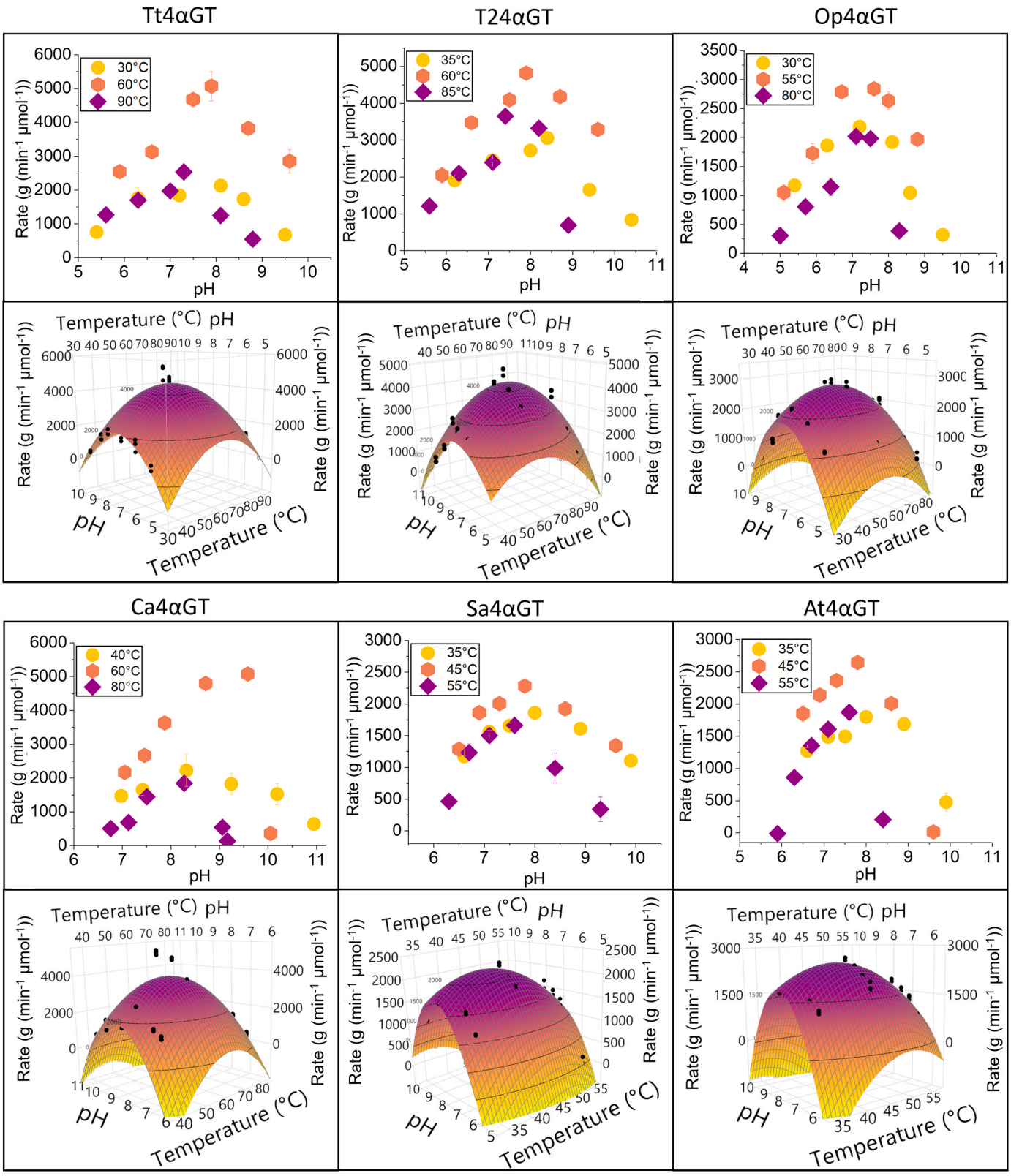


**Fig. 3.** pH-temperature optima of activity towards 1% (w/w) gelatinized potato starch for six 4αGTs. pH curves were performed at three different temperatures with one center temperature (the highest activity) flanked by a higher and a lower temperature. Specific rates of amylose depletion are shown (g ($min^{-1}$ $\mu mol^{-1}$). Datapoints obtained in the pH-activity assessments are shown as black dots in the surface plots. Enzyme concentrations were 20 nM Tt4αGT, 25 nM T24αGT, 50 nM Op4αGT, 30 nM Ca4αGT, 60 nM Sa4αGT, 50 nM At4αG.

**Table 2**
Optimal reaction conditions and stability of the five selected 4αGTs and the benchmark enzyme Tt4αGT (enzyme abbreviations as outlined in Table 1).

| Enzyme | Optimum[a] | | Thermal stability[b] | |
|---|---|---|---|---|
| | Temperature (°C) | pH | Stable for 4 h up to minimum (°C)[c] | $T_{1/2}$ (min) |
| Tt4αGT | 59–60 | 7.3 | 70 | ND[d] |
| T24αGT | 59–60 | 7.8 | 70 | ND[d] |
| Op4αGT | 50–51 | 7.8 | 70 | ND[d] |
| Ca4αGT | 55–56 | ~9.0 | 70 | ND[d] |
| Sa4αGT | 43–45 | 8.0 | 45 | 0.7[e] |
| At4αGT | 43–45 | 7.8 | 45 | 4.4[e] |

[a] Combined pH-temperature optimum according to RSM model, see Fig. 3;
[b] Thermal stability data. Detailed in Suppl. Fig. S4;
[c] The highest tested temperature with 100% residual activity after 4 h;
[d] Enzyme was stable at the highest tested temperature (70 °C) and no half-life was derived;
[e] Half-life at 55 °C.

**Table 3**
Amylose depletion and hydrolysis of potato starch catalyzed by the five selected 4αGTs and the Tt4αGT benchmark. The amylose depletion rate data were obtained as rates of amylose depletion quantified with the iodine assay using potato amylose and amylopectin standard series. The hydrolytic rates were obtained by measuring the specific reducing sugars formation rate ($min^{-1}$) in μM per minute per μM enzyme, using 5 μM enzyme. The reducing sugars were quantified with the BCA assay using a glucose standard series. All data are averages of the rates obtained for three replicate reactions.

| Enzyme | Amylose depletion rate (g ($min^{-1}$ $\mu mol^{-1}$)) | Hydrolytic rate ($min^{-1}$) |
|---|---|---|
| Tt4αGT[a] | 18.4 ± 0.6 | 1.11 ± 0.03 |
| T24αGT[b] | 21.8 ± 1.2 | 0.53 ± 0.01 |
| Op4αGT[c] | 13.9 ± 2.1 | 0.10 ± 0.03 |
| Ca4αGT[d] | 9.7 ± 1.9 | 0.68 ± 0.06 |
| Sa4αGT[e] | 6.2 ± 0.2 | 0.22 ± 0.03 |
| At4αGT[e] | 6.1 ± 0.5 | 0.08 ± 0.02 |

[a] reaction performed at 60 °C, pH 7.5;
[b] reaction performed at 60 °C, pH 8.0;
[c] reaction performed at 55 °C, pH 8.0;
[d] reaction performed at 60 °C, pH 9.0;
[e] reaction performed at 45 °C, pH 8.0

all 23 of the identified CUPP groups, reflecting a peptide-based similarity categorization [16]. The CUPP groups were generally congruent with defined phylogenetic clusters demonstrating a strong correlation between phylogeny and the unsupervised peptide-based CUPP categorization. The tree can roughly be divided into three large clusters reflecting the different 4αGT domain architectures (different colors around the circle edge in Fig. 1). A and B architecture sequences thus form two large separate clusters, while architectures C and D together form a smaller cluster. Most CUPP groups comprise sequences of microbial origin, except for CUPP groups 1.1 and 5.1 that represent plant sequences.

Based on the CUPP grouping and the phylogenetic analysis, 27 microbial sequences were selected for further analysis (Table 1) according to the following criteria: 1. Selection of a range of sequences representing the phylogenetic breadth and different CUPP groups to discover functional diversity and explore potato starch modifying abilities across the entire GH77 family; 2. Selection focused on sequences originating from thermophilic microbes; 3. A specific focus on enzymes clustering with the thermostable benchmark enzyme, Tt4αGT, previously demonstrated to have desired starch modification abilities and high efficacy [4, 14]. This cluster (denoted as the *thermus cluster*, blue box in Fig. 1) holds a high concentration of previously characterized enzymes (white stars in Fig. 1). Yet, the 27 selected GH77 sequences are spread throughout the phylogenetic tree and represent 12 CUPP groups and three domain architectures: A, B, and C. The minor D architecture group was excluded in the current study, as only low 4αGT activity was found from this group in a previous study [14]. One enzyme, Ag4αGT, was not categorized in any CUPP group (denoted CUPP 0.0).

According to criteria 2 and 3 described above, 11 sequences were selected from CUPP 1.2, which holds most of the sequences originating from microbes belonging to genera conventionally regarded as thermophilic, such as *Thermus*, *Meiothermus*, and *Rhodothermus* (Suppl. Table S2). Four of the 27 selected enzymes were characterized previously, *i.e.*, Ag4αGT, Tt4αGT, Bt4αGT, Cg4αGT [14], while the remaining are new putative 4αGTs. Six selected enzymes represent unexplored CUPP groups and clusters in the phylogenetic tree, namely enzymes from CUPP groups 2.4, 7.2, 9.1, 10.1, 13.1, and 18.1 (Table 1).

The 27 selected enzymes were fused with a His-tag, recombinantly produced in *E. coli*, and purified using His-trap chromatography. For domain architecture type A enzymes, the His-tag was positioned N-terminally, whereas the tag was in the C-terminal for the remaining enzymes to prevent interference with additional N-terminal domains. All the enzymes were successfully produced, but with substantial variations in the final yields (Suppl. Table S1). The final production yields were independent of the position of the His-tag, and no correlation was found between the yields and domain architecture or position in the phylogenetic tree.

### *Potato starch modification activity of 4αGTs*

The starch modification efficacy of the 27 selected, recombinantly expressed, purified 4αGT proteins was evaluated based on amylose depletion and amylopectin chain elongation degree [3,4]. Most of the enzymes induced potato starch modification, as evident from the changes in the chain length distributions, and the enzymes produced nearly similar chain length shift profiles (Suppl. Fig. S2). Hence, the 4αGTs generally induced a decrease in the quantity of DP11–20 chains, and an increase of DP1–10 and DP21–50 chains. The sum of the DP21–50 chain growth was used as an indication of degree of elongation of the amylopectin branch chains (Fig. 2).

Most of the investigated enzymes induced amylose depletion, which correlated with the amylopectin elongation degree. It was noted, though, that Co4αGT, T94αGT and As4αGT showed neither detectable amylose depletion nor amylopectin elongation under standard conditions. T94αGT and As4αGT originate from extremophilic microbes, and when screened at extreme pH, amylose depletion took place (Suppl. Fig. S3). Several of the 4αGTs had higher activity at 50 than at 30 °C, while some lost activity at 50 °C. Five novel 4αGTs were selected for further investigation. T24αGT, Op4αGT, and Ca4αGT were selected because of their high starch modifying efficacy (Ty4αGT was excluded due to its high similarity to T24αGT). At4αGT and Sa4αGT were selected for their high overall efficacy, their low similarity to CUPP 1.2 enzymes, and for being from unexplored CUPP groups (CUPP 2.4 and 13.1, respectively). Tt4αGT was included as the benchmark enzyme.

### *Reaction optima and thermostability*

Combined pH-temperature optima of all the five selected enzymes and Tt4αGT, were based on pH-activity curves performed at three different temperatures (Fig. 3). The datapoints obtained in the curves were fitted to a multiple regression model to derive the combined pH-temperature optimum (Table 2). Tt4αGT, T24αGT, Op4αGT, and Ca4αGT exhibited high temperature optima in the range of 50-60 °C, while Sa4αGT and At4αGT had optimum activity at 43-45 °C. Ca4αGT stood out by exhibiting reaction optimum at pH ~9.0, while the other five enzymes were most active in the range of pH 7.3–8.0 (Fig. 3).

Thermal stability of the enzymes was assessed by determining residual activity after incubation with 1% w/w gelatinized potato starch (Suppl. Fig. S4). This was done to account for the potential enzyme stabilizing effect of the substrate. Four enzymes, Tt4αGT, T24αGT, Op4αGT, and Ca4αGT, maintained 100% activity at all investigated

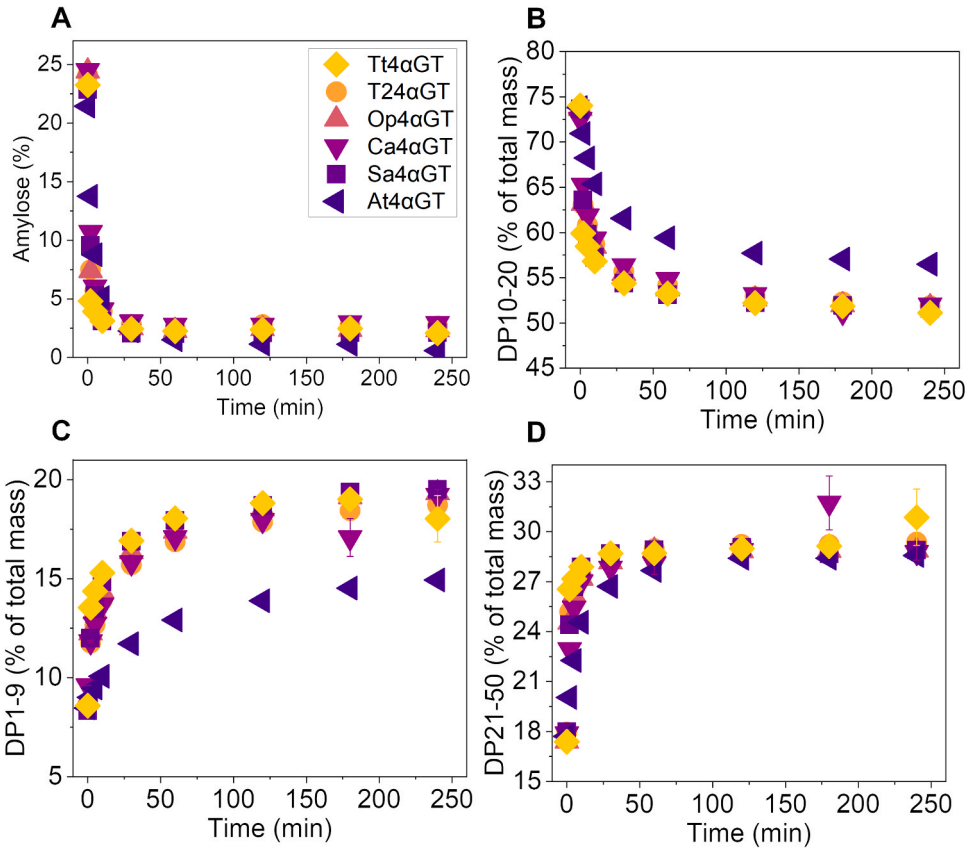


**Fig. 4.** Amylose depletion and chain length distribution change by six 4αGTs. Time dependent modification of 8% potato starch from batch experiments. (A) Amylose depletion over time. (B-D) Time dependent change in the percentage content of chains of three fractions: (B), DP10–20; (C), DP1–9; (D), DP21–50. All reactions were performed at the optimal reaction conditions of the enzymes and were the same as used in Table 3.

temperatures (Table 2), and were thus stable at least up to 70 °C for 4 h. Sa4αGT and At4αGT maintained 100% activity after 4 h at 45 °C but showed short half-lives (<5 min) at 55 °C (Table 2). All six enzymes were stable at their optimum temperature for at least 4 h.

*Enzymatic starch modification*

The potato starch modifying efficacy of the selected enzymes was evaluated by determining the initial amylose depletion rate at high substrate load (8%). The rates (Table 3) were determined at the optimum reaction conditions (Table 2 and Fig. 3). The initial hydrolysis rate of starch was determined based on the production of reducing sugars. The hydrolytic rate scaled somewhat with the amylose depletion, though the rate of the benchmark Tt4αGT exceeded the remaining 4αGTs (Table 3).

When dosed based on initial rates (Table 3) to aim for comparable degrees of starch modification, the enzymes induced similar starch modifications and thus catalyzed nearly similar amylose depletion (Fig. 4A), reduction of the sum of DP10–20 chains (Fig. 4B), and growth of DP1–9 and DP21–50 chains (Fig. 4C and D, Suppl. Fig. S5). However, the action progress of At4αGT differed from the other enzymes by leveling off at a lower amylose content and DP1–9 quantity, and at a higher DP10–20 quantity.

HPSEC analysis of the enzyme modified potato starch samples revealed a gradual decrease in molecular weight distribution as a function of the modification time (Fig. 5). This phenomenon appeared faster for modification catalyzed by Ca4αGT, Sa4αGT, and At4αGT when compared to the other enzymes (Table 3). The final molecular weight distribution peak (240 min) of starches modified with Ca4αGT, Sa4αGT, At4αGT appeared narrower and had shifted more towards lower molecular weights than those modified with Tt4αGT, T24αGT, and Op4αGT. Starch modified with At4αGT had a unique molecular weight profile that was narrower and more symmetrical after 240 min of modification.

The gelling tendency of 4αGT modified potato starch was evaluated using the tube turn test (TTT) [14], which reflects the capability of the resulting gel (stored for 18 h at 4 °C) to stay at the bottom of a test tube after it has been vortexed and turned (Suppl. Fig. S6). The test thus qualitatively indicates gel firmness. All six enzymes were able to produce firm gels (Table 4), though starch modified with Ca4αGT and Sa4αGT only stayed partially in the tube bottom, reflecting gels that were less firm ("+" in Table 4). The firm gels were produced after 2–5 min of enzymatic modification, while after 10–30 min the firmness decreased resulting in complete liberation from the tube bottom (Table 4).

## Discussion and conclusions

Microbial 4αGTs of CAZy family GH77 can catalyze starch chain elongation, and have recently gained industrial interest for their ability to convert native starch into clean label ingredients with designed textural properties via such reaction [3,11,12]. In this work, starch modification, including chain elongation was catalyzed by a range of microbial 4αGTs representing different bioinformatically categorized CUPP groups (Fig. 1). The starch modifying 4αGTs were found in distant CUPP groups, indicating that high starch modifying efficacy was not restricted to a specific group or phylogenetic cluster. Yet, there were profound efficacy variations among the enzymes belonging to different CUPP groups (Fig. 1, Fig. 2). Notably, all B and C domain architecture 4αGTs, except for At4αGT, exerted only weak potato starch modifying efficacy (below mean in Fig. 2). The data suggest a potential correlation between the lower efficacy and the presence of additional inserts and

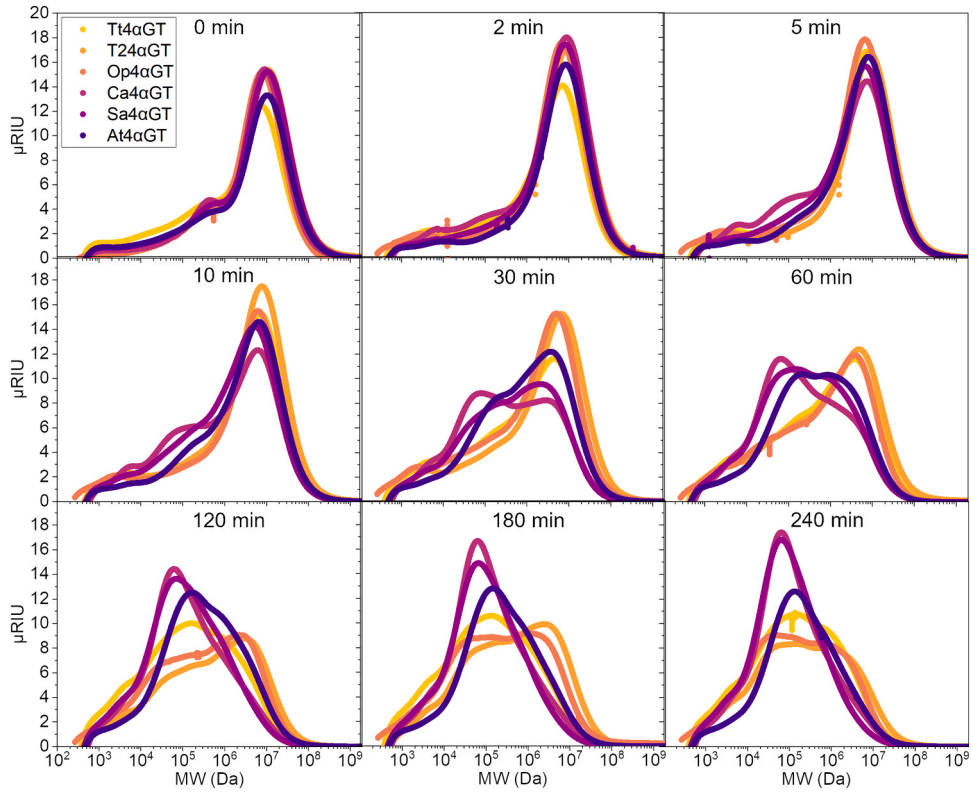


**Fig. 5.** Molecular weight distribution of potato starch modified with 4αGTs. modification of 8% w/w potato starch. The time indicates the enzymatic modification time. Reactions were performed at the optimal reaction condition for each enzyme.

**Table 4**
Gelling tendency of potato starch after 4αGT treatment assessed by the tube turn test (TTT) [14].

| Enzyme | Time (min)[a] | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 0 | 2 | 5 | 10 | 30 | 60 | 120 | 180 | 240 |
| | Gelling tendency[b] | | | | | | | | |
| Tt4αGT | -[c] | ++[d] | ++[d] | ++[d] | -[c] | -[c] | -[c] | -[c] | -[c] |
| T24αGT | -[c] | ++[d] | +[e] | ++[d] | -[c] | -[c] | -[c] | -[c] | -[c] |
| Op4αGT | -[c] | ++[d] | ++[d] | ++[d] | -[c] | -[c] | -[c] | -[c] | -[c] |
| Ca4αGT | -[c] | +[e] | +[e] | -[c] | -[c] | -[c] | -[c] | -[c] | -[c] |
| Sa4αGT | -[c] | +[e] | +[e] | +[e] | -[c] | -[c] | -[c] | -[c] | -[c] |
| At4αGT | -[c] | -[c] | ++[d] | +[e] | -[c] | -[c] | -[c] | -[c] | -[c] |

[a] Enzymatic treatment time;
[b] Gelling tendency based on TTT of 4% w/w potato starch modified at the conditions described in Table 3;
[c] did not form a firm gel;
[d] formed a firm gel;
[e] formed a less firm gel.

domains (Fig. 1), potentially influencing the openness or compactness of the catalytic cleft [14]. The study design did not allow us to firmly decipher the molecular basis for high or low starch modifying efficacy of GH77 4αGTs, as this will require more detailed structural analysis and ligand interaction investigations of the enzymes. Yet, the data establish that the vast majority of 4αGTs exhibiting high starch modifying efficacy are type A architecture enzymes, consisting of only the catalytic domain. Furthermore, especially type A enzymes found in CUPP group 1.2, localized in or close to the *Thermus* cluster (blue box in Fig. 1), appear to constitute a hot spot for efficient starch modifying 4αGTs.

To expand the options for discovery of additional enzymes in the CUPP group 1.2 hot spot, a BLAST search of Tt4αGT was performed, leading to the identification of T24αGT. Remarkably, this enzyme exhibited the highest amylose depletion rate of all the enzymes studied, and had thermostability and starch modification abilities on par with the benchmark Tt4αGT, and with significantly lower hydrolytic activity (Table 2, Table 3).

We also identified two enzymes with high starch modifying efficacy outside CUPP group 1.2, namely Sa4αGT and At4αGT, categorized in CUPP group 13.1 and 2.4, respectively. These findings suggest that also these previously unexplored CUPP groups might be hot spots for finding 4αGT enzymes exerting unique starch modifying activity, and illustrate that identification of such hot spots can be guided by CUPP categorization [30].

Thermostable enzymes are advantageous in industrial processes [31, 32], including industrial starch processing. Thus, their ability to maintain activity at high temperatures supports lower substrate viscosity, fast reaction rates, and prevents microbial contamination. In the present study we therefore deliberately targeted the discovery of thermostable 4αGTs by selecting enzymes (in the relevant CUPP groups) that originate from thermophilic bacteria. The bioinformatics analysis of the non-redundant GH77 sequences showed that most sequences originating from microbes belonging to genera conventionally regarded as thermophilic, were architecture A proteins (Suppl. Table S2). The experimental results confirmed that most of the investigated architecture A enzymes, including Tt4αGT, T24αGT, Ca4αGT, and Op4αGT, had higher activity at 50 °C than at 30 °C, whereas it was opposite for most of the B and C architecture enzymes (Fig. 2). We therefore propose that the simpler and more compact structure of architecture A enzymes is beneficial, if not a prerequisite, for thermostability. Indeed, fusion between an architecture A 4αGT enzyme and additional domains was recently shown to reduce thermostability [33].

The pH-temperature optima and the thermal stability of all the more deeply characterized enzymes, *i.e.*, T24αGT, Ca4αGT, Op4αGT, Sa4αGT, At4αGT, as well as Tt4αGT, correlate with the high thermal growth optima of the thermophilic bacteria from which the enzymes originate [34–39]. The optimum determined for the benchmark Tt4αGT enzyme,

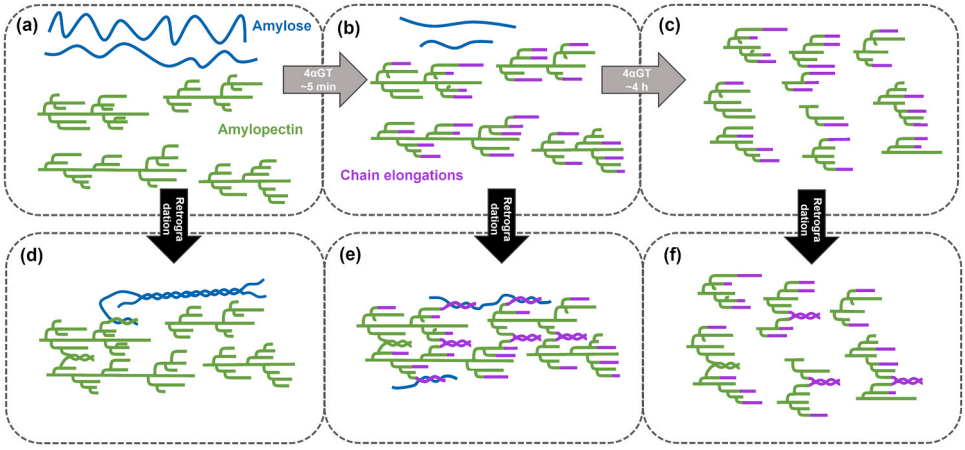


**Fig. 6.** A schematic model of 4αGT mediated starch modification and its effect on gelling and retrogradation. (a) Gelatinized native (untreated) potato starch. (b) Starch treated with 4αGT for ~5 min leading to amylose fragmentation and transfer to amylopectin chains. (c) Starch treated with 4αGT for ~4 h causing almost complete amylose depletion and fragmentation of modified amylopectin molecules (d) Retrogradation of native (untreated) potato starch causing amylose-amylose helix formation, amylopectin-amylopectin helix formation to a lower extent, and perhaps amylose-amylopectin helix formation, producing a weak gel network (e) Retrogradation of ~5 min 4αGT modified starch where longer amylopectin chains and shorter amylose fragments result in more helix formation, and a stronger gel network (f) Retrogradation of ~4 h modified starch where fragmentation of modified amylopectin prevents strong gel network formation despite high prevalence of intermolecular helixes.

~60 °C, pH 7.3 was in perfect agreement with recent data reported by others [40], but differed from our own previously published data using maltotriose as substrate, that gave reaction optimum at ~70–75 °C, pH ~5.5 [14]. We ascribe the difference in these reaction optima to differences in the composition of the reaction solution and infer that the careful adjustment of the total sodium salt level (40 mM NaCl), use of true starch as substrate (1% w/w gelatinized potato starch), and maybe even the extra addition of glycine to extend the pH range, influence the reaction optimum. This finding indicates that the use of a true starch substrate, and elevated salt level (ionic strength) may control the pH-temperature optimum of the Tt4αGT.

Batch experiments of potato starch modification revealed that all six characterized enzymes catalyzed amylose depletion and a characteristic chain length distribution change (Fig. 4). These observations are in line with previous results, and support assumptions that 4αGT mediated disproportionation leads to amylose fragmentation and transfer of amylose to amylopectin branch chains, causing their elongation [3,4]. Prolonged enzyme treatment resulted in a gradual molecular weight decrease of the modified starch. The phenomenon is in agreement with previous observations, and might indicate fragmentation of the elongated amylopectin structures through hydrolytic cleavage, intra-molecular cyclization, or transfer to small residual fragments by disproportionation [3,41,42]. The molecular weight reduction was observed at a later stage than the amylose depletion and amylopectin chain elongation, suggesting disproportionation as being the dominating activity, as also previously proposed [4,7]. The five new 4αGTs and Tt4αGT introduced varying molecular weight modifications leading to different HPSEC profiles of the products. This outcome may be a result of the significantly different hydrolytic activity of the enzymes (Table 3). We note, however, that the enzyme with highest hydrolytic activity, namely Tt4αGT, appeared to cause the slowest molecular weight reduction (on par with T24αGT and Op4αGT). Thus, we suggest that the molecular fragmentation is not only caused by hydrolysis, but may involve a mix of hydrolytic, and intermolecular and intramolecular transfer reactions [3,41,42].

Clearly, 4αGT catalyzed modification of starch initially leads to increased gel tendency, but the gel firmness then decreases with prolonged reaction time (Table 4). Our molecular interpretation of this phenomenon is illustrated in Fig. 6, and further elucidated below.

During cooling and retrogradation of native potato starch, a gradual recrystallization process occurs, most importantly triggered by an initial formation of amylose-amylose double-helixes, and strengthened by subsequent formation of amylopectin-amylopectin and probably amylose-amylopectin helixes [43–45]. Several factors govern the gelation behavior of starch including the amylose/amylopectin content and chain lengths [44,46–48]. Under the conditions used here, the unmodified potato starch was negative in the TTT, which might partly be due to the characteristic long amylose chains of potato starch probably hampering formation of a strong gel network [47]. After 5–10 min of 4αGT treatment, amylopectin branch chains are elongated and can possibly mimic the requisite amylose recrystallization behavior [42,49]. In addition, amylose is still present to a certain extent (~3–9%), but fragments into shorter chains possibly promoting formation of intermolecular helixes [41], which in combination with amylopectin chain elongation might increase the gelling tendency and gel firmness. However, after extended treatment, here ~30–240 min of modification, the firmness of the starch products decreases. At this point, the amylopectin elongation should be more pronounced, which might increase the number intermolecular helixes formed, but at the same time, a low amylose content (0.5–3%) and reduced overall molecular weight become evident. Previous results have demonstrated that low amylose and amylopectin molecular weight are unfavorable for formation of firm gels [50]. Hence, at prolonged enzymatic modification these events produce a molecular composition that does not favor the formation of a strong gel network. We note that a high modification degree, similar to the 4 h modification performed here, has previously been applied for production of texturizing agents [5]. The gelling tendency described in that patent was also highly dependent on the degree of modification (i.e., the reaction time), which is most likely related to the ability to generate a gel network via strengthened intermolecular double helixes. Overall, the available data infer that the modification degree depends on the choice of enzyme, and this can potentially be utilized to create novel starch-based ingredients.

Three of the novel enzymes identified in the present work, T24αGT, Ca4αGT and At4αGT, appear to be particularly promising 4αGT candidates for industrial starch modification. T24αGT demonstrated high thermostability and starch modification comparable to Tt4αGT, which is already industrially used, but T24αGT exhibited slightly higher catalytic transferase efficacy. Ca4αGT exhibited the highest pH optimum (pH ~9.0) and catalyzed extensive molecular weight reduction of potato starch. T24αGT, Ca4αGT and Tt4αGT were all categorized in CUPP-group 1.2, and are all type A enzymes, comprising one catalytic

domain only. In contrast, At4αGT, which bioinformatically is distantly related to the 4αGTs from CUPP group 1.2 and was the only type B architecture enzyme with high starch modification efficacy, performed exceptionally well among the type B enzymes, and introduced distinctly different molecular starch modifications than the CUPP-group 1.2 enzymes. This enzyme moreover had low hydrolytic activity and produced almost complete amylose depletion of potato starch. These unique properties of At4αGT make this enzyme particularly interesting for industrial starch modification.

## CRediT authorship contribution statement

**Christensen Stefan Jarl:** Conceptualization, Data curation, Formal analysis, Methodology, Writing – original draft, Writing – review & editing. **Hedberg Christinne:** Formal analysis, Project administration, Resources, Validation. **Sørensen Ole Bandsholm:** Funding acquisition, Investigation, Project administration, Resources, Supervision. **Madsen Michael Schmidt:** Formal analysis, Investigation. **Zinck Signe Schram:** Formal analysis, Investigation. **Svensson Birte:** Supervision, Writing – review & editing. **Meyer Anne S.:** Conceptualization, Funding acquisition, Methodology, Supervision, Writing – review & editing.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper: Michael S. Madsen, Birte Svensson, and Anne S. Meyer declare no competing interests. Stefan Jarl Christensen, Signe Schram Zinck, Christinne Hedberg, and Ole Bandsholm Sørensen are employed by the company KMC, a manufacturer of potato-based food ingredients, but declare no competing or potentially competing interests.

## Acknowledgements

This work was financially supported by Innovation Fund Denmark, Project 9066-00044B, and DTU.

## Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.nbt.2023.12.002.

## References

[1] Drula E, Garron ML, Dogan S, Lombard V, Henrissat B, Terrapon N. The carbohydrate-active enzyme database: functions and literature. Nucleic Acids Res 2022;50:D571–7. https://doi.org/10.1093/NAR/GKAB1045.
[2] Palmer T.N., Ryman B.E., Whelan W.J. The action pattern of amylomaltase from *Escherichia coli* 1976;115:105–15. https://doi.org/10.1111/j.1432–1033.1976.tb10863.x.
[3] van der Maarel MJEC, Capron I, Euverink GJW, Bos HT, Kaper T, Binnema DJ, et al. A novel thermoreversible gelling product made by enzymatic modification of starch. Starch/Staerke 2005;57:465–72. https://doi.org/10.1002/star.200500409.
[4] Hansen MR, Blennow A, Pedersen S, Nørgaard L, Engelsen SB. Gel texture and chain structure of amylomaltase-modified starches compared to gelatin. Food Hydrocoll 2008;22:1551–66. https://doi.org/10.1016/j.foodhyd.2007.10.010.
[5] Euverink G.J.W., Binnema D.J. WO1998015347A1 - Use of modified starch as an agent for forming a thermoreversible gel, 2002.
[6] Li X, Wang Y, Wu J, Jin Z, Dijkhuizen L, Hachem MA, et al. *Thermoproteus uzoniensis* 4-α-glucanotransferase catalyzed production of a thermo-reversible potato starch gel with superior rheological properties and freeze-thaw stability. Food Hydrocoll 2023;134. https://doi.org/10.1016/J.FOODHYD.2022.108026.
[7] Kaper T, Talik B, Ettema TJ, Bos H, van der Maarel MJEC, Dijkhuizen L. Amylomaltase of *Pyrobaculum aerophilum* IM2 produces thermoreversible starch gels. Appl Environ Microbiol 2005;71:5098–106. https://doi.org/10.1128/AEM.71.9.5098-5106.2005.
[8] Jiang H, Miao M, Ye F, Jiang B, Zhang T. Enzymatic modification of corn starch with 4-α-glucanotransferase results in increasing slow digestible and resistant starch. Int J Biol Macromol 2014;65:208–14. https://doi.org/10.1016/j.ijbiomac.2014.01.044.
[9] Alting, Fred van de Velde AC, Kanning MW, Burgering M, Mulleners L, Sein A, et al. Improved creaminess of low-fat yoghurt: the impact of amylomaltase-treated starch domains. Food Hydrocoll 2009;23:980–7. https://doi.org/10.1016/j.foodhyd.2008.07.011.
[10] Do VH, Mun S, Kim YL, Rho SJ, Park KH, Kim YR. Novel formulation of low-fat spread using rice starch modified by 4-α-glucanotransferase. Food Chem 2016;208:132–41. https://doi.org/10.1016/j.foodchem.2016.03.101.
[11] Chen Y, McClements DJ, Peng X, Chen L, Xu Z, Meng M, et al. Research progresses on enzymatic modification of starch with 4-α-glucanotransferase. Trends Food Sci Technol 2023;131:164–74. https://doi.org/10.1016/j.tifs.2022.11.025.
[12] Miao M, BeMiller JN. Enzymatic approaches for structuring starch to improve functionality. Annu Rev Food Sci Technol 2023;14:271–95. https://doi.org/10.1146/ANNUREV-FOOD-072122-023510.
[13] Miao M, Jiang B, Jin Z, BeMiller JN. Microbial starch-converting enzymes: recent insights and perspectives. Compr Rev Food Sci Food Saf 2018;17:1238–60. https://doi.org/10.1111/1541-4337.12381.
[14] Christensen SJ, Madsen MS, Zinck SS, Hedberg C, Sørensen OB, Svensson B, et al. Enzymatic potato starch modification and structure-function analysis of six diverse GH77 4-alpha-glucanotransferases. Int J Biol Macromol 2023;224:105–14. https://doi.org/10.1016/J.IJBIOMAC.2022.10.107.
[15] Kuchtová A, Janeček Š. In silico analysis of family GH77 with focus on amylomaltases from borreliae and disproportionating enzymes DPE2 from plants and bacteria. Biochim Biophys Acta - Proteins Proteom 2015;1854:1260–8. https://doi.org/10.1016/j.bbapap.2015.05.009.
[16] Barrett K, Lange L. Peptide-based functional annotation of carbohydrate-active enzymes by conserved unique peptide patterns (CUPP). Biotechnol Biofuels 2019;12:1–21. https://doi.org/10.1186/s13068-019-1436-5.
[17] Barrett K, Hunt CJ, Lange L, Meyer AS. Conserved unique peptide patterns (CUPP) online platform: peptide-based functional annotation of carbohydrate active enzymes. Nucleic Acids Res 2020;48:W110–5. https://doi.org/10.1093/NAR/GKAA375.
[18] Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 2004;32:1792–7. https://doi.org/10.1093/nar/gkh340.
[19] Waterhouse AM, Procter JB, Martin DMA, Clamp M, Barton GJ. Jalview Version 2-A multiple sequence alignment editor and analysis workbench. Bioinformatics 2009;25:1189–91. https://doi.org/10.1093/bioinformatics/btp033.
[20] Gouveia-Oliveira R, Sackett PW, Pedersen AG. MaxAlign: maximizing usable data in an alignment. BMC Bioinforma 2007;8:312. https://doi.org/10.1186/1471-2105-8-312.
[21] Miller M.A., Pfeiffer W., Schwartz T. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gatew. Comput. Environ. Work. GCE 2010, 2010, p. 1–8. https://doi.org/10.1109/GCE.2010.5676129.
[22] Nguyen LT, Schmidt HA, Von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol Biol Evol 2015;32:268–74. https://doi.org/10.1093/molbev/msu300.
[23] Hoang DT, Chernomor O, Von Haeseler A, Minh BQ, Vinh LS. UFBoot2: improving the ultrafast bootstrap approximation. Mol Biol Evol 2018;35:518–22. https://doi.org/10.1093/molbev/msx281.
[24] Le SQ, Gascuel O. An improved general amino acid replacement matrix. Mol Biol Evol 2008;25:1307–20. https://doi.org/10.1093/molbev/msn067.
[25] Gill SC, von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem 1989;182:319–26. https://doi.org/10.1016/0003-2697(89)90602-7.
[26] Blennow A, Bay-Smidt AM, Wischmann B, Olsen CE, Møller BL. The degree of starch phosphorylation is related to the chain length distribution of the neutral and the phosphorylated chains of amylopectin. Carbohydr Res 1998;307:45–54. https://doi.org/10.1016/S0008-6215(98)00015-9.
[27] Brooke D, Movahed N, Bothner B. Universal buffers for use in biochemistry and biophysical experiments. AIMS Biophys 2015;2:336–42. https://doi.org/10.3934/biophy.2015.3.336.
[28] Zinck SS, Christensen SJ, Sørensen OB, Svensson B, Meyer AS. Importance of inactivation methodology in enzymatic processing of raw potato starch: NaOCl as efficient α-amylase inactivation agent. Molecules 2023;28:2947. https://doi.org/10.3390/molecules28072947.
[29] Biel-Nielsen TL, Li K, Susanne OS, Sejberg JJP, Meyer AS, Holck J. Utilization of industrial citrus pectin side streams for enzymatic production of human milk oligosaccharides. Carbohydr Res 2022;519:108627. https://doi.org/10.1016/j.carres.2022.108627.
[30] Barrett K, Hunt CJ, Lange L, Grigoriev IV, Meyer AS. Conserved unique peptide patterns (CUPP) online platform 2.0: implementation of +1000 JGI fungal genomes. Nucleic Acids Res 2023;51:108–14. https://doi.org/10.1093/nar/gkad385.
[31] Zamost BL, Nielsen HK, Starnes RL. Thermostable enzymes for industrial applications. J Ind Microbiol 1991;8:71–81. https://doi.org/10.1007/BF01578757.
[32] Kaper T, van der Maarel MJEC, Euverink GJW, Dijkhuizen L. Exploring and exploiting starch-modifying amylomaltases from thermophiles. Biochem Soc Trans 2004;32:279–82. https://doi.org/10.1042/BST0320279.
[33] Wang Y, Wu Y, Christensen SJ, Janeček Š, Bai Y, Møller MS. Impact of starch binding domain fusion on activities and starch product structure of 4-α-glucanotransferase. Molecules 2023;28:1320. https://doi.org/10.3390/molecules28031320.
[34] Zhao ZL, Ming H, Ding CL, Ji WL, Cheng LJ, Niu MM, et al. *Azospirillum thermophilum* sp. Nov., isolated from a hot spring. Int J Syst Evol Microbiol 2020;70:550–4. https://doi.org/10.1099/ijsem.0.003788.

[35] Oshima T, Imahori K. Description of *Thermus thermophilus* (Yoshida and Oshima) comb. nov., a nonsporulating thermophilic bacterium from a Japanese thermal spa. Int J Syst Bacteriol 1974;24:102–12. https://doi.org/10.1099/00207713-24-1-102.
[36] Siozios S, Cestaro A, Kaur R, Pertot I, Rota-Stabelli O, Anfora G. Draft genome sequence of *Thermus* sp. isolate 2.9, obtained from a hot water spring located in Salta, Argentina. Genome Announc 2013;1:e00032–13. https://doi.org/10.1128/genomeA.00032-13.
[37] Miroshnichenko ML, L'Haridon S, Jeanthon C, Antipov AN, Kostrikina NA, Tindall BJ, et al. *Oceanithermus profundus* gen. nov., sp. nov., a thermophilic, microaerophilic, facultatively chemolithoheterotrophic bacterium from a deep-sea hydrothermal vent. Int J Syst Evol Microbiol 2003;53:747–52. https://doi.org/10.1099/ijs.0.02367-0.
[38] Mohr KI, Garcia RO, Gerth K, Irschik H, Müller R. *Sandaracinus amylolyticus* gen. nov., sp. nov., a starch-degrading soil myxobacterium, and description of Sandaracinaceae fam. nov. Int J Syst Evol Microbiol 2012;62:1191–8. https://doi.org/10.1099/ijs.0.033696-0.
[39] Sekiguchi Y, Yamada T, Hanada S, Ohashi A, Harada H, Kamagata Y. *Anaerolinea thermophila* gen. nov., sp. nov. and *Caldilinea aerophila* gen. nov., sp. nov., novel filamentous thermophiles that represent a previously uncultured lineage of the domain bacteria at the subphylum level. Int J Syst Evol Microbiol 2003;53:1843–51. https://doi.org/10.1099/ijs.0.02699-0.
[40] Wan H, Ouyang X, Yang T, Ye T, Jin M, Huan J. A 4-αglucanotransferase from *Thermus thermophilus* HB8; Secretory expression and characterization. Curr Microbiol 2022;79:202–10. https://doi.org/10.1007/s00284-022-02856-y.
[41] Do HV, Lee EJ, Park JH, Park KH, Shim JY, Mun S, et al. Structural and physicochemical properties of starch gels prepared from partially modified starches using *Thermus aquaticus* 4-α-glucanotransferase. Carbohydr Polym 2012;87:2455–63. https://doi.org/10.1016/j.carbpol.2011.11.021.
[42] Boonna S, Rolland-Sabaté A, Lourdin D, Tongta S. Macromolecular characteristics and fine structure of amylomaltase-treated cassava starch. Carbohydr Polym 2019;205:143–50. https://doi.org/10.1016/J.CARBPOL.2018.10.042.
[43] Gudmundsson M. Retrogradation of starch and the role of its components. Thermochim Acta 1994;246:329–41. https://doi.org/10.1016/0040-6031(94)80100-2.
[44] Fredriksson H, Silverio J, Andersson R, Eliasson AC, Åman P. The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. Carbohydr Polym 1998;35:119–34. https://doi.org/10.1016/S0144-8617(97)00247-6.
[45] Morris VJ. Starch gelation and retrogradation. Trends Food Sci Technol 1990;1:2–6. https://doi.org/10.1016/0924-2244(90)90002-G.
[46] Jane J, Chen JF. Effect of amylose molecular size and amylopectin branch chain length on paste properties of starch. Cereal Chem 1992;69:60–5.
[47] Reyniers S, Ooms N, Gomand SV, Delcour JA. What makes starch from potato (*Solanum tuberosum* L.) tubers unique: a review. Compr Rev Food Sci Food Saf 2020;19:2588–612. https://doi.org/10.1111/1541-4337.12596.
[48] Bertoft E, Annor GA, Shen X, Rumpagaporn P, Seetharaman K, Hamaker BR. Small differences in amylopectin fine structure may explain large functional differences of starch. Carbohydr Polym 2016;140:113–21. https://doi.org/10.1016/j.carbpol.2015.12.025.
[49] Srichuwong S, Jane J. Physicochemical properties of starch affected by molecular composition and structures: a review. Food Sci Biotechnol 2007;16:663–74.
[50] Mua JP, Jackson DS. Retrogradation and gel textural attributes of corn starch amylose and amylopectin fractions. J Cereal Sci 1998;27:157–66. https://doi.org/10.1006/jcrs.1997.0161.