



OPEN

## A novel GH3- $\beta$ -glucosidase from soda lake metagenomic libraries with desirable properties for biomass degradation

Oliyad Jeilu<sup>1,2,3✉</sup>, Erik Alexandersson<sup>1</sup>, Eva Johansson<sup>1</sup>, Addis Simachew<sup>3</sup> & Amare Gessesse<sup>3,4</sup>

Beta-glucosidases catalyze the hydrolysis of the glycosidic bonds of cellobiose, producing glucose, which is a rate-limiting step in cellulose biomass degradation. In industrial processes,  $\beta$ -glucosidases that are tolerant to glucose and stable under harsh industrial reaction conditions are required for efficient cellulose hydrolysis. In this study, we report the molecular cloning, *Escherichia coli* expression, and functional characterization of a  $\beta$ -glucosidase from the gene, CelGH3\_f17, identified from metagenomics libraries of an Ethiopian soda lake. The CelGH3\_f17 gene sequence contains a glycoside hydrolase family 3 catalytic domain (GH3). The heterologous expressed and purified enzyme exhibited optimal activity at 50 °C and pH 8.5. In addition, supplementation of 1 M salt and 300 mM glucose enhanced the  $\beta$ -glucosidase activity. Most of the metal ions and organic solvents tested did not affect the  $\beta$ -glucosidase activity. However, Cu<sup>2+</sup> and Mn<sup>2+</sup> ions, Mercaptoethanol and Triton X-100 reduce the activity of the enzyme. The studied  $\beta$ -glucosidase enzyme has multiple industrially desirable properties including thermostability, and alkaline, salt, and glucose tolerance.

**Keywords** Beta-glucosidase, Glycoside hydrolase family 3 (GH3), Soda lakes, Enzyme characterisation

Glycoside Hydrolase Family 3 (GH3), particularly beta-glucosidases, are pivotal in a variety of biological processes due to their ability to hydrolyze glycosidic bonds in complex sugars. These enzymes are known for their substrate versatility, playing a significant role in carbohydrate metabolism<sup>1</sup>. GH3 contains enzymes that can hydrolyze the terminal glycosidic bond in the non-reducing end of various glycosides and glycoconjugates<sup>2,3</sup>. GH3 enzymes are of interest for industrial applications<sup>3</sup>, and their structure, specificity, and biological role have been found to be extraordinarily diverse. Members of this family have been demonstrated as having a variety of enzymatic activities, including those of  $\beta$ -glucosidase,  $\beta$ -xylosidase,  $\alpha$ -L-arabinofuranosidase, and N-acetyl  $\beta$ -glucosaminidase<sup>3,4</sup>. GH3 enzymes have been widely found in bacteria, fungi, and plants<sup>5</sup>. Thus, GH3 is one of the most prevalent families of carbohydrate-active enzymes.

The  $\beta$ -glucosidases of the GH3 are known to hydrolyze terminal non-reducing  $\beta$ -D-glucosidic linkages, primarily cleaving cellobiose into glucose<sup>6,7</sup>. These enzymatic activities are of potential use in numerous biotechnological processes, including biomass degradation, pharmaceutical production, and the food industry<sup>8,9</sup>. For the degradation of cellulosic biomass, the  $\beta$ -glucosidase activity is regarded as the rate-limiting factor<sup>9,10</sup>.

To increase substrate accessibility in many industrial processes, including biomass degradation, the  $\beta$ -glucosidases need to operate under extreme conditions. Thus, enzymes like  $\beta$ -glucosidases must withstand high temperatures, extreme pH levels, elevated substrate and product concentrations, and the presence of organic solvents or detergents. These factors, prevalent in industries such as biofuel production, can significantly impact enzyme structure and function, necessitating robust enzymes capable of maintaining their activity<sup>11–13</sup>. Unfortunately, glucose which is the desired end product of biomass degradation, is known to inhibit and limit the activity of the  $\beta$ -glucosidases<sup>14</sup>. As a result,  $\beta$ -glucosidases that are stable both under harsh industrial conditions<sup>15,16</sup> and high glucose concentrations are required<sup>10</sup>. Consequently, most of the current industrially important enzymes,

<sup>1</sup>Department of Plant Breeding, Swedish University of Agricultural Sciences, Box 190, 23422 Lomma, Sweden. <sup>2</sup>Department of Civil and Environmental Engineering, Northwestern University, Evanston, IL 60208, USA. <sup>3</sup>Institute of Biotechnology, Addis Ababa University, P O Box 1176, Addis Ababa, Ethiopia. <sup>4</sup>Department of Biological Sciences and Biotechnology, Botswana International University of Science and Technology, Private Bag 16, Palapye, Botswana. ✉email: oliyad.jeilu.oumer@slu.se; oliyadjeilu@northwestern.edu

including the  $\beta$ -glucosidases, originate from mesophilic microorganisms. Despite their numerous benefits, these enzymes frequently lack stability under harsh industrial processes<sup>11</sup>. As a result, there is a high need for the exploration of more stable enzymes from extremophiles.

One source of new enzymes for industrial applications are extremophiles, which are able to live in harsh environments such as soda lakes<sup>17,18</sup>. Soda lakes present a variety of physical and chemical extremes, including alkalinity, and salinity (NaCl)<sup>19</sup>. Some of the are found in the East African Rift Valley in Ethiopia and Kenya<sup>20,21</sup>. Despite the particular environment, soda lakes are known to be incredibly productive and to harbour a range of haloalkaliphilic microbial communities that are responsible for the cycling of essential elements such as carbon, nitrogen, and sulfur<sup>22–24</sup>. These microorganisms offer huge potential as sources of biocatalysts that work under extreme conditions, useful for conditions in many industrial operations<sup>25,26</sup>.

To discover novel biocatalytic enzymes with potentially high functionality in industry, appropriate approaches are needed for screening and evaluation of a large number of microorganisms from extreme environments. Harnessing the microbial genetic pool in any environment, is possible by isolating a given microbial strain in pure culture form, followed by screening for enzyme production and cloning of the gene for further manipulation and/or large-scale production<sup>27–29</sup>. Even though this culture-dependent approach provides many novel microorganisms, only a small fraction of microbes (according to many estimates  $\leq 1\%$ ) can be recovered in culture media<sup>30,31</sup>.

Recently, culture-independent methods such as metagenomics have provided a powerful tool to access the genetic and metabolic diversity of microorganisms in any environment, which overcomes the limitations of culture-dependent approaches<sup>32</sup>. In functional metagenomics, genes are isolated from microorganisms of a certain environment and then cloned into libraries. Thus, functional metagenomics is favourable over other commonly used methods, as it allows the detection of microorganisms that are not culturable<sup>33,34</sup>. Functional metagenomics has already been extremely useful and effectively employed to identify enzymes with specific functions and novel enzyme families, including glycoside hydrolases family 3<sup>4,35,36</sup>. From the constructed library, a single gene or a group of genes are then identified based on certain enzymatic activities. The gene products are thereafter further assessed using a suitable screening assay<sup>11</sup>.

Based on our previous study, the aim of this study was to investigate and characterize the enzymatic properties of a gene (CelGH3\_f17) previously identified in a metagenomic library from an Ethiopian soda lake<sup>37</sup>, to elucidate its potential application in industrial processes. This included the cloning and expression of the gene, to study the properties and stability of the encoded  $\beta$ -glucosidase enzyme under extreme industrial conditions, such as high temperatures and varying pH levels. The study aimed to assess the enzyme's efficiency in biomass degradation and its resistance to inhibition by glucose, thereby contributing to the development of more robust biocatalysts for biomass degradation.

## Results

### Identification and selection of CelGH3\_f17

In our preceding research<sup>37</sup>, we constructed metagenomic libraries from Ethiopian Soda lakes, which underwent a functional screening process targeting hydrolytic enzymes, including  $\beta$ -glucosidase activity. Subsequent to the screening, clones demonstrating positive results were chosen for sequencing. These sequenced clones then underwent comprehensive bioinformatics analysis and were annotated utilizing the carbohydrate-active enzyme (CAZy) database<sup>37</sup>. Among the genes identified through this process, a gene named CelGH3\_f17, was chosen for more detailed downstream analysis due to its higher identity percentage and cloning success.

