



OPEN

Functional characterisation of a new halotolerant seawater active glycoside hydrolase family 6 cellobiohydrolase from a salt marsh

Daniel R. Leadbeater[✉] & Neil C. Bruce[✉]

Realising a fully circular bioeconomy requires the valorisation of lignocellulosic biomass. Cellulose is the most attractive component of lignocellulose but depolymerisation is inefficient, expensive and resource intensive requiring substantial volumes of potable water. Seawater is an attractive prospective replacement, however seawater tolerant enzymes are required for the development of seawater-based biorefineries. Here, we report a halophilic cellobiohydrolase SMECel6A, identified and isolated from a salt marsh meta-exo-proteome dataset with high sequence divergence to previously characterised cellobiohydrolases. SMECel6A contains a glycoside hydrolase family 6 (GH6) domain and a carbohydrate binding module family 2 (CBM2) domain. Characterisation of recombinant SMECel6A revealed SMECel6A to be active upon crystalline and amorphous cellulose. Mono- and oligosaccharide product profiles revealed cellobiose as the major hydrolysis product confirming SMECel6A as a cellobiohydrolase. We show SMECel6A to be halophilic with optimal activity achieved in 0.5X seawater displaying $80.6 \pm 6.93\%$ activity in 1 × seawater. Structural predictions revealed similarity to a characterised halophilic cellobiohydrolase despite sharing only 57% sequence identity. Sequential thermocycling revealed SMECel6A had the ability to partially reversibly denature exclusively in seawater retaining significant activity. Our study confirms that salt marsh ecosystems harbour enzymes with attractive traits with biotechnological potential for implementation in ionic solution based bioprocessing systems.

Keywords Cellobiohydrolase, 1,4-β-glucanase, Halotolerant, Halophilic, GH6, Salt tolerant, Reversible denaturation, Thermostability, CBH, Sea water, CAZyme

Abbreviations

AA	Auxiliary activity enzyme
CAZyme	Carbohydrate-active enzymes
CBH	Cellobiohydrolase
CBM	Carbohydrate-binding module
CMC	Carboxy methyl cellulose
G1	Glucose
G2	Cellobiose
G3	Cellotriose
G4	Cellotetraose
G5	Cellopentaose
G6	Cellohexaose
GH	Glycoside hydrolase
LPMO	Lytic polysaccharide monooxygenase
PASC	Phosphoric acid swollen cellulose
PBS	Phosphate buffered saline

Attaining a fully circular bioeconomy requires the efficient and effective utilisation of all biologically derived by-products. Lignocellulosic residues offer the greatest opportunity for a renewable and sustainable feedstock to

Centre for Novel Agricultural Products, Department of Biology, University of York, Heslington, York YO10 5DD, UK.
✉email: daniel.leadbeater@york.ac.uk; neil.bruce@york.ac.uk

support the bioeconomy through the production of sustainable fuel, valuable chemicals and sustainable materials. Lignocellulose is a composite of complex polysaccharides in the form of cellulose, hemicellulose and pectins and an aromatic polymer lignin that interconnect to form a complex network of polymers in the secondary plant cell wall. The molecular and structural complexity of lignocellulose confers extreme recalcitrance requiring extreme pre-treatments for efficient enzymatic hydrolysis, however; these processes require a substantial volume of potable water potentially threatening water security local to bioprocessing developments with estimates ranging between 1.9 and 9.8 L of potable water consumed during processing only per litre of ethanol produced, at scale these volumes can have considerable local and regional impacts^{1–3}.

A prospective alternative is the use of seawater, an inexhaustible and inexpensive solvent that would liberate expensive and increasingly scarce potable water for agricultural use and human consumption. In doing so, process economics and sustainability would improve as well as mitigating water security issues. Incorporating untreated seawater as the main solvent in bioprocessing systems requires full process level compatibility. Seawater is a complex solvent with approximate pH 8.15 containing around 3.5% (w/v) salt of which 2% comprises of over 82 trace elements existing as diverse chemical species^{4–6} while the two most abundant components are sodium chloride and potassium chloride at 14% and 84% respectively.

Often seawater is an unfavourable solvent for enzymes as in the presence of ionic solutions, many enzymes aggregate, precipitate and lose activity. The mechanisms underpinning halotolerant characteristics are not fully known; however, adaptations include exhibition of increased numbers of acidic surface residues (aspartic acid and glutamic acid) and reduced numbers of lysine residues and reduced abundance of hydrophobic residues (glycine, alanine, valine) that in combination produce a greater negative electrostatic surface potential are a common feature^{7–9}. These adaptations putatively facilitate the formation of salt-bridges between the dissolved cationic salt and anionic carboxylate that configures an ordered hydration shell facilitating the dissolution and stabilisation of proteins in concentrated ionic solutions^{10–12}. The Hofmeister series describes electro-selectivity of ion and ion-protein interactions causing proteins to salt in and salt out, with ion shielding preventing protein-protein interaction increasing solubility and protein dehydration increasing protein-protein interactions decreasing solubility at low and high ion concentrations respectively. As such, initial resistance to dehydration is a major factor determining tolerance to ionic solutions. In the context of seawater Ca^{2+} , Mg^{2+} , Na^+ , K^+ , Br^- and Cl^- are highly abundant destabilisers¹³. More recently, halotolerance is of particular interest due to the emergence of efficient ionic liquid based lignocellulose pre-treatments¹⁴ with which consolidated bioprocessing using halotolerant cocktails and raw ionic liquid treated biomass has significant valorisation potential.

Cellulose is often the most abundant constituent of lignocellulose that exists as a linear polysaccharide comprised of repeating β -(1–4) *D*-glucose units with successive residues rotated by 180° around the β -(1–4) glycosidic linkages, conferring the disaccharide cellobiose as the basic repeating unit. Crystalline cellulose is formed by tightly packed cellulose chains to form micro fibrils and this crystallinity confers stability in the cell wall and recalcitrance to degradation through inaccessibility and compaction. Cellulases are glycosyl hydrolase (GH) family enzymes that synergistically act to depolymerise cellulose but differ substantially in their modes of action, interaction with cellulose and products produced. Cellulases are present in families GH1, GH3, GH5–9, GH12, GH44–45, GH48 and GH74¹⁵. Endoglucanases (EC 3.2.1.4) perform cleavage events at internal β -1,4-glucan positions within the cellulose chain producing free ends, multiplying assessable initiation sites for exo-acting cellulases. Exoglucanases (EC 3.2.1.91 and 3.2.1.176) are collectively known as cellobiohydrolases (CBH) at the reducing ends (CBH-I) or non-reducing ends (CBH-II) that act progressively to perform cleavage events at the repeating units producing cellobiose and β -glucosidases (EC 3.2.1.21) hydrolyse cellobiose into glucose. Lytic polysaccharide monoxygenases (LPMOs; EC 1.14.99.54 and EC 1.14.99.56) currently exist as auxiliary activity (AA) families AA9, AA10, AA15 and AA16 are metalloenzymes that require electron donors to orchestrate oxidative glycosidic cleavage events^{16,17}.