### Sequence and structural analysis of the CelGH3\_f17 protein

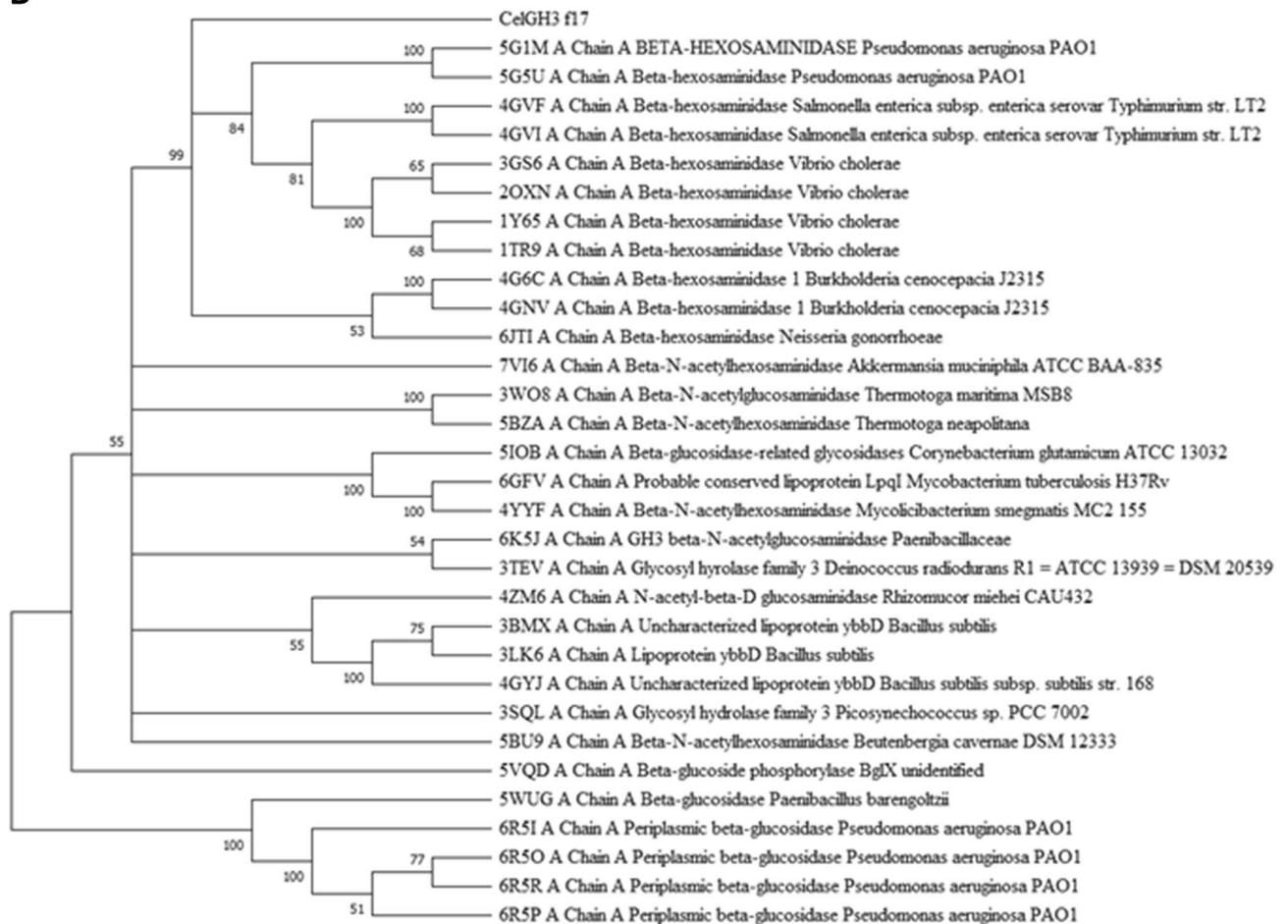
CelGH3\_f17 had an amino acid sequence of 229 and using the ExPASy's Compute pI/MW tool ([www.expasy.org/resources/compute-pi-mw](http://www.expasy.org/resources/compute-pi-mw)) the theoretical isoelectric point and molecular mass of the protein encoded by CelGH3\_f17 were 6.16 and  $\sim 25$  kDa, respectively. The sequence analysis with the SMART software package indicated that the putative protein had one GH3 region in the C-terminal in the amino acid positions 1–190 (Fig. 1A). Conserved domain searches in NCBI confirmed that the CelGH3\_f17 gene is likely to encode GH3  $\beta$ -glucosidase/ $\beta$ -hexosaminidase. BLASTP showed that the CelGH3\_f17 protein has a sequence similarity of 24–54% with 31 other proteins in the protein database (Supplementary Table S1). The majority of these proteins were classified as GH3 enzymes and were taxonomically linked to 21 organisms. Among them, 19 were identified as bacteria; where most of them were from *Pseudomonas spp.*, one was traced back to a fungal family, and the origin of one remained unidentified (Fig. 1B). Nevertheless, the similarity ranging from 24 to 54% is relatively low, signifying the specific sequence of the CelGH3\_f17 protein.

Phyre2 identified a total of 51 homologous templates with a confidence level of 100% (Supplementary Table S2). These templates exhibited a sequence identity ranging from 51% down to 15%. Among these templates, the crystal structure of NagZ from *Pseudomonas aeruginosa*<sup>38</sup> was identified as the most closely related, sharing a sequence identity of 51% with the target protein. Consequently, this high-confidence template, NagZ, was utilized as the primary basis for modelling the three-dimensional structure of the CelGH3\_f17 protein. The 214 residues (93%) of the CelGH3\_f17 protein were aligned to the template crystal structure of NagZ (Fig. 2D). In the predicted structure of the CelGH3\_f17 protein, the conserved amino acid residues of aspartic acid D187, glutamic acid E180, leucine L162, leucine L169, histidine H183, and asparagine N190 (Fig. 2A,C) form the pocket-shaped substrate recognition domain (Fig. 2C). The amino acid residues, aspartic acid D28, glutamic acid E51, arginine R213, and histidine H47 are predicted as the enzymatic active sites among the conserved amino acids (Fig. 2B). In addition, the aspartic acid D28 and glutamic acid E51 are the predicted catalytic active sites. The aspartic acid residue often acts as the nucleophile, and it is likely that E51 serves as the proton donor.

A



B

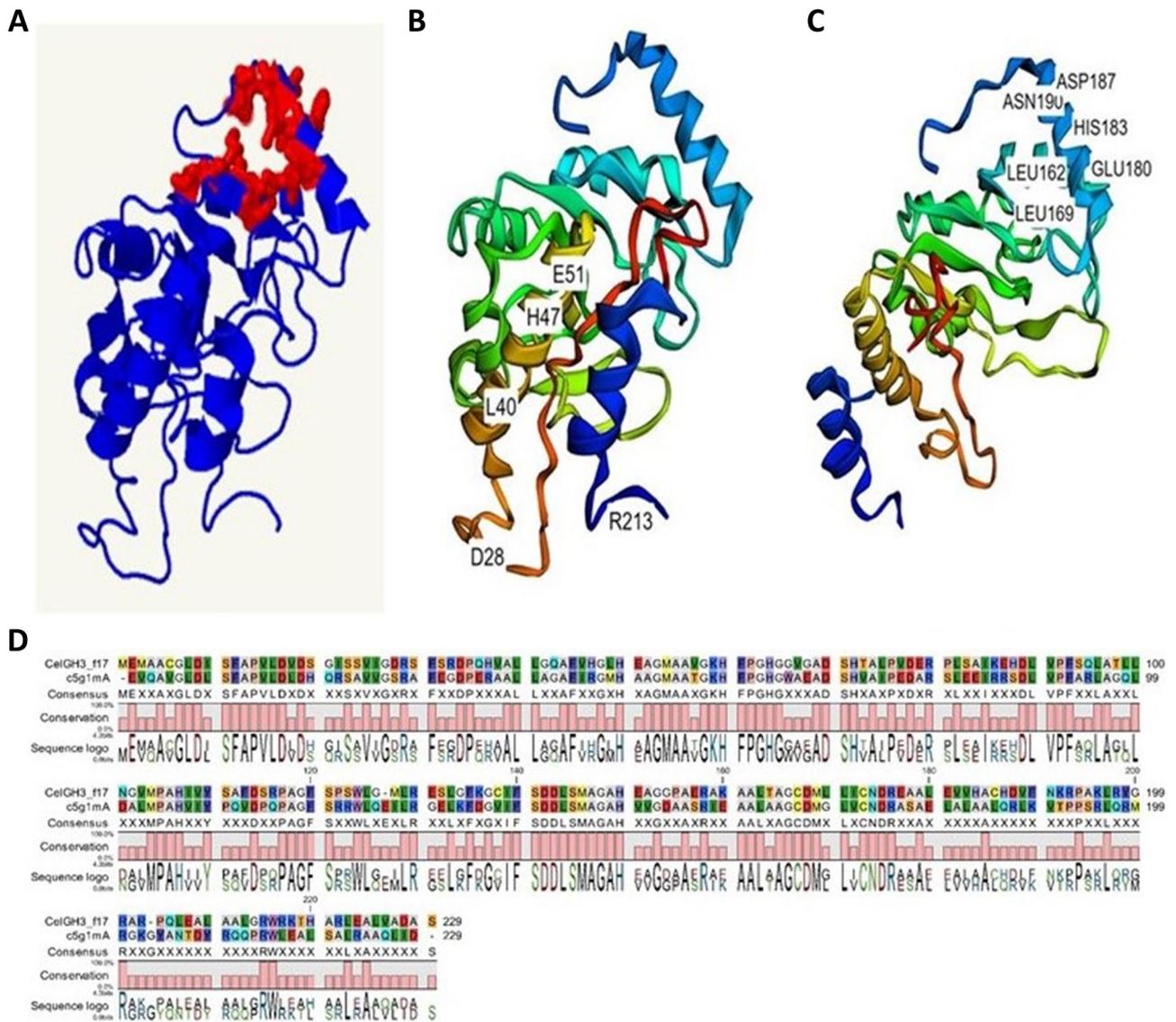


**Figure 1.** Sequence analysis of CelGH3\_f17 (A) The predicted modular architecture of CelGH3\_f17. (B) The phylogenetic tree was constructed using the Neighbor-Joining method<sup>80</sup>, based on the blast analysis of CelGH3\_f17 with 31 amino acid sequences of proteins in the protein data bank. Branch reliability is indicated by the percentage of 500 bootstrap replicates supporting each cluster of taxa next to the branches. The evolutionary study was performed using MEGA11 software<sup>75</sup>.

### Expression and purification of the CelGH3\_f17 protein

To characterize the biochemical properties of the CelGH3\_f17 protein, it was heterologously expressed with an N-terminal His-tag fusion protein using the pET-28a expression system under the control of the T7 lac promoter in *E. coli* BL21. A black halo around the colonies of CelGH3\_f17 on the agar plate indicated the presence of  $\beta$ -glucosidase activities (Fig. 3A).

The recombinant CelGH3\_f17 protein was secreted and expressed in a soluble form. The molecular mass of the purified enzyme was estimated at approximately 25 kDa as a single band by SDS-PAGE, which is consistent with the predicted molecular weight (Fig. 3B).



**Figure 2.** The predicted three-dimensional (3D) structural model of the CelGH3\_f17 protein (A). The predicted active sites of the CelGH3\_f17 protein (B). The predicted substrate recognition-pocket site of the CelGH3\_f17 protein (C). The pairwise alignment of the CelGH3\_f17 with c5g1mA, Crystal structure of NagZ from *P. aeruginosa* (D).

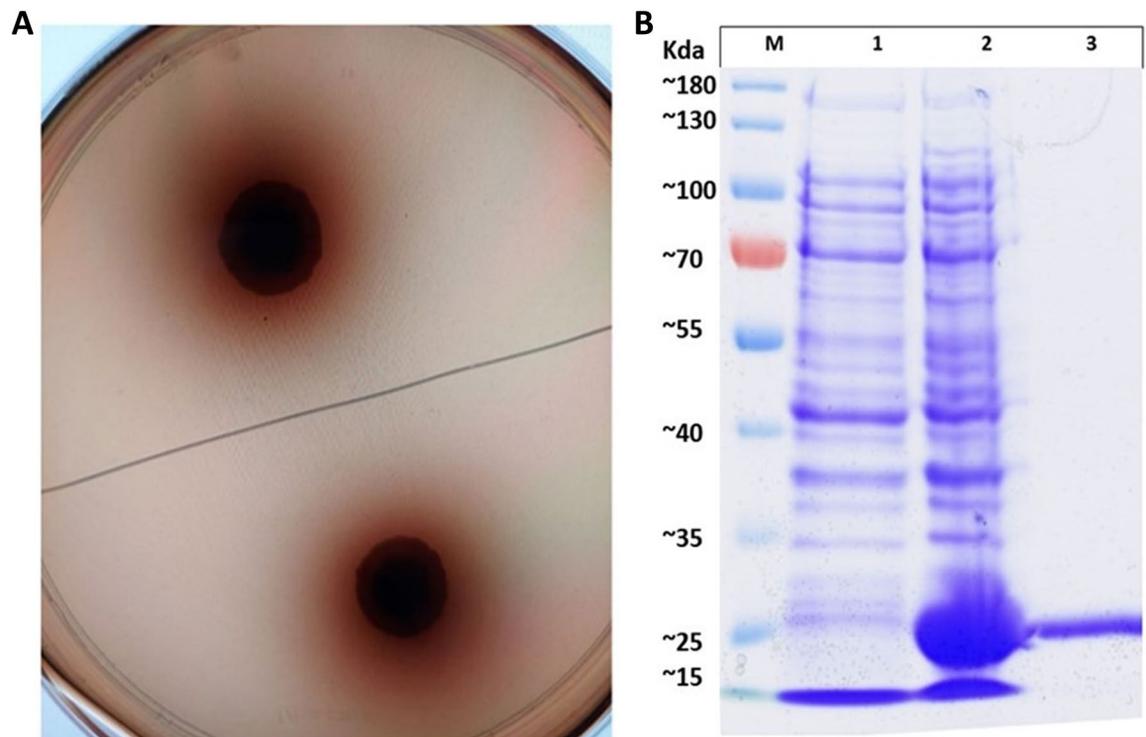
### Characterization of the $\beta$ -glucosidase activity of the CelGH3\_f17 protein

The enzyme encoded by CelGH3\_f17 displayed its highest level of  $\beta$ -glucosidase activity at a temperature of 50 °C. Remarkably, even within a range spanning from 30 to 60 °C, this enzyme retained more than 50% of its relative activity (Fig. 4A). The enzyme's stability was assessed over time, and it exhibited high stability at 37 °C and 50 °C, as evidenced by the gradual decrease in residual activity over a 16 h period. At 70 °C, the enzyme's residual activity decreased significantly (Fig. 4B).

The purified  $\beta$ -glucosidase peak hydrolytic activity was observed at pH8.5. Furthermore, it was observed that this enzyme exhibited a substantial level of relative activity exceeding 70% across a broad pH range, reaching as high as pH 10.5 (Fig. 4C). Further investigation into the stability of  $\beta$ -glucosidase at various pH levels showed that the enzyme was most stable at pH 7.5 and 8.5, maintaining high activity over an 16-h period. In contrast, at 10.5, the enzyme's activity reduced significantly over time (Fig. 4D).

The  $\beta$ -glucosidase activity of the CelGH3\_f17 was increased when supplemented with a salt concentration of 1 M (Fig. 5A). The enzyme still exhibited substantial activity, around 60% residual activity, even when exposed to a much higher salt concentration of 4 M. Furthermore, the  $\beta$ -glucosidase activity of the CelGH3\_f17 showed a noticeable improvement when supplemented with glucose, with enhancements observed up to a concentration of 300 mM. However, beyond the threshold of 300 mM glucose, there was a sharp decline in the  $\beta$ -glucosidase activity of the enzyme (Fig. 5B).

The addition of Fe did not affect the activity of this  $\beta$ -glucosidase. However,  $\text{Cu}^{2+}$ ,  $\text{Mg}^{2+}$  and  $\text{Mn}^{2+}$  ions resulted in a decreased activity of  $\beta$ -glucosidase, while,  $\text{Al}^{3+}$  and  $\text{K}^{1+}$  ions had little impact on the activity of the enzyme



**Figure 3.** (A) Agar plate  $\beta$ -glucosidase assay of CelGH3\_f17. (B) SDS-PAGE analysis of the purified, recombinant CelGH3\_f17 protein. M, marker proteins; lane 1, IPTG-uninduced crude extract of BL21 (DE3) carrying pET-CelGH3\_f17; lane 2, IPTG-induced crude extract of BL21 (DE3) carrying pET-CelGH3\_f17; lane 3, Ni-NTA chromatography purified extract of BL21 (DE3) carrying pET-CelGH3\_f17.

(Table 1). Furthermore, Glycerol and SDS enhanced the  $\beta$ -glucosidase activity of the CelGH3\_f17 protein, while, Mercaptoethanol and Triton X-100 decreased the activity considerably (Table 2).

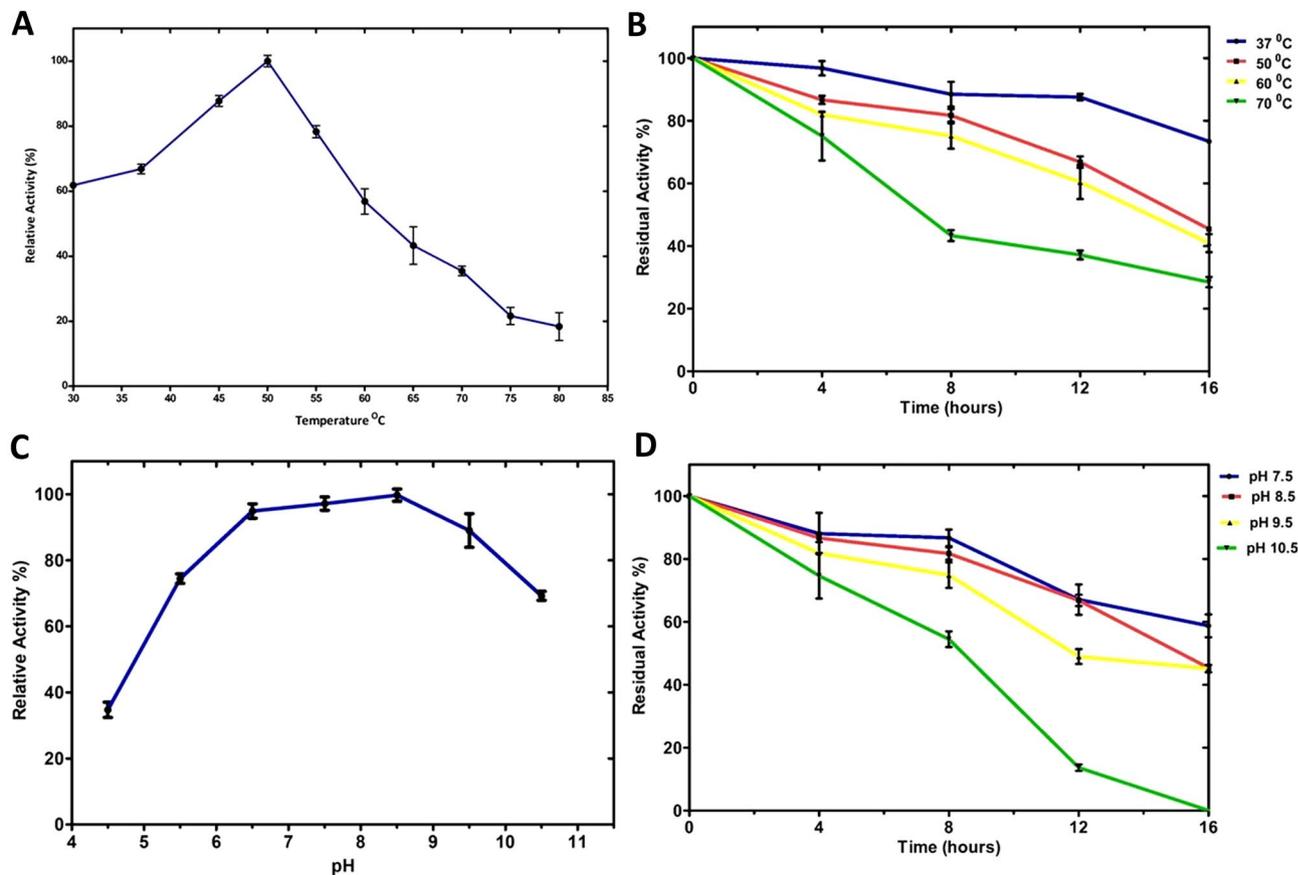
CelGH3\_f17 had the best hydrolyzing capacity against pNPG. CMC was hydrolyzed at 30%, but there was no hydrolytic activity on xylan (Table 3).

## Discussion

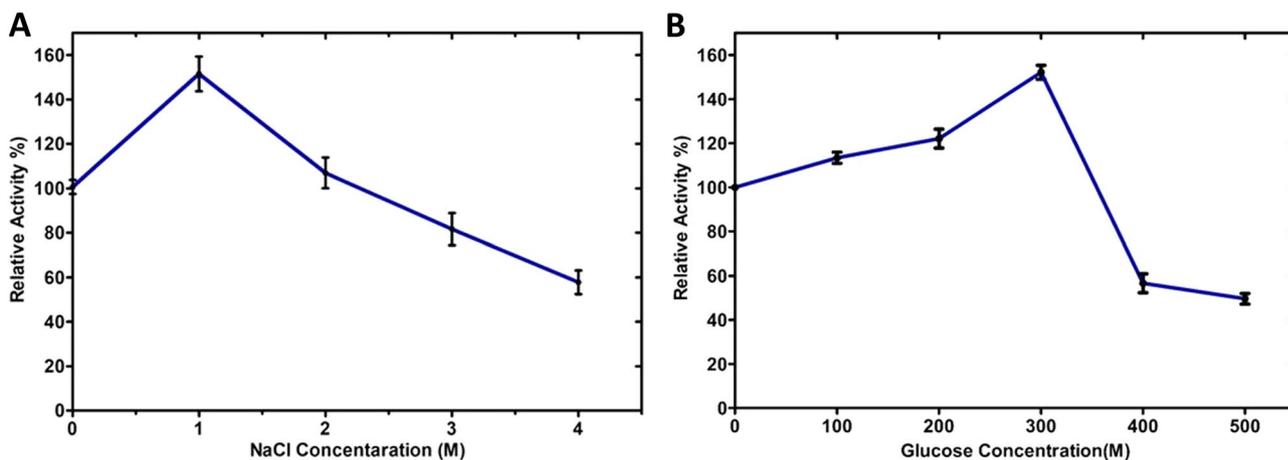
Enzyme discovery via functional metagenomics, followed by biochemical characterization, is critical for linking genes and gene sequences to practical applications in industry<sup>39</sup>. Thus, in this study, we report the molecular cloning, expression in *E. coli*, and functional characterization of the gene product of the CelGH3\_f17, which was isolated from our previous metagenomics libraries constructed from the Ethiopian soda lakes<sup>37</sup>.

The amino acid sequence analysis places CelGH3\_f17 among the GH3, which includes  $\beta$ -glucosidases, N-acetylglucosaminidases,  $\beta$ -xylosidases, and  $\alpha$ -arabinofuranosidases<sup>5</sup>. The CelGH3\_f17 encoded for 229 amino acids corresponding to an approximate protein molecular weight of 25 kDa. The modest amino acid sequence length and lower molecular weight of CelGH3\_f17 compared to, e.g., the Cen502  $\beta$ -glucosidase identified from *Bursaphelenchus xylophilus*<sup>9</sup>, and a novel GH3  $\beta$ -glucosidase<sup>40</sup> illustrates the wide-ranging structural and functional diversity of GH3<sup>41</sup>. This particular attribute of CelGH3\_f17 could be due to its evolutionary adaptation to the Ethiopian soda lake environment where it was discovered. To survive in extreme environments such as Soda lakes, microorganisms have evolved various adaptive strategies<sup>42</sup>, particularly in their protein structures and compositions. Enzymes such as CelGH3\_f17, with their smaller size, may have developed to function optimally in such niches, possibly conferring benefits such as increased efficiency in their synthesis and maintenance<sup>43–45</sup>.

The sequence relatedness of CelGH3\_f17 to the other proteins was low, with a maximum of 54% to beta-hexosaminidase from *Pseudomonas aeruginosa* PAO1<sup>38</sup>, where the lowest hit was 24% to periplasmic beta-glucosidase from *P. aeruginosa* PAO1<sup>46</sup>. This indicates that CelGH3\_f17 might be a new enzyme. This distinctness could be due to a unique evolutionary path or to specific adaptations in its enzymatic functions. The active site of GH3 families is composed of a well-conserved pair of aspartic acid residues that act as catalytic nucleophile and glutamic acid residues that act as catalytic acid/base<sup>4,5,47</sup>. GH3 enzymes also have three additional highly conserved amino acid residues (Arg, Lys, and His) that are predicted to be involved in catalysis in addition to its two critical residues<sup>4</sup>. Despite the fact that CelGH3\_f17 showed a low sequence relation to the other GH3  $\beta$ -glucosidases, it shared some common features with other GH3 enzymes, e.g. a conserved aspartate residue which acts as a catalytic nucleophile across different species, and the position of the general acid/base residue, which is phylogenetically variable<sup>41</sup>. Furthermore, the crystal structure of NagZ in *P. aeruginosa*, which served



**Figure 4.** The CelGH3\_f17's  $\beta$ -glucosidase activity and stability analysis. (A) The relative activity of the enzyme across a range of temperatures, demonstrating optimal activity. (B) Residual activity of the enzyme over a 16-h period at various temperatures, demonstrating the stability (C) The relative activity of the enzyme across a range of pHs, demonstrating optimal activity. (D) Residual activity of the enzyme over a 16-h period at various pHs demonstrating the stability of CelGH3\_f17 over time.



**Figure 5.** Effect of NaCl and glucose concentrations on the CelGH3\_f17's  $\beta$ -Glucosidase activity. (A) Relative enzyme activity at varying molar concentrations of NaCl. (B) Relative enzyme activity at different molar concentrations of glucose.

as the model for the CelGH3\_f17 enzyme, is a monomer. Based on this, we infer that CelGH3\_f17 may also be monomeric. However, additional research is required to definitively determine its oligomeric state.

We used an agar plate that contained an esculin and ferric ammonium citrate to confirm the  $\beta$ -glucosidase activity of the cloned CelGH3\_f17 gene. Esculin is comprised of a glucose molecule attached to esculin that, when released by hydrolysis, combines with ferric ammonium citrate in the medium to generate a black precipitate. The black precipitate detected suggests the presence of glucosidase enzymatic activity<sup>48</sup>.

Metal ions	Enzyme activity (U)	Relative activity (%)
Control	53.6 ± 0.4	100.0
CuCl <sub>2</sub>	25.7 ± 2.8*	48.0
FeCl <sub>3</sub>	53.2 ± 1.8	99.2
KCl	43.4 ± 6.2	80.9
CaCl <sub>2</sub>	37.6 ± 0.3*	70.0
ZnCl <sub>2</sub>	41.7 ± 3.0	77.7
MnCl <sub>2</sub>	23.7 ± 0.7*	44.2
MgSO <sub>4</sub>	35.3 ± 0.9*	65.8
AlCl <sub>3</sub>	48.3 ± 3.6	89.9

**Table 1.** The effect of metal ions on the  $\beta$ -glucosidase activity of the CelGH3\_f17. The enzyme activity is given with mean and standard error. Statistical significance between treatments and a control group was assessed using post-hoc Tukey's Honestly Significant Difference (HSD) test. Differences were considered statistically significant at a p-value of < 0.05. \*Significant difference of the treatment with control.

Surfactant and chelating agents(10mM)	Enzyme activity (U)	Relative activity (%)
Control	53.6 ± 0.4	100
EDTA	51.4 ± 1.3	96
SDS	76.6 ± 2.0*	143
DMSO	33.2 ± 2.2*	62
Glycerol	56.3 ± 0.6*	105
Mercaptoethanol	7.5 ± 0.1*	14
Triton X-100	12.8 ± 0.8*	24

**Table 2.** The effect of surfactant and chelating agents on the  $\beta$ - activity of CelGH3\_f17. The enzyme activity is given with mean and standard error. Statistical significance between treatments and a control group was assessed using post-hoc Tukey's Honestly Significant Difference (HSD) test. Differences were considered statistically significant at a p-value of < 0.05. \*Significant difference of the treatment with control.

Substrate	Linkage of a glycosyl group	Enzyme activity (U)
pNPG	$\beta$ Glc	53.6 ± 0.4
CMC	$\beta$ (1,4) Glc	16.0 ± 1.5*
Xylan	$\beta$ (1,3) Glc	0.0 ± 0.0*

**Table 3.** Effects of the  $\beta$ -glucosidase on different substrates. The enzyme activity is given with mean and standard error. Statistical significance between treatments and a control group was assessed using post-hoc Tukey's Honestly Significant Difference (HSD) test. Differences were considered statistically significant at a p-value of < 0.05. \*Significant difference of the treatment with control.

The fact that the recombinant CelGH3\_f17 protein was secreted in soluble form is attractive for industrial applications, since purification becomes less expensive and time-consuming<sup>49</sup>. In addition, extracellular enzyme production is more effective than intracellular enzyme production since it improves protein folding and reduces proteolytic degradation<sup>50,51</sup>.

The CelGH3\_f17 exhibited optimum activity in alkaline pHs with a maximum hydrolytic activity at pH 8.5 and temperature 50 °C. In line with other studies, the  $\beta$ -glucosidase, Cen502, was shown to exhibit optimal activity at 38 °C and pH 8.0, and is stable across a pH range of 7.0 to 9.0<sup>9</sup>. An engineered  $\beta$ -glucosidase from a previous study was found with optimal activity at 75 °C and pH 4.5<sup>52</sup>.

The optimal temperature for CelGH3\_f17  $\beta$ -glucosidase is consistent with many  $\beta$ -glucosidases in literature and the wide range of temperature over which the enzyme retains significant activity suggests it may be robust for various biotechnological applications. The activity of enzymes in elevated temperatures is a crucial factor in the breakdown process of biomass<sup>47</sup>. However, the loss of stability at higher temperatures is a common characteristic among enzymes due to thermal denaturation, which affects the enzyme's tertiary and quaternary structures<sup>53</sup>. The CelGH3\_f17  $\beta$ -glucosidase had a higher activity and stability at alkaline pH compared to most other  $\beta$ -glucosidases characterized. Most  $\beta$ -glucosidases that have been identified to date exhibit maximal activity in the acidic pH range (pH 4–6.5), retaining only minimal activity or displaying instability in mildly alkaline conditions<sup>54–57</sup>. The CelGH3\_f17  $\beta$ -glucosidase's ability to retain considerable activity up to pH 10.5 indicates its potential usefulness in slightly alkaline conditions where other enzymes might not perform well. The reduced

stability at higher pH values has in previous studies been attributed to the disruption of ionic bonds and electrostatic interactions within the enzyme's structure<sup>52</sup>.

$\beta$ -glucosidases have great application prospects in many industrial fields, including aroma improvement of juice and wine, and lignocellulosic biomass conversion in biofuel production<sup>58</sup>. Typically, lignocellulosic biomass conversion takes place at high temperatures in the presence of alkaline conditions<sup>59</sup>. The neutralisation of these alkalines results in the formation of salts, which requires massive amounts of water and energy to be removed for subsequent processes to proceed. From previous studies, it is known that the alkaline conditions needed have a negative effect on the activity and the stability of most known enzymes<sup>60</sup>. For the enzyme reported here, the  $\beta$ -glucosidase activity remained consistent even when exposed to high salt concentrations, and the activity even increased with the addition of 1 M NaCl. This reflects that the CelGH3\_f17 gene comes from metagenomic libraries derived from Ethiopian soda lakes, which hold extreme conditions characterized by high alkalinity and salinity<sup>22,61</sup>.

One commonly observed unfavourable characteristic that impedes the efficient breakdown of cellulose via enzymatic hydrolysis is glucose, which competitively inhibits the activity of  $\beta$ -glucosidase<sup>58</sup>. This results in the accumulation of cellobiose and oligosaccharides, which repress endoglucanase and exoglucanase activities, ultimately impeding the entire process of cellulose hydrolysis<sup>14,50,62</sup>. The majority of well-known GH3 are inhibited when exposed to glucose concentrations lower than 100 mM<sup>58,63,64</sup>. In this study, up to 300 mM of glucose rather increased the  $\beta$ -glucosidase activity of CelGH3\_f17. This strong glucose tolerance offers an additional advantage in mitigating the catabolic inhibition imposed on endoglucanase and cellobiohydrolase due to the accumulation of cellobiose during the bioconversion process.

It has been proposed that metal ions are responsible for inhibiting  $\beta$ -glucosidases' catalytic activity<sup>65</sup>. Most of the supplemented metal ions did not have a significant impact on the activity of  $\beta$ -glucosidases evaluated here, with the exception of the Cu<sup>2+</sup> and Mn<sup>2+</sup> ions. However, these metal ions have been found to significantly inhibit the activity of  $\beta$ -glucosidases evaluated in previous studies as well<sup>14,66,67</sup>. This result implies that the CelGH3\_f17 protein active catalytic sites may include thiol groups, which may be necessary for the preservation of the 3D structure of the active protein. Disulfide bonds are crucial to maintain the natural shape of a protein thereby controlling its biological activity<sup>68–70</sup>. Disulfide bonds can be chemically reduced by reducing agents like mercaptoethanol, and these modifications may result in changes in the structure of the protein, which have an impact on their activity<sup>68</sup>. Indeed, mercaptoethanol decreased the  $\beta$ -glucosidases which indicates that disulfide bonds are essential for the activity of the  $\beta$ -glucosidases of CelGH3\_f17. Moreover, the fact that the chelating chemical EDTA did not affect the activity of  $\beta$ -glucosidase indicates that the CelGH3\_f17 protein is not a metalloprotein.

The moderate sequence similarity of the CelGH3\_f17 protein to other known GH3 enzymes indicates that the CelGH3\_f17 protein may possess unique structural features conducive to its stability and activity profile. The structural modeling based on the homologous NagZ<sup>71</sup> suggests that CelGH3\_f17 has a well-defined active site with specific amino acids contributing to its catalytic process. This information could be valuable for further enzyme engineering efforts aimed at enhancing the desirable traits of CelGH3\_f17. The  $\beta$ -glucosidase enzyme of CelGH3\_f17, with optimal activity at 50 °C, pH 8.5, and enhanced activity in the presence of salt and glucose, has promising applications in several industrial processes due to its unique properties. Primarily, the enzyme's thermostability and alkaline tolerance suggest its potential use in the food and bioenergy industries<sup>72</sup>. The CelGH3\_f17  $\beta$ -glucosidases halotolerance and glucose enhancement up to a certain threshold is an advantageous trait not commonly observed in  $\beta$ -glucosidases<sup>58</sup>. Its tolerance to high glucose concentrations is particularly valuable in biofuel production, as it overcomes the common challenge of product inhibition in this process<sup>63,64</sup>. In the food industry, glucose tolerance  $\beta$ -glucosidase can play a vital role in processing, particularly in the degradation of cellulose or soy isoflavone in organic solvents<sup>73</sup>.

## Conclusion

In this study we cloned, expressed, and characterised a  $\beta$ -glucosidase, CelGH3\_f17, from a previously constructed metagenomics libraries derived from an Ethiopian soda lake. This enzyme belongs to glycoside hydrolase family 3 (GH3) and performed optimum activity at 50 °C and pH 8.5. It was particularly resistant to high salt concentrations and glucose. Furthermore, the majority of metal ions and solvents examined showed little effect on its activity. Overall, CelGH3\_f17 has favourable properties such as stability at alkaline pH, and high salt and glucose tolerance, making it a good choice for industrial applications, particularly in biomass degradation.

## Methods

### Screening and identification of CelGH3\_f17 from a soda lake library

The CelGH3\_f17 gene was identified from our previously constructed metagenomic libraries derived from Ethiopian soda lakes<sup>37</sup>. Briefly, metagenomics libraries were constructed from the water and sediment samples of the soda lakes. Following, functional screening of these libraries for  $\beta$ -glucosidase activity, clones exhibiting  $\beta$ -glucosidase activity were sequenced, and a comprehensive bioinformatics analysis were conducted. A sequence named CelGH3\_f17 due to its success in cloning and high hit rate, was chosen for further investigation in this study.

### Sequence and structure analysis of CelGH3\_f17 protein

Domain architecture and conserved domains of the CelGH3\_f17 protein were identified with the Simple Modular Architecture Research Tool (SMART) program (<http://smart.embl-heidelberg.de>)<sup>74</sup>. To identify potential homologs of the CelGH3\_f17 protein, we conducted a BLASTP search against the Protein Data Bank (PDB) using the National Center for Biotechnology Information (NCBI) platform, employing the default algorithm settings. The amino acid sequence of the CelGH3\_f17 protein served as the query. Following this, the resulting

FASTA sequences were incorporated into MEGA software<sup>75</sup>. The Neighbor-Joining method was employed to deduce the evolutionary history of these sequences. The three-dimensional (3D) structural model of CelGH3\_f17 protein was predicted utilizing the Protein Homology/analogy Recognition Engine (Phyre2; V 2.0) web server<sup>76</sup>. This online program enabled the prediction of the tertiary structure of the CelGH3\_f17 protein. Furthermore, ExPASy's Compute pI/MW tool<sup>77</sup>, was used to calculate the theoretical pI (isoelectric point) and MW (molecular weight) of the CelGH3\_f17 protein.

### Cloning of the $\beta$ -glucosidase gene

The CelGH3\_f17 was synthesized and codon-optimized for expression in *E. coli* using Twist Bioscience's gene synthesis service. Briefly, the pET-28a (+) vector was prepared for cloning by digestion with BamHI and XhoI restriction enzymes, followed by purification. The synthesized CelGH3\_f17 was similarly digested and purified. The gene was then ligated into the prepared vector using T4 DNA ligase, and the ligation mixture was transformed into chemically competent *E. coli* DH5 $\alpha$  cells. Chemically competent DH5 $\alpha$  cells were thawed gently on ice, and the ligation mixture with the CelGH3\_f17 insert was added, followed by careful mixing to avoid cellular damage. The cells were then incubated on ice for 30 min to allow DNA adsorption, succeeded by a 45 s heat shock at 42 °C to increase membrane permeability and facilitate DNA uptake, and subsequently cooled on ice for 2 min. SOC medium was added to the transformation mixture for cell recovery, and the mixture was incubated at 37 °C with shaking for 1 h to enable antibiotic resistance gene expression. The transformed cells were then plated on LB agar plates containing kanamycin and incubated overnight at 37 °C. To confirm the  $\beta$ -glucosidase activity, the transformed cells were plated onto LB agar plates containing esculin hydrate (0.1%), ferric ammonium citrate (0.25%), and chloramphenicol (12.5  $\mu$ g/ml). Colonies with  $\beta$ -glucosidase activity was identified by observing a surrounding black halo after 20–24 h incubation at 37 °C<sup>78</sup>.

### Expression and purification of the recombinant CelGH3\_f17 protein

A single colony of transformant was inoculated into LB media supplemented with kanamycin (50  $\mu$ g/ml) and incubated overnight at 37 °C with 220 rpm shaking. Then, this starter culture was used to inoculate a larger culture (1 L) at 1:50 dilution and again incubated at 37 °C with 220 rpm shaking. The expression was induced by adding 0.5 mM IPTG (Sigma); when the cell density reached an  $A_{600}$  of 0.4–0.6. The culture was then incubated for an additional 3.5 h at 37 °C. Ni-NTA resin (Qiagen) was used to purify the protein following the pET System manufacturer's instructions. Briefly, an appropriate amount of Ni-NTA resin slurry was added to a gravity-flow column. Then, a binding buffer was passed through the column to equilibrate the resin. Thereafter, the filtered soluble protein was added to the equilibrated resin and incubated for 2 h at 4 °C. Then, after incubation, the flow-through was allowed to pass and the resin was washed with wash buffer three times. Finally, the purified protein bonded to the resin was eluted with an elution buffer. The purity and integrity of the protein were assessed by SDS-PAGE. Briefly, for SDS-PAGE analysis, each sample including IPTG-uninduced crude extract, IPTG-induced crude extract, and Ni-NTA chromatography purified extract, were mixed with SDS-PAGE loading buffer, heated at 95 °C for 5 min for denaturation, and then loaded onto an 12% SDS-PAGE gel. Electrophoresis was performed at a 120 voltage until the dye front reached the bottom of the gel. After electrophoresis, the gel was stained with Coomassie Brilliant Blue and destined to visualize protein bands.

### Enzyme activity assays

The  $\beta$ -glucosidase activity was measured by determining the hydrolysis of p-nitrophenyl- $\beta$ -D-glucopyranoside (pNPG; Sigma) using the initial rate of accumulation of coloured reaction product as described by<sup>9</sup>. One hundred eighty microliter of 5 mM pNPG substrate was diluted in 50 mM Tris HCl buffer (pH 8.5) and mixed with a 20  $\mu$ l aliquot of the acquired enzyme. Then, the mixture was incubated at 50 °C for 10 min, and the reaction was terminated by adding 100  $\mu$ l of ice-cold 0.5 M Na<sub>2</sub>CO<sub>3</sub>. The release of p-nitrophenol (pNP) via enzymatic hydrolysis was indicated by the appearance of a yellow colour and the absorbance was measured with a UV/Vis microplate spectrophotometer (Multiskan GO, Thermo Scientific) at 405 nm. One unit (U) of  $\beta$ -glucosidase activity was defined as the amount of enzyme that released 1  $\mu$ mol of pNP per minute from the substrate. All experiments were performed in three technical replicates. Statistical significance between treatments and a control group was assessed using post-hoc Tukey's Honestly Significant Difference (HSD) test. Differences were considered statistically significant at a p-value of < 0.05.

Relative activity is as follows:

$$\text{Relative Activity (\%)} = (\text{Activity of sample (U)} / \text{Maximum enzyme activity (U)}) * 100$$

### Characterization of the CelGH3\_f17 protein

The effect of temperature on  $\beta$ -glucosidase activity was studied by assaying the enzyme in 50 mM Tris-HCl buffer (pH 8.5) at temperatures ranging from 30 to 80 °C. The effect of pH on  $\beta$ -glucosidase activity was investigated by incubating 180  $\mu$ l of a substrate prepared at different pH buffers with 20  $\mu$ l of an enzyme for 10 min at 50 °C. The substrate (5 mM pNPG) was prepared in 50 mM pH buffers ranging from 4.5 to 10.5. The buffers used were sodium acetate (pH 4.5–5.5), phosphate (pH 6.5), Tris-HCl (pH 7.5), and glycine-NaOH (pH 8.5–10.5). The stability of  $\beta$ -glucosidase at different temperatures was also evaluated. The enzyme was pre-incubated at temperatures of 37 °C, 50 °C, 60 °C, and 70 °C over a period of 16 h. Similarly, the pH stability of  $\beta$ -glucosidase was investigated by pre-incubating the enzyme in buffers with pH values of 7.5, 8.5, 9.5, and 10.5 at the optimal temperature for varying time points up to 16 h. After incubation, the residual activity of the enzyme was assessed as described above.

The effects of metal ions on the  $\beta$ -glucosidase activity were studied by introducing the metal ions ( $\text{CuCl}_2$ ,  $\text{FeCl}_3$ ,  $\text{KCl}$ ,  $\text{CaCl}_2$ ,  $\text{ZnCl}_2$ ,  $\text{MnCl}_2$ ,  $\text{MgSO}_4$  and  $\text{AlCl}_3$ ) into the reaction at a final concentration of 5 mM.

The effect of glucose on the activity of  $\beta$ -glucosidase was studied ranging from 0 to 500 mM concentrations. In the meantime, the effect of salt on the  $\beta$ -glucosidase activity was studied by supplementing salt, ranging from 0 to 4 M concentrations.

The effect of surfactants and inhibitors on  $\beta$ -glucosidase activity was studied by directly incorporating these in the reaction to a final concentration of 10 mM. The studied surfactants and inhibitors include mercaptoethanol, Ethylenediaminetetraacetic Acid (EDTA), Dimethyl Sulfoxide (DMSO), Glycerol, Sodium Dodecyl Sulphate (SDS), and Triton X-100.

The substrates *p*NPG, CMC, and Xylan were used to study the specificity of substrate on  $\beta$ -glucosidase activity. The enzyme assay of the tested polysaccharides, CMC, and Xylan was based on the determination of reducing sugars by the dinitrosalicylic acid reagent (DNS) method<sup>79</sup>. One unit of cellulase activity was defined as the amount of enzyme capable of breaking down CMC to produce 1  $\mu\text{mol}$  of glucose within a 1-min reaction. Similarly, one unit of xylanase activity was determined as the amount of enzyme that can release 1  $\mu\text{mol}$  of xylose per minute reaction.

## Data availability

Sequence data that supports the findings of this study has been deposited in NCBI's GenBank database and is publicly available with an accession number of OR883674 (<https://www.ncbi.nlm.nih.gov/nucleotide/OR883674>).

Received: 16 November 2023; Accepted: 25 April 2024

Published online: 01 May 2024

## References

- Lombard, V., Golaconda Ramulu, H., Drula, E., Coutinho, P. M. & Henrissat, B. The carbohydrate-active enzymes database (CAZY) in 2013. *Nucleic Acids Res.* **42**, 1 (2014).
- Davies, G. & Henrissat, B. Structures and mechanisms of glycosyl hydrolases. *Structure* **3**, 1 (1995).
- Ara, K. Z. G. *et al.* Characterization and diversity of the complete set of GH family 3 enzymes from *Rhodothermus marinus* DSM 4253. *Sci. Rep.* **10**, 1 (2020).
- Dodd, D., Kiyonari, S., Mackie, R. I. & Cann, I. K. O. Functional diversity of four glycoside hydrolase family 3 enzymes from the rumen bacterium *Prevotella bryantii* B14. *J. Bacteriol.* **192**, 1 (2010).
- Macdonald, S. S., Blaukopf, M. & Withers, S. G. N-acetylglucosaminidases from CAZy family GH3 are really glycoside phosphorylases, thereby explaining their use of histidine as an acid/Base catalyst in place of glutamic acid. *J. Biol. Chem.* **290**, 1 (2015).
- Qu, X. *et al.* Characterization of a GH3 halophilic  $\beta$ -glucosidase from *Pseudoalteromonas* and its NaCl-induced activity toward isoflavones. *Int. J. Biol. Macromol.* **164**, 1 (2020).
- Fry, S. C. Comprehensive natural products chemistry, volume 3: Carbohydrates and their derivatives including tannins. *Cell. Relat. Lignins. Phytochem.* **59**, 1 (2002).
- Karnchanat, A. *et al.* Purification and biochemical characterization of an extracellular  $\beta$ -glucosidase from the wood-decaying fungus *Daldinia eschscholzii* (Ehrenb.:Fr.) Rehm. *FEMS Microbiol. Lett.* **270**, 1 (2007).
- Zhang, L. *et al.* Identification and characterization of a novel  $\beta$ -glucosidase via metagenomic analysis of *Bursaphelenchus xylophilus* and its microbial flora. *Sci. Rep.* **7**, 1 (2017).
- Tiwari, R. *et al.* Bioprospecting of novel thermostable  $\beta$ -glucosidase from *Bacillus subtilis* RA10 and its application in biomass hydrolysis. *Biotechnol. Biofuels* **10**, 1 (2017).
- Mirete, S., Morgante, V. & González-Pastor, J. E. Functional metagenomics of extreme environments. *Curr. Opin. Biotechnol.* **38**, 1 (2016).
- Aguirre, A. *et al.* The production, properties, and applications of thermostable steryl glucosidases. *World J. Microbiol. Biotechnol.* **34**, 1–9 (2018).
- Costa, L. S. C. *et al.* Molecular dynamics gives new insights into the glucose tolerance and inhibition mechanisms on  $\beta$ -glucosidases. *Molecules* **24**, 1 (2019).
- Ariaeenejad, S. *et al.* A novel high glucose-tolerant  $\beta$ -glucosidase: Targeted computational approach for metagenomic screening. *Front. Bioeng. Biotechnol.* **8**, 1 (2020).
- Rashid, M. H. & Siddiqui, K. S. The stability of extracellular  $\beta$ -glucosidase from *Aspergillus niger* is significantly enhanced by non-covalently attached polysaccharides. *Folia Microbiol. (Praha)* **41**, 341–346 (1996).
- Horn, S. J., Vaaje-Kolstad, G., Westereng, B. & Eijsink, V. G. H. Novel enzymes for the degradation of cellulose. *Biotechnol. Biofuels* **5**, 1–12 (2012).
- Ersoy Omeroglu, E. *et al.* Microbial community of soda Lake Van as obtained from direct and enriched water, sediment and fish samples. *Sci. Rep.* **11**, 18364 (2021).
- Kumar, L., Awasthi, G. & Singh, B. Extremophiles: A novel source of industrially important enzymes. *Biotechnology* **10**, 1 (2011).
- Grant, W. D. & Jones, B. E. Bacteria, archaea and viruses of soda lakes. In *Soda Lakes of East Africa* 97–147 (Springer International Publishing, 2016). [https://doi.org/10.1007/978-3-319-28622-8\\_5](https://doi.org/10.1007/978-3-319-28622-8_5).
- Schagerl, M. & Renaut, R. W. Dipping into the Soda Lakes of East Africa. *Soda Lakes East Afr.* [https://doi.org/10.1007/978-3-319-28622-8\\_1](https://doi.org/10.1007/978-3-319-28622-8_1) (2016).
- Grant, W. D. & Sorokin, D. Y. Distribution and diversity of soda lake alkaliphiles. in *Extremophiles Handbook* (ed. Horikoshi, K.) 27–54 (2011). [https://doi.org/10.1007/978-4-431-53898-1\\_3](https://doi.org/10.1007/978-4-431-53898-1_3).
- Antony, C. P. *et al.* Microbiology of Lonar Lake and other soda lakes. *ISME J.* **7**, 1 (2013).
- Sorokin, D. Y., Banciu, H. L. & Muyzer, G. Functional microbiology of soda lakes. *Curr. Opin. Microbiol.* **25**, 1 (2015).
- Jeilu, O., Gessesse, A., Simachew, A., Johansson, E. & Alexandersson, E. Prokaryotic and eukaryotic microbial diversity from three soda lakes in the East African Rift Valley determined by amplicon sequencing. *Front. Microbiol.* **13**, 1 (2022).
- Demirjian, D. C., Moris-Varas, F. & Cassidy, C. S. Enzymes from extremophiles. *Curr. Opin. Chem. Biol.* **5**, 1 (2001).
- Niehaus, F., Bertoldo, C., Kähler, M. & Antranikian, G. Extremophiles as a source of novel enzymes for industrial application. *Appl. Microbiol. Biotechnol.* **51**, 1 (1999).
- Haile, G. & Gessesse, A. Properties of alkaline protease C45 produced by alkaliphilic *Bacillus* Sp. isolated from Chitu. *Ethiopian Soda Lake. J. Biotechnol. Biomater.* **02**, 1 (2012).
- Gessesse, A. & Gashe, B. A. Production of alkaline protease by an alkaliphilic bacteria isolated from an alkaline soda lake. *Biotechnol. Lett.* **19**, 1 (1997).
- Martins, R. F. *et al.* Starch-hydrolyzing bacteria from Ethiopian soda lakes. *Extremophiles* **5**, 1 (2001).

30. Berini, F., Casciello, C., Marcone, G. L. & Marinelli, F. Metagenomics: Novel enzymes from non-culturable microbes. *FEMS Microbiol. Lett.* **364**, 1–19 (2017).
31. Vitorino, L. C. & Bessa, L. A. Technological microbiology: Development and applications. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2017.00827> (2017).
32. Pabbathi, N. P. P. *et al.* Role of metagenomics in prospecting novel endoglucanases, accentuating functional metagenomics approach in second-generation biofuel production: A review. *Biomass Convers. Biorefin.* **11**, 1–28 (2021).
33. Robinson, S. L., Piel, J. & Sunagawa, S. A roadmap for metagenomic enzyme discovery. *Nat. Prod. Rep.* **38**, 1 (2021).
34. Lam, K. N., Cheng, J., Engel, K., Neufeld, J. D. & Charles, T. C. Current and future resources for functional metagenomics. *Front. Microbiol.* **6**, 1–8 (2015).
35. Li, Y. *et al.* Cloning and characterization of a new  $\beta$ -Glucosidase from a metagenomic library of Rumen of cattle feeding with *Miscanthus sinensis*. *BMC Biotechnol.* **14**, 1 (2014).
36. Wang, L., Zhang, G., Xu, H., Xin, H. & Zhang, Y. Metagenomic analyses of microbial and carbohydrate-active enzymes in the rumen of holstein cows fed different forage-to-concentrate ratios. *Front. Microbiol.* **10**, 1–14 (2019).
37. Jeilu, O., Simachew, A., Alexandersson, E., Johansson, E. & Gessesse, A. Discovery of novel carbohydrate degrading enzymes from soda lakes through functional metagenomics. *Front. Microbiol.* **13**, 1 (2022).
38. Acebrón, I. *et al.* Catalytic cycle of the N-acetylglucosaminidase NagZ from *Pseudomonas aeruginosa*. *J. Am. Chem. Soc.* **139**, 1 (2017).
39. de Morais Souto, B. *et al.* Functional screening of a Caatinga goat (*Capra hircus*) rumen metagenomic library reveals a novel GH3  $\beta$ -xylosidase. *PLoS One* **16**, 1 (2021).
40. Su, H. *et al.* High diversity of  $\beta$ -glucosidase-producing bacteria and their genes associated with scleractinian corals. *Int. J. Mol. Sci.* **22**, 3523 (2021).
41. Drula, E. *et al.* The carbohydrate-active enzyme database: Functions and literature. *Nucleic Acids Res.* <https://doi.org/10.1093/nar/gkab1045> (2021).
42. Shu, W. S. & Huang, L. N. Microbial diversity in extreme environments. *Nat. Rev. Microbiol.* **20**, 219–235 (2021).
43. Panja, A. S., Maiti, S. & Bandyopadhyay, B. Protein stability governed by its structural plasticity is inferred by physicochemical factors and salt bridges. *Sci. Rep.* **10**, 1–9 (2020).
44. Babu, P., Chandel, A. K. & Singh, O. V. Survival Mechanisms of Extremophiles. 9–23. [https://doi.org/10.1007/978-3-319-12808-5\\_2](https://doi.org/10.1007/978-3-319-12808-5_2) (2015).
45. Reed, C. J., Lewis, H., Trejo, E., Winston, V. & Evilia, C. Protein adaptations in archaeal extremophiles. *Archaea* **2013**, (2013).
46. Mahasanen, K. V. *et al.* Catalytic cycle of glycoside hydrolase BglX from *Pseudomonas aeruginosa* and its implications for biofilm formation. *ACS Chem Biol* **15**, 189–196 (2020).
47. Colabardini, A. C. *et al.* Expression of two novel  $\beta$ -Glucosidases from *Chaetomium atrobrunneum* in *Trichoderma reesei* and characterization of the heterologous protein products. *Mol. Biotechnol.* **58**, 1 (2016).
48. Mahadevan, P. R. & Eberhart, B. The  $\beta$ -glucosidase system of *Neurospora crassa*. *Arch. Biochem. Biophys.* **108**, 1 (1964).
49. Sørensen, H. P. & Mortensen, K. K. Soluble expression of recombinant proteins in the cytoplasm of *Escherichia coli*. *Microbial. Cell Fact.* **4**, 1 (2005).
50. Chamoli, S., Kumar, P., Navani, N. K. & Verma, A. K. Secretory expression, characterization and docking study of glucose-tolerant  $\beta$ -glucosidase from *B. subtilis*. *Int. J. Biol. Macromol.* **85**, 1 (2016).
51. Patel, A. K., Singhanian, R. R. & Pandey, A. Production, Purification, and Application of Microbial Enzymes. In *Biotechnology of Microbial Enzymes: Production, Biocatalysis and Industrial Applications*. <https://doi.org/10.1016/B978-0-12-803725-6.00002-9> (2017).
52. Xia, W. *et al.* Engineering a highly active thermophilic  $\beta$ -glucosidase to enhance its pH stability and saccharification performance. *Biotechnol. Biofuels* **9**, 1–12 (2016).
53. Lenz, F., Zurek, P., Umlauf, M., Tozakidis, I. E. P. & Jose, J. Tailor-made  $\beta$ -glucosidase with increased activity at lower temperature without loss of stability and glucose tolerance. *Green Chem.* **22**, 2234–2243 (2020).
54. Biver, S., Stroobants, A., Portetelle, D. & Vandenbol, M. Two promising alkaline  $\beta$ -glucosidases isolated by functional metagenomics from agricultural soil, including one showing high tolerance towards harsh detergents, oxidants and glucose. *J. Ind. Microbiol. Biotechnol.* **41**, 1 (2014).
55. Bergmann, J. C. *et al.* Discovery of two novel  $\beta$ -glucosidases from an Amazon soil metagenomic library. *FEMS Microbiol. Lett.* **351**, 1 (2014).
56. Monteiro, L. M. O. *et al.* A highly glucose tolerant  $\beta$ -glucosidase from *Malbranchea pulchella* (MpBg3) enables cellulose saccharification. *Sci. Rep.* **10**, 1 (2020).
57. Sun, N. *et al.* Characterization of a novel recombinant halophilic  $\beta$ -glucosidase of *Trichoderma harzianum* derived from Hainan mangrove. *BMC Microbiol.* **22**, 185 (2022).
58. Salgado, J. C. S., Meleiro, L. P., Carli, S. & Ward, R. J. Glucose tolerant and glucose stimulated B-glucosidases—A review. *Bioresour. Technol.* **267**, 1 (2018).
59. Zhu, D. *et al.* Recent development of extremophilic bacteria and their application in biorefinery. *Front. Bioeng. Biotechnol.* **8**, 1–18 (2020).
60. Sweeney, M. D. & Xu, F. Biomass converting enzymes as industrial biocatalysts for fuels and chemicals: Recent developments. *Catalysts* **2** (2012).
61. Grant, W. D. Alkaline environments and biodiversity. *Encycl. Life Support Syst.* (2006).
62. Singhanian, R. R., Patel, A. K., Sukumaran, R. K., Larroche, C. & Pandey, A. Role and significance of beta-glucosidases in the hydrolysis of cellulose for bioethanol production. *Bioresour. Technol.* **127**, 1 (2013).
63. Teugjas, H. & Våljamäe, P. Selecting  $\beta$ -glucosidases to support cellulases in cellulose saccharification. *Biotechnol. Biofuels* **6**, 1 (2013).
64. Kim, I. J., Bornscheuer, U. T. & Nam, K. H. Biochemical and Structural Analysis of a Glucose-Tolerant  $\beta$ -Glucosidase from the Hemicellulose-Degrading Thermoanaerobacterium saccharolyticum. *Molecules* **27**, 1 (2022).
65. Chen, T. *et al.* Enhancing catalytic performance of  $\beta$ -glucosidase via immobilization on metal ions chelated magnetic nanoparticles. *Enzyme Microb. Technol.* **63**, (2014).
66. Mao, X., Hong, Y., Shao, Z., Zhao, Y. & Liu, Z. A novel cold-active and alkali-stable  $\beta$ -glucosidase gene isolated from the marine bacterium *martelella mediterranea*. *Appl. Biochem. Biotechnol.* **162**, 1 (2010).
67. Liu, D. *et al.* Characterization of a thermostable  $\beta$ -glucosidase from *Aspergillus fumigatus* Z5, and its functional expression in *Pichia pastoris* X33. *Microb. Cell Fact.* **11**, 1 (2012).
68. Cao, P., Wang, L., Wang, Y., Zhou, N. & Chen, Y. Alkali-tolerant  $\beta$ -glucosidase produced by newly isolated *Aspergillus fumigatus* WL002 from rotten wood. *Int. Biodeterior. Biodegrad.* **105**, 1 (2015).
69. Tang, Z. *et al.* Cloning and expression of A. oryzae  $\beta$ -glucosidase in *Pichia pastoris*. *Mol. Biol. Rep.* **41**, 1 (2014).
70. Riou, C., Salmon, J. M., Vallier, M. J., Günata, Z. & Barre, P. Purification, characterization, and substrate specificity of a novel highly glucose-tolerant  $\beta$ -glucosidase from *Aspergillus oryzae*. *Appl. Environ. Microbiol.* **64**, 1 (1998).
71. Acebrón, I. *et al.* Catalytic cycle of the N-acetylglucosaminidase NagZ from *Pseudomonas aeruginosa*. *J. Am. Chem. Soc.* **139**, 6795–6798 (2017).

72. Nieto-Domínguez, M. *et al.* Novel pH-stable glycoside hydrolase family 3  $\beta$ -xylosidase from *Talaromyces amestolkiae*: An enzyme displaying regioselective transxylosylation. *Appl. Environ. Microbiol.* **81**, 6380–6392 (2015).
73. Hwang, E. J., Lee, Y. S. & Choi, Y. L. Cloning, purification, and characterization of the organic solvent tolerant  $\beta$ -glucosidase, OaBGL84, from *Olleya aquimaris* DAU311. *Appl. Biol. Chem.* **61**, 325–336 (2018).
74. Letunic, I. & Bork, P. 20 years of the SMART protein domain annotation resource. *Nucleic Acids Res.* **46**, D493–D496 (2018).
75. Tamura, K., Stecher, G. & Kumar, S. MEGA11: Molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* **38**, 3022–3027 (2021).
76. Kelley, L. A., Mezulis, S., Yates, C. M., Wass, M. N. & Sternberg, M. J. E. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 1 (2015).
77. Gasteiger, E. *et al.* ExPASy: The proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* **31**, 3784 (2003).
78. Eberhart, B., Cross, D. F. & Chase, L. R. Beta-glucosidase system of *neurospora crassa* I. Beta-glucosidase. *J. Bacteriol.* **87**, 761–770 (1964).
79. Miller, G. L. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* **31**, 1 (1959).
80. Saitou, N. & Nei, M. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425 (1987).

## Author contributions

All authors contributed to the study's conception and design. The experiment was designed by OJ, AS, EA, EJ, and AG, data collection and analysis were performed by OJ. The first draft of the manuscript was written by OJ. All authors read and approved the final manuscript.

## Funding

Open access funding provided by Swedish University of Agricultural Sciences. This study was financed by the Swedish International Development Cooperation Agency (SIDA) through the research and training grant awarded to Addis Ababa University and the Swedish University of Agricultural Sciences (AAU-SLU Biotech) and FORMAS (2019-00527).

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-60645-y>.

**Correspondence** and requests for materials should be addressed to O.J.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024