The seawater based biorefining concept has additional benefits greater than reducing water requirements including cost savings, improved land use efficiency, flexibility in facility locations and mitigation of environmental impacts¹⁸. Seawater also has processing advantages, passive seawater treatments on lignocellulosic rice straw and rice husk have shown a significant reduction in hemicellulose (19% and 12.4% respectively) and lignin (6.1% and 9.3% respectively) with physical changes in roughness, flexibility and crystallinity¹⁹. While more intensive hydrothermal lignocellulose pre-treatments have shown effective removal and degradation of hemicellulose putatively catalysed by the diversity of salts present^{20–22}, concordant with high sodium chloride treatments^{23–25} and seawater infused ionic liquid pre-treatments²⁶. Utilisation of these findings could alleviate industrial scale harsh physio-chemical pre-treatments and their associated adverse environmental impacts. Early investigations demonstrating the seawater based biorefining concept have ranged from fermentations targeting ethanol production from *Kluyveromyces marxianus* and artichoke in seawater mixtures²⁷ to more conventional approaches using seawater-based sugarcane molasses fermentations²⁸. The concept has been assessed in the context of more complex lignocellulosic feedstocks such as rice straw, palm residues, sugarcane bagasse, papaya and the use of macroalgae and microalgae^{13,21,29–34}. However, significant further comprehensive research is required to advance the seawater-based biorefinery concept. One area of focus is to identify an expanded repertoire of halophilic enzymes as the majority of biorefining research has focussed on non-ionic based solvents using enzymes from terrestrial systems, therefore it is likely optimal saccharification in seawater has yet to be achieved.

Despite cellulases being ubiquitous across the tree of life, natural diversity for these enzymes and particularly for their halotolerant representatives have yet to be fully explored. This is due in part to the scarce number of examples where the saline environments and lignocellulose derived polymers co-exist where selection pressures have evolved adaptations to ionic solutions in lignocellulolytic enzymes. In addition, halotolerant adaptations confer deep sequence level divergence between non-salt tolerant enzymes and salt tolerant enzymes that restricts identification. For this reason, annotation reliability for halotolerant enzymes using archive databases is often

inadequate due to low levels of representation particularly for high confidence annotations from characterisation-level proteins. The GH6 family CBH in particular is poorly represented in terms of functionally characterised proteins^{35–37} with even fewer reports of halotolerant variants^{38,39}. A correlation between thermostability and halotolerance has long been known eluded to^{40–42} and recent engineering approaches have yielded simultaneous improvements in CBH thermostability and tolerance to ionic liquids and seawater⁴². These findings strengthen the need for a greater depth of representative sequences for use as templates in engineering studies and further discovery, an approach that holds promise given the success of previous cellulase engineering efforts⁴³.

In this study, we report a novel exoglucanase with significant sequence level divergence to characterised proteins but with significant structural similarity, SMECel6A; that contains GH6 and CBM2 domains identified using a targeted meta-exo-proteomic approach. We describe the cloning, recombinant expression, purification and biochemical characterisation of SMECel6A. We demonstrate SMECel6A as a cellobiohydrolase (CBH) which is halophilic with optimal activity 0.5X seawater, consistent with the environmental origin, and tolerant to high salt conditions up to 3M NaCl. SMECel6A displays considerable plasticity in its active pH (4–7.5) and temperature range (25–50 °C). We also demonstrate considerable thermostability with reversible denaturation with retained activity of SMECel6A exclusively in the presence of seawater not observed in ion-free solution.

Materials and methods

Heterologous expression, purification and characterization

An open reading frame (a_c570829_g1_i1_1) containing glycoside hydrolase family 6 (GH6) and carbohydrate binding module family 2 (CBM2) annotation, referred to here as SMECel6A, was identified as the most abundant CAZyme within a salt marsh exo-meta-proteome dataset⁴⁴. A signal peptide was identified from the open reading frame SMECel6A using SingalP⁴⁵ which was trimmed from the coding sequence between positions 28 and 29: AQA-AS and amplified from genomic DNA harvested from Welwick salt marsh (Hull, UK) using forward GCA TCTTGTTTCGATTTCATTGA (5′–3′) and reverse TTCCGCAGGAGGACCTG (5′–3′) primers. The amplicon was cloned into pET-52b(+) expression vector with an introduced N-terminal streptavidin II affinity peptide tag (WSHPQFEK) and transformed into *Escherichia coli* Rosetta-gami[™] 2 (DE3) competent cells (Novagen). The full sequence is available under GenBank accession OR614076. Bacterial cultures were grown at 175 rpm at 37 °C and heterologous expression was induced at A_{600nm} of 0.6 with 0.1mM isopropyl β-D-1-thiogalactopyranoside (IPTG) at 16 °C for 16 h. Protein purification was performed using a StrepTrap[™] HP column (GE healthcare) with fast protein liquid chromatography system (FPLC, GE Healthcare) and protein was eluted with 2.5mM d-desthiobiotin at 1 mL min⁻¹.

Reducing sugar assay

Reducing sugars were detected using 4-hydroxybenzoic acid hydrazide (PAHBAH)⁴⁶. Sugar hydrolysate was incubated with a solution containing 50mM PAHBAH, 10mM bismuth nitrate, 10 mM sodium, potassium tartrate in 0.5M NaOH at 72 °C for 10 min. Absorbance changes were detected at 405 nm and quantified using glucose standards and against relevant boiled enzyme or no enzyme controls.

Substrate affinity and activity

A range of complex polysaccharide, oligosaccharide and respective derivatives were used to determine substrate specificity. Reactions were performed for 6 h at 250 rpm with 39.4nM SMECel6A (3μg mL⁻¹) at 43.8 °C and 50mM sodium acetate pH 5.2. The polysaccharides utilized and their final concentration (w/v) in reaction mixtures included; 0.5% Avicel[™] (m/v; Sigma), 1% (m/v) phosphoric acid swollen cellulose (PASC), 0.5% carboxy-methyl cellulose sodium salt (m/v; Sigma), 0.5% (m/v; Megazyme) beechwood xylan (1–4-β-xylose), 0.1% (m/v; Megazyme) barley glucan (1–4-β-glucan), 0.1% (m/v; Megazyme) lichenan (1,3:1,4-β-glucan), 0.1% (m/v; Megazyme) pachyman (1,3-β-glucan), 0.1% (m/v; Megazyme) arabinoxylan, 0.1% (m/v; Megazyme) arabinogalactan, 0.1% (m/v; Megazyme) xyloglucan, 0.1% (m/v; Megazyme) Lupin galactan, 0.1% (m/v; Megazyme) glucomannan, 0.1% (m/v; Megazyme) galactomannan, 0.1% (m/v; Megazyme) mannan (1–4-β-mannan). Activity was determined using the PAHBAH method⁴⁶.

Mono- and oligosaccharides (glucose, cellobiose, cellotriose, cellotetraose, cellopentaose and cellohexaose; Megazyme) were utilized at a final concentration of 0.1mM or 0.5mM with 131nM SMECel6A (10 μg mL⁻¹) in a final volume of 100 μL for 1 h. Determination of monosaccharide and oligosaccharide product profiles was achieved using ion exchange chromatography using an ICS-3000 Ion Chromatography System (Dionex, Thermo, UK) against relevant standards after⁴⁷. The 4-nitrophenyl-β-D-R (Carbosynth Limited) fluorometric derivatives (pNP-β-D-glucopyranoside, pNP-β-D-cellobioside, pNP-β-D-cellotrioside, pNP-β-D-cellotetraoside, pNP-β-D-cellopentaoside and pNP-β-D-cellohexaoside) were utilized in a final concentration of 0.5mM as above. Reactions were terminated with the addition of 1M sodium carbonate (Na₂CO₃) to a final concentration of 0.5M. Activity was assessed against 4-nitrophenol standards and measured at 405nm.

Bivariate pH, temperature optima

Two-dimensional pH and temperature experiments were conducted in triplicate with relative activity profiles obtained using a 100 mM citrate-phosphate (CPB) buffer system in a pH range of 4–8⁴⁸. The reaction mixture contained 13 1nM SMECel6A (10 μg mL⁻¹) and 2% PASC for 24 h in a total reaction volume of 100 μL. Reducing sugar release was quantified against a boiled enzyme negative control with the PAHBAH method.

Halotolerance

Salt tolerance was determined using artificial seawater (ASW; with typical ionic concentrations at 1 × concentrate: Chloride 19,336 (mg/L), Sodium 10,752 (mg/L), Sulfate 2657 (mg/L), Magnesium 1317 (mg/L), Potassium 421

(mg/L), Calcium 380 (mg/L), Carbonate/Bicarbonate 142 (mg/L), Strontium 9.5 (mg/L), Boron 16 (mg/L), Bromide 56 (mg/L), Iodide 0.060 (mg/L), Lithium 0.3 (mg/L), Silicon <0.1 (mg/L), Iron 0.0098 (mg/L), Copper 0.0003 (mg/L), Nickel <0.015 (mg/L), Zinc 0.0107 (mg/L), Manganese 0.0023 (mg/L), Molybdenum 0.0098 (mg/L), Cobalt 0.0004 (mg/L), Vanadium <0.015 (mg/L), Selenium <0.019 (mg/L), Rubidium 0.118 (mg/L), Barium 0.04 (mg/L) (Marine salt, SeaChem, UK) and sodium chloride concentrations equivalent to the ionic strength of estuarine waters, seawater and above. Enzyme preparations were dialyzed three times in 5 000× volume of > 18 MΩ·cm dH₂O for 2, 16 and 2 h respectively to remove ionic background. Reactions were conducted at the respective pH and temperature optima; 43.8 °C and 50mM sodium acetate pH 5.2 in 0.5% Avicel™ with 131nM SMECel6A (10 μg mL⁻¹) for 6 h at 250 rpm.

Thermostability

Thermal shift assays were performed using SYPRO™ orange protein dye (Sigma S5692) at a final dye concentration of 1–5× and assayed on a Stratagene Mx300 rtPCR for 70 cycles between 25 °C and 95 °C at increments of 2 °C min⁻¹. Fluorescence data was normalised to elucidate the protein melting point (T_m). Thermostability assays were conducted by heat treatment of SMECel6A at 100 °C in various ionic solutions followed by 1h of renaturation at 25 °C. Sequential thermostability and reversible renaturation of SMECel6A was explored in a Prometheus nanoDSF (Nanotemper Technologies, London, UK) at 13.1μM (1mg/mL) in NT.48 standard capillaries during ten thermocycles between 25 °C for 25 min and 90 °C for 5 min at a ramp rate of 1 °C min⁻¹ at pH 5.2 in sodium chloride, artificial seawater and 100 mM citrate phosphate buffer. Denaturation and renaturation were assessed as an index of intrinsic tryptophan fluorescence at 350nm/330nm. Precipitation was assessed as a function of light scattering (mAU). Confirmation of function post heat treatment was validated using the PAHBAH method using 0.5% Avicel™ with 126nM SMECel6A (10 μg mL⁻¹).

Processivity

Avicel (0.2%), phosphoric acid swollen cellulose (0.2%) and carboxy-methyl cellulose (0.05%) at 13.12nM SMECel6A (1 μg mL⁻¹) for 1 h at optimum conditions. To sample, the reaction was centrifuged at 4600 rpm for 1 min and 4 volumes of 100% 4 °C ethanol added to precipitate proteins and impurities. The hydrolysate was dehydrated in a refrigerated vapor trap (Thermo). Samples were re-suspended in the original volume taken and analysed with ion exchange as above. Apparent processivity (P_{app}) was determined semiquantitatively as the molar ratio of cellobiose/(glucose + cellotriase)^{49,50}.

Structural analysis and predictions

The structure of SMECel6A was predicted using AlphaFold version 2.1.1⁵¹. Five models were generated with a mean pLDDT of 91.72 ± 0.75; the model selected for analysis exhibited a pLDDT of 92.54. Structural alignments and electrostatic potentials were generated using ChimeraX⁵².

Bioinformatics and statistics

Analysis of Variance (ANOVA) and Kruskal–Wallis H tests were performed using SciPy⁵³ and Scikit⁵⁴. All data was assessed for normality using the Shapiro–Wilk test.

Results

Protein production and properties

A multi-domain protein originating from open reading frame a_c570829_g1_i1_1⁴⁴, designated here as SMECel6A, was identified as the most abundant CAzyme within a salt marsh meta-exo-proteome dataset⁴⁴. SMECel6A has maximum identity of 50% to a characterised *Cellulomonas fimi* 1,4-β-cellobiohydrolase (Swiss-Prot accession P50401.1). SMECel6A is multi domain protein containing a carbohydrate-binding module (CBM) family 2 domain (IPR012291) and glycoside hydrolase (GH) family 6 domain (IPR016288) separated by two sequential predicted PKD domains (IPR035986) with three predicted disulphide bridges ostensibly acting as a linker between the CBM2 and GH6 domains. Recombinant expression was performed using a pET-52b(+) expression vector in *Escherichia coli* Rosetta-gami™ 2 (DE3) cells and purified to homogeneity using an introduced N-terminal streptavidin affinity tag. The recombinant protein exhibits a predicted molecular weight of 76.1kDa and theoretical isoelectric point (pI) of 4.32.

Substrate specificity, optima, and product profiles of SMECel6A

The hydrolytic activity of SMECel6A was tested against 4-nitrophenyl-β-D-R glucooligosaccharide derivatives (pNP-β-D-glucopyranoside, pNP-β-D-cellobioside, pNP-β-D-cellotrioid, pNP-β-D-cellobetraoside and pNP-β-D-cellopentaoside) but no detectable activity was observed after incubation for 24h. However; screening of polysaccharides revealed substantial sugar release and a high affinity for Avicel™ that was significantly higher than both PASC (ANOVA, $F_{1,8} = 36.7$, $p < 0.00$) and CMC (ANOVA, $F_{1,8} = 479.35$, $p < 1.99 \times 10^{-8}$) with relative activities of 62.25 ± 10.7% and 28.75 ± 5.57% respectively. Insignificant activity (Kruskal–Wallis test: H = 1, DF = 1, $p = 0.317$) was also detected for lichenan (1,3:1,4-β-D-Glucan) and glucomannan (Fig. 1a).

To assess the optimal pH and temperature conditions for SMECel6A, a bivariate experiment was undertaken which elucidated the optima as 43.8°C and pH 5.2 (Fig. 1d). Interestingly, both the pH and temperature had a broad range of high functioning in which the enzyme remained highly active between pH 4 and 7.5 and retained 60% activity between 50 °C and 25 °C. A thermal shift assay was performed to confirm the structural integrity of GH6 and CBM2 domains, revealing two unfolding events (Fig. 1b) and these observations are also confirmed by the bivariate pH and temperature optima assessment profile which displayed two activity apices in agreement

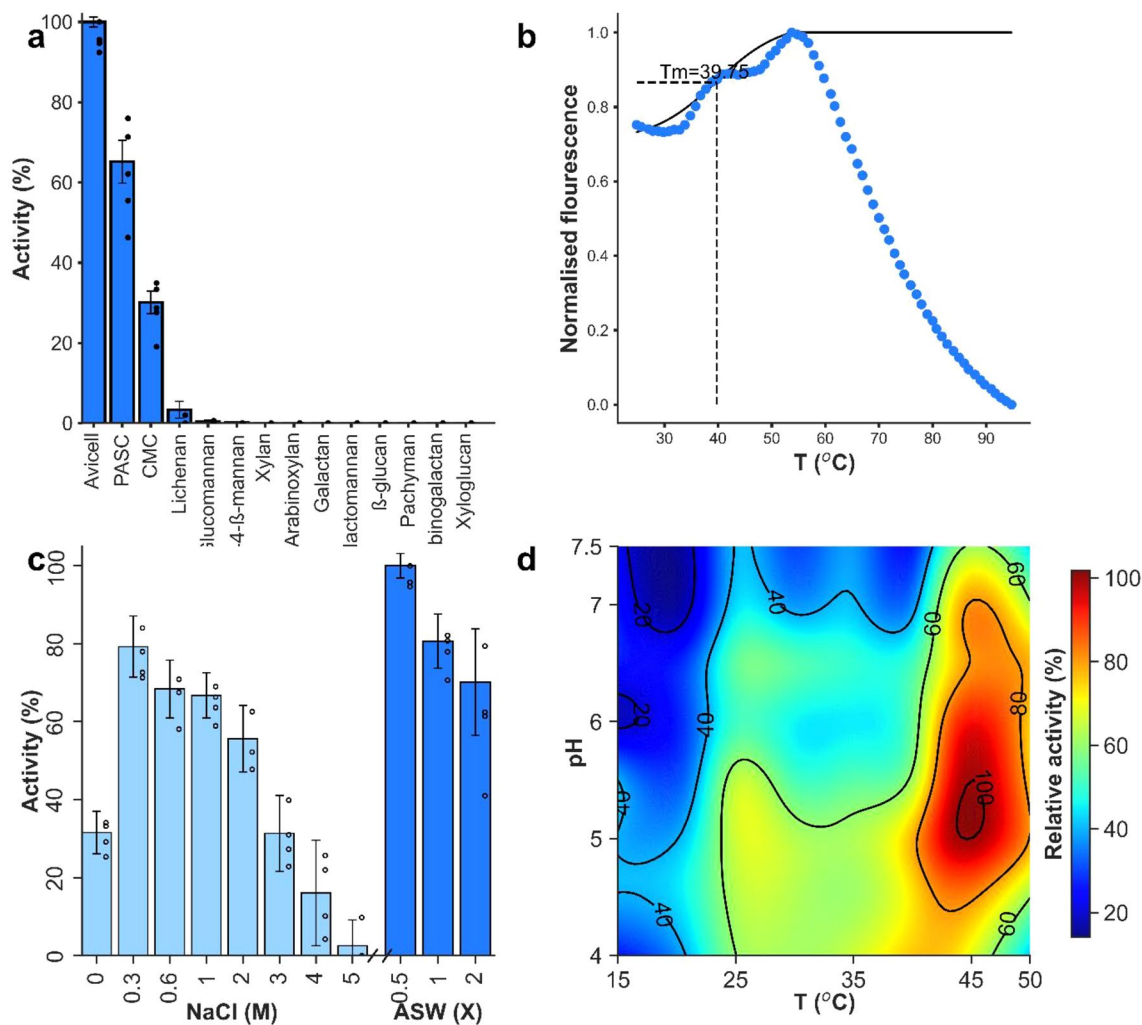


Figure 1. Characteristics of SMECel6A. **(a)** Substrate affinities of SMECel6A. Error bars represent standard error of five technical replicates. Asterisks indicate significance at a confidence interval of $p < 0.05$ assessed by a one-way ANOVA. **(b)** A representative thermal shift profile of SMECel6A. A sigmoidal curve between the maximum and minimum before the maximum is shown, the median value of this fit was determined as the melting point (T_m ; 39.75). **(c)** Halotolerance of SMECel6A as a function of activity in solutions of increasing salt concentrations. **(d)** Bivariate pH/temperature optima of SMECel6A (optima; 43.8 °C, pH 5).

with the double peak observed within the thermal shift profile suggesting two independent unfolding events with independent effects on catalytic efficiency (Fig. 1b). The melting point (T_m) was calculated to be 39.75 °C however this is likely biased lower due to the double peak.

To determine the mode of action for SMECel6A, product profiles were elucidated using ion exchange chromatography after incubation with monosaccharides and oligosaccharides (G1–G6). It was apparent SMECel6A was unable to hydrolyse G2 and showed very weak activity toward G3 with negligible quantities of G1 and G2 detected during incubation (Fig. 2a). The major product from incubation with G4 was cellobiose, incubation with G5 evolved only G2 and G3, and incubation with G6 revealed only G2, G3 and G4. Further product profiles were determined for polysaccharides where activity was highest (Avicell™, PASC and CMC; Fig. 2b). In all cases, synonymous with results observed with oligosaccharide substrate, the major product G2, with G1 and G3 evolved as minor products in all reactions. The product profiles for each polysaccharide was quantified and relative proportions were used to determine enzyme apparent processivity (P_{app}) on each substrate which were determined to be 2.65 ± 0.06 , 11.77 ± 1.76 and 3.57 ± 0.80 for Avicell™, PASC and CMC respectively (Fig. 2c).

Salt tolerance, thermal stability and reversible renaturation

Using the optima determined above, SMECel6A tolerance to ionic conditions was explored using NaCl and artificial seawater (ASW). Activity was significantly higher in ionic reaction media compared to dialyzed H_2O conditions where $31.6 \pm 5.45\%$ relative activity was observed (Fig. 1c). Activity was highest in 0.5X seawater and significantly higher than the highest activity observed in NaCl (0.3M) where $79.2 \pm 7.8\%$ activity was retained. Activity remained high in high ionic strength solutions up to 2M NaCl ($55.6 \pm 8.45\%$) before declining rapidly in

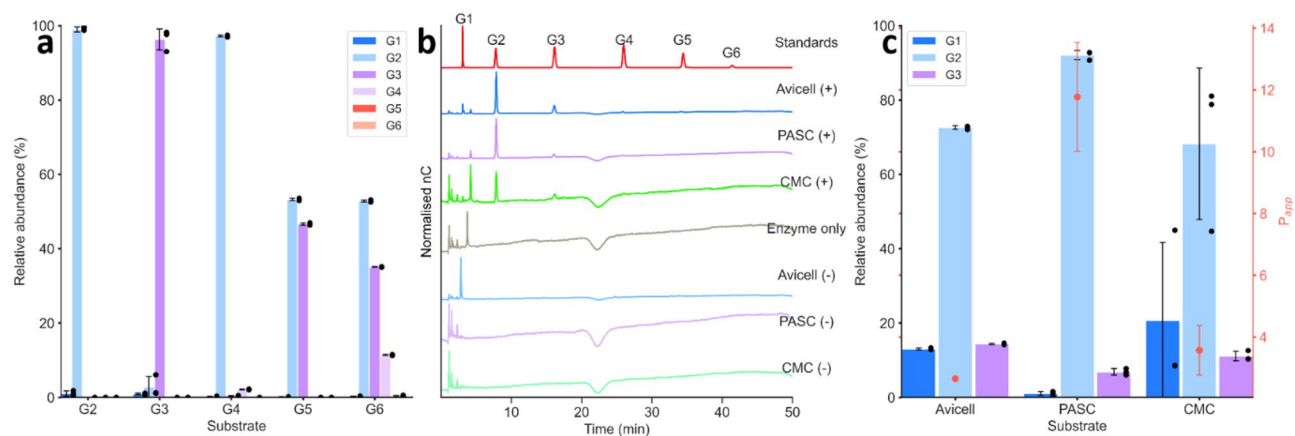


Figure 2. Product profiles for SMECel6A post reaction with oligosaccharides substrates. **(a)** Product profile SMECel6A after incubation with oligosaccharides. **(b)** Ion exchange chromatography product profiles SMECel6A after incubation with polysaccharides. **(c)** Relative abundance of products and apparent processivity (P_{app}) after incubation with Phosphoric acid swollen (PASC), Carboxy-methyl cellulose (CMC) and Avicell™. G1; glucose, G2; cellobiose, G3; celotriose, G4; cellotetraose, G5; cellopentaose, G6; cellohexaose.

3M ($31.34 \pm 9.76\%$) and 4M ($16.05 \pm 13.54\%$) NaCl solutions, however higher artificial seawater concentrations retained higher levels of activity at and for 1X ($80.86 \pm 6.93\%$) and 2X ($70.17 \pm 13.62\%$) concentrated seawater respectively.

The observation that salt is required for optimal performance was explored further during thermostability assessments. We observed a resistance to inactivation post heat treatment of SMECel6A in complex ionic solutions. Consecutive thermal cycling between 5 and 90 °C revealed a profile suggestive of reversible denaturation. The thermal shift profile was notably different between ultrapure H₂O and ionic solutions (Fig. 3a,b). Irreversible protein aggregation as measured as light scattering (maU) during thermal cycling was observed in sodium chloride solutions; however, in seawater the protein appeared to largely return to solution upon cooling. No aggregation was observed in the ultrapure H₂O condition, however, activity assays performed with the heat-treated enzyme revealed almost complete inactivation for H₂O and NaCl conditions, while the enzyme retained $71.2 \pm 3.67\%$ 0.5X concentrated seawater (Fig. 3c).

Despite major sequence level divergence, there were significant structural similarities between HmCel6A and the predicted structure of SMECel6A (Fig. 4a) with calculated root mean square deviation (RMSD) value of 0.655Å for the CBH domains. Analysis of surface residue properties reveals a marked predominance of negatively charged residues as expected from highly ion-tolerant protein (Fig. 4b). Surface charge analysis suggests that

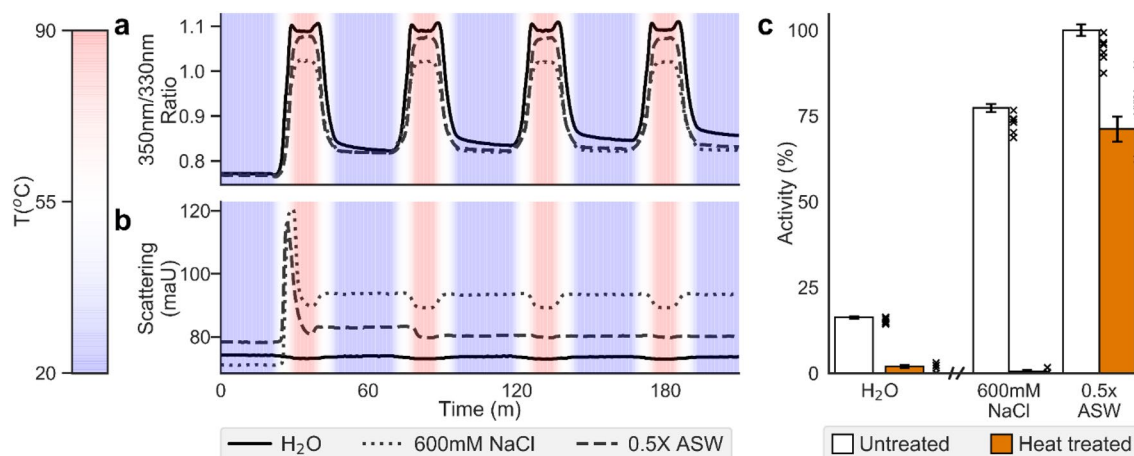


Figure 3. Thermal reversible denaturation profile of SMECel6A. **(a)** Thermal shift profile as a function of fluorescence at 330nm and 350nm. The enzyme was subjected to four sequential 20 °C to 90 °C treatment cycles. The fluorescence ratio increases as the protein denatures. The results show that the thermal shift profile was notably different between H₂O and ionic solutions. **(b)** Light scattering as a function of enzyme aggregation for enzyme SMECel6A when subjected to four sequential 20 °C to 90 °C treatment cycles. Scattering increases as proteins precipitate. Partial and almost complete re-solubilisation is visible for 600mM NaCl and 0.5X ASW solutions respectively. **(c)** Relative activity of untreated and heat treated SMECel6A, means are displayed as re-normalised percentages, raw data is plotted individually. ASW: artificial seawater. Error bars represent SE (n=6).

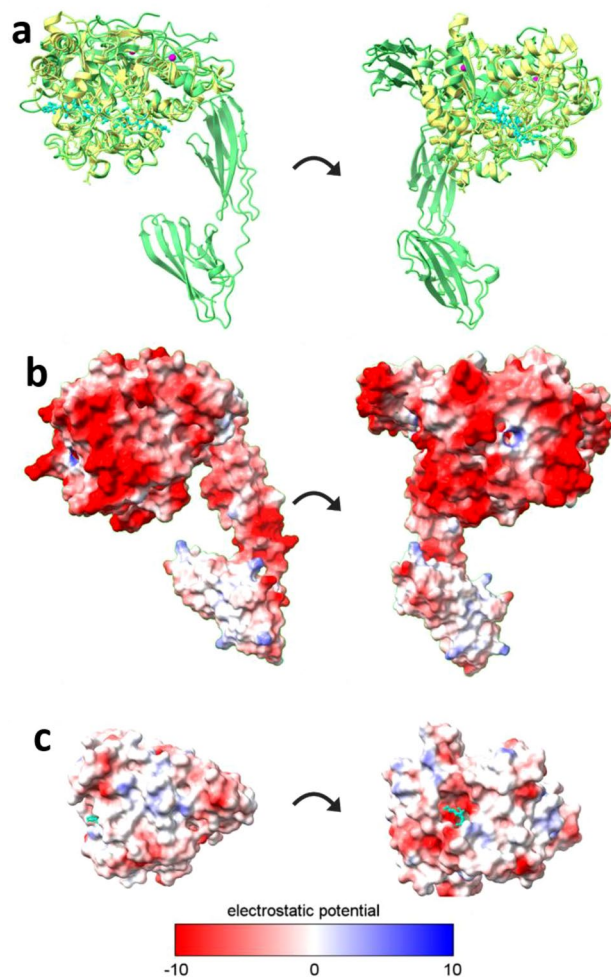


Figure 4. Predicted structure and structural comparisons of SMECel6A. (a) Structural alignment between SMECel6A (cellobiohydrolase domain and carbohydrate binding module family 2; green) and HmCel6A (10.2210/pdb6K54/pdb; yellow) in complex with trisaccharide (cyan). Calcium ions are displayed in pink. (b) Electrostatic potential of SMECel6A shown as a colored gradient between red (acidic) to blue (basic). (c) Electrostatic potential of HmCel6A shown as a colored gradient between red (acidic) to blue (basic).

SMECel6A exhibits a highly negatively charged surface greater than that of its most similar counterpart HmCel6A (Fig. 4b,c). Both of which exhibited a disproportionate abundance of negative and neutral charges coupled with a distinctive lack of positively charged residues.

Discussion

The GH6 family comprises exo-1,4- β -glucanases (cellobiohydrolases; CBH) and endo-1,4- β -glucanases (endoglucanase; EG). These activities, together with β -glucosidases and lytic polysaccharide monooxygenases (LPMOs) act to synergistically depolymerise cellulose into smaller oligosaccharides. Here, we explored SMECel6A, a CBM2 containing GH6 family enzyme identified from an environmental meta-omics analysis of a lignocellulosic decomposition performed in situ in a salt marsh ecosystem⁴⁴. Our investigations of SMECel6A revealed product profiles in which G2 was the major product upon hydrolysis of oligosaccharide and polysaccharide substrates suggesting a processive mechanism, while G1 and G3 were identified at trace levels, putatively originating from initial or false start events⁵⁵.

Semiquantitative delineation of apparent processivity (P_{app}), the efficiency and ability with which an enzyme can perform successive successful hydrolytic events prior to dissociation or obstruction⁵⁶; processive catalysis is thought to be advantageous over distributive catalysis for biopolymeric substrates with processive catalysis having been reported in nucleases, proteases and motor proteins as well as exo- and endo- acting glycoside hydrolases, despite this, values are not widely reported⁵⁷. The P_{app} for SMECel6A revealed a significantly higher processivity for the hydrolysis of amorphous cellulose (PASC; 11.77 ± 1.76) than crystalline (Avicell™; 2.65 ± 0.06) or soluble cellulose (CMC; 3.57 ± 0.80). The observed disparity likely arises from the greater accessibility of amorphous cellulose microfibrils as processive enzymes complex with the chain terminus⁵⁸ reducing the number of false start events. These results fall within the spectrum of previously determined values for CBH belonging to broad

families of 2.7–20.8^{49,59–63}. The P_{app} of SMECel6A for crystalline cellulose are lower than previously reported bacterial CBH, such as wild type TfCel6B (*Thermobifida fusca*) with P_{app} values of 7.2–11.8 on crystalline cellulose (filter paper)^{59,62}, as well as LqCel7b (*Limnoria quadripunctata*) and Cel7D (*Phanerochaete chrysosporium*) where a P_{app} values greater than 7 were recorded on PASC and Avicel^{50,64}.

The affinity of SMECel6A for crystalline cellulose may be due to the CBM2 domain which is known to bind crystalline cellulose and are thought to enhance catalytic activity by retaining proximity and interface with the substrate which could explain the dual peak in activity profiles during the assessment of pH and temperature optima⁶⁵. This feature may be particularly important for this enzyme originating from a highly turbid salt marsh environment to prevent the loss of enzyme from the originating location to retain the benefit to the secreting organism.

These results are concordant with previous observations, where GH6 family cellobiohydrolase catalytic domains are known to be highly processive on crystalline cellulose such as CfCel6B⁶⁶. However, the relationship between processivity and efficiency with regard to turnover and rates of reaction are complex and not yet fully understood⁶⁷, evidenced here by the significantly higher rates of activity on crystalline cellulose where processivity was observed to be lowest (Fig. 2c). Our results suggest SMECel6A is a processive CBH and CBHs within the GH6 family putatively act from the non-reducing ends of cellulose (EC 3.2.1.91)⁶⁸. Similar substrate affinities reported here have been observed for other GH6 family cellobiohydrolase, such as PcCel6A from *Paenibacillus curdlanolyticus* B-6³⁶ but also contrasted by CBM3 containing cellobiohydrolase Cel6D from *Paenibacillus barcinonensis* that displayed almost exclusive affinity for amorphous cellulose³⁷.

Cellobiohydrolases occupy 40–60% of lignocellulolytic cocktails⁶⁹ which signifies their importance, despite this there are few known salt tolerant cellobiohydrolase representatives^{50,70–72}, ostensibly due to a lack of assessment and limited interaction of cellulose in extreme saline environments restricting the pool of candidates for investigation as few saline environments, with the exceptions of salt marshes, seagrass meadows and mangroves, contain reliable supplies of pristine lignocellulose. Salt marshes vary greatly in their form, biotope and salinity and exist in sheltered intertidal regions transitioning the terrestrial to marine environment that are characterised by halotolerant fauna and flora with high primary productivity. Their periodic inundation and disturbance with seawater mediates an array of physicochemical stresses that vary in extremity on diurnal, seasonal, and inter-annual timescales in relation to other factors (air exposure, precipitation levels, seawater nutrient loading, tidal cycle), which underpin huge surface level fluctuations in pH (below 7.2–8.3 and above), salinity (2–3% and above), temperature (sub-zero to 25 °C and greater), pCO₂ and O₂ (1–14 mg L⁻¹) with heterogeneous pockets of greater extremities⁷³. It is therefore probable that enzymes from these ecosystems have tolerances reflective of their native environment that confers these ecosystems as valuable resources for bioprospecting. Instead, structure-based engineering approaches have been employed to enhance catalytic activity, thermal and pH stability and ionic tolerance with various degrees of success^{42,74}. Similar plasticity in environmental tolerance been observed for an endo-acting halophilic GH6 isolated from *Thermobifida haloterans* YIM 90,462(T)⁷⁵. Our observations of SMECel6A, where optimal activity can only be achieved in ionic solutions designate the enzyme as halophilic. Halophilic enzymes commonly feature adaptations that reduce hydrophobicity, increased proportions of acidic surface residues particularly aspartic acid, reduced cysteine residues and tendencies to form coiled structures over helices^{76,77}. Often, these adaptations confer polyextremophilic properties where resistance to one environmental extreme also confers resistance to others, mediated by the presence of salt ions⁷⁸. In support of this, for SMECel6A these adaptations also appear to have consigned great plasticity in pH and temperature tolerances within which high activity is retained across a broad range.

Exo-proteomic approaches are advantageous as they are culture independent and enable the identification of proteins in environmental conditions in situ. Indeed, similar approaches have yielded successful results, the crystal structure of a novel hyper-thermostable cellobiohydrolase, HmCel6A (<https://doi.org/10.2210/pdb6K54/pdb>) was the highest identity match to SMECel6A GH6 domain at 57% was identified by employing a metagenomics mining approach from extreme environments³⁹. This exemplifies the degree to which sequence divergence exists for enzymes among and between extreme environments and strengthens the case for further exploration. Despite these difference, structural homology between HmCel6A and SMECel6A was strikingly high suggesting a high degree of compositional heterogeneity within the natural diversity within and among halophilic proteins.

Taxonomic delineation of SMECel6A, originating from a meta-exo-proteomic dataset is challenging as the parent organism cannot be precisely determined, however, SMECel6A has significant nucleotide identity (> 80%; CP033077.1) to *Vibrio spp.*, also confirmed within the original dataset⁴⁴. The protein also contains a CBM2 domain, prevalence of which outside of bacteria is rare, supports this hypothesis⁷⁹.

The assessment of biocatalysts and in particular cellulases in seawater as a hydrolytic solvent to date is limited. Seawater concentration has been shown to be generally inhibitive to cellulase activity, with the enzyme-dependent inhibitory effects. Wild type CelA2 lost > 50% activity in seawater, where directed evolution generated a mutant with a 1.6-fold relative improvement in activity⁸⁰. Industrial cellulase cocktails, including Accellerase-1500 have been shown to be effective on amorphous and crystalline cellulose but exhibited minor inhibitory effects⁸¹. In accordance with this, the production of seawater optimised cocktails has been explored using halophilic organisms in seawater-based production media such as *Fusarium subglutinans* MTCC11891 that produced thermostable, pH stable, and halophilic cellulases with optimal activity observed 0–8-1M NaCl⁸². Halophilic organisms such as *Aspergillus spp.* have also shown promising halophilic enzyme products such as cellulases from *Aspergillus flavus* tolerant up to 200 g L⁻¹ NaCl⁸³, endo-1,4-β-xylanase from *Aspergillus clavatus* tolerant to 3M NaCl⁸⁴ and cellulases and xylanases from halophiles *Aspergillus flavus* and *Aspergillus penicilloides* co-culture tolerant up to 30% NaCl⁸⁵.

Exposing SMECel6A to sequential denaturing thermal treatments followed by a period of rest to facilitate renaturation revealed reversible denaturation ability for CMECel6a exclusively in seawater conditions. While some precipitation and possible conformational changes were observed, the degree of retained activity after

the first thermocycle suggested the observed conformational changes were only partially deleterious and likely associated with the CBM2 domain that was poorly thermostable in pH and temperature optimisation experiments (Fig. 1d). It is possible the reduced rate of activity is a result of the loss of the CBM2 domain, observed in the pH and temperature optima profile to boost activity, reducing the overall rate of activity and that the catalytic domain was largely unaffected. Additionally, refolding of halophilic proteins can take several days and so the 1 h period employed in this study may have been inadequate to facilitate complete refolding⁸⁶. This confirms that ions in seawater play a protective role, supported by our observation that tertiary structure was differentially affected during denaturation in H₂O, NaCl and ASW conditions (Fig. 3a), putatively from dehydration for the catalytic domain, from dehydration, denaturation and aggregation⁴¹. The lack of aggregation in ultrapure H₂O is an interesting observation, it is also possible that exposed hydrophilic regions in the denatured form are sufficient to maintain solubilisation, or that exposed hydrophobic regions become solvated, maintaining solubility. The lack of activity, taken together with the lack of possible ionic interactions that usually promote aggregation and precipitation of proteins, and the large degree of conformational change observed during thermocycling compared to the ionic solutions (Fig. 3a) suggest the formation of a destabilised intermediary complex that changes conformation during thermocycling while retaining solubility in ultrapure H₂O.

An adaptation of halotolerant enzymes to combat initial dehydration is to display a greater resistance to initial dehydration, likely a result of a greater negative surface charge characterised by acidic surface residues. This feature has been suggested to promote function and maintains solvation in lower water activity environments however the mechanism is not yet fully understood⁸⁷. The negative surface charge improves the ability of halophilic proteins to bind hydrated cations, where they would otherwise disrupt the ordered solvation shell, that help maintain a surface hydration layer, reducing hydrophobicity and preventing denaturation and aggregation⁸⁸. Recent rational design experiments of surface charge engineering of a β -glucosidase have demonstrated these features to have significant effects on stability and catalytic capacity highlighting both the complexity and importance of this characteristic⁸⁹. These features are apparent in SMECel6A which displayed a higher negative surface charge with a greater number of negative and neutral surface residues. This is known to be a shared feature among extracellular enzymes identified from a diversity of taxa including archaea, bacteria and fungi, spanning broad functionalities including lipases, proteases and lignocellulose degrading enzymes with multiple industrial uses^{90–92}. The SMECel6A negative surface charges shares similarities to other halophilic cellulases including *Hu*-CBH1 from *Halorhabdus utahensis* that was stable in high concentrations of the ionic solvent 1-allyl-3-methylimidazolium chloride⁷⁰, a halophilic CBH-I from *Penicillium verrucosum* that displayed significant activity in up to 2 times concentrated seawater with minimal inhibition⁴², nmGH45; a halophilic endo- β -1,4-glucanase with negative surface charges demonstrated 88% relative activity in 4M NaCl with similar activity levels in 10% 1-butyl-3-methylimidazolium chloride⁹³ and a halophilic and β -glucosidase B9L147 from *Thermomicrobium roseum* that additionally demonstrated significant activity in seawater⁹⁴. Identifying a broader scope of intrinsically halophilic enzymes will also aid future rational engineering efforts.

Here, we determined the mode of action and halotolerance of a divergent bacterial cellobiohydrolase SMECel6A identified using meta-exo-proteomics in a salt marsh ecosystem. SMECel6A has significant industrial potential in the development of consolidated bioprocessing using ionic liquid pre-treatments or seawater based biorefining practices as well as embellishing the available templates from naturally occurring sequence diversity to aid future intelligent design protein engineering studies. The novel sequence, domain architecture and characteristics of SMECel6A will strengthen archive databases to enhance future identification of further halophilic cellobiohydrolases with attractive traits, in particular those with greater negative surface residue charged, which are currently underrepresented using homology and design-based approaches.

Data availability

Metaproteomic and metatranscriptomic databases mined during this research are available at MassIVE (<https://massive.ucsd.edu/>) and ProteomeXchange (<http://www.proteomexchange.org/>) accessions MSV000083872 and PXD014068 respectively. Sequence data for SMECel6A is available under GenBank accession OR614076.

Received: 18 October 2023; Accepted: 6 February 2024

Published online: 08 February 2024

References

1. Alsaleh, M., Abdul-Rahim, A. S. & Abdulwakil, M. M. EU28 region's water security and the effect of bioenergy industry sustainability. *Environ. Sci. Pollut. Res.* **28**, 9346–9361. <https://doi.org/10.1007/s11356-020-11425-4> (2021).
2. Scapini, T. *et al.* Seawater-based biorefineries: A strategy to reduce the water footprint in the conversion of lignocellulosic biomass. *Bioresour. Technol.* <https://doi.org/10.1016/j.biortech.2021.126325> (2022).
3. Wu, M., Mintz, M., Wang, M. & Arora, S. Water consumption in the production of ethanol and petroleum gasoline. *Environ. Manage.* **44**, 981–997. <https://doi.org/10.1007/s00267-009-9370-0> (2009).
4. Elderfield, H., Holland, H. D. & Turekian, K. K. *The Oceans and Marine Geochemistry* 1st edn. (Elsevier, 2006).
5. Turner, D. R., Whitfield, M. & Dickson, A. G. The equilibrium speciation of dissolved components in fresh-water and seawater at 25-degrees-C and 1 atm pressure. *Geochim Cosmochim Acta* **45**, 855–881. [https://doi.org/10.1016/0016-7037\(81\)90115-0](https://doi.org/10.1016/0016-7037(81)90115-0) (1981).
6. Byrne, R. H., Kump, L. R. & Cantrell, K. J. The influence of temperature and pH on trace-metal speciation in seawater. *Mar. Chem.* **25**, 163–181. [https://doi.org/10.1016/0304-4203\(88\)90062-x](https://doi.org/10.1016/0304-4203(88)90062-x) (1988).
7. Enache, M. & Kamekura, M. Hydrolytic enzymes of halophilic microorganisms and their economic values. *Rom. J. Biochem.* **47**, 47–59 (2010).
8. Delgado-Garcia, M., Valdivia-Urdiales, B., Aguilar-Gonzalez, C. N., Contreras-Esquivel, J. C. & Rodriguez-Herrera, R. Halophilic hydrolases as a new tool for the biotechnological industries. *J. Sci. Food Agric.* **92**, 2575–2580. <https://doi.org/10.1002/jsfa.5860> (2012).
9. Liu, X. S., Huang, Z. Q., Zhang, X. N., Shao, Z. Z. & Liu, Z. D. Cloning, expression and characterization of a novel cold-active and halophilic xylanase from *Zunongwangia profunda*. *Extremophiles* **18**, 441–450. <https://doi.org/10.1007/s00792-014-0629-x> (2014).

10. Madern, D., Ebel, C. & Zaccai, G. Halophilic adaptation of enzymes. *Extremophiles* **4**, 91–98. <https://doi.org/10.1007/s007920050142> (2000).
11. Nayek, A., Sen Gupta, P. S., Banerjee, S., Mondal, B. & Bandyopadhyay, A. K. Salt-bridge energetics in halophilic proteins. *Plos One* **9**, 11. <https://doi.org/10.1371/journal.pone.0093862> (2014).
12. Warden, A. C. *et al.* Rational engineering of a mesohalophilic carbonic anhydrase to an extreme halotolerant biocatalyst. *Nat. Commun.* **6**, 10. <https://doi.org/10.1038/ncomms10278> (2015).
13. Fang, C. J. *et al.* Seawater as alternative to freshwater in pretreatment of date palm residues for bioethanol production in coastal and/or arid areas. *ChemSusChem* **8**, 3823–3831. <https://doi.org/10.1002/cssc.201501116> (2015).
14. Pramanik, S., Dhoke, G. V., Jaeger, K. E., Schwaneberg, U. & Davari, M. D. How to engineer ionic liquids resistant enzymes: Insights from combined molecular dynamics and directed evolution study. *ACS Sustain. Chem. Eng.* **7**, 11293–11302. <https://doi.org/10.1021/acssuschemeng.9b00752> (2019).
15. Talamantes, D., Biabini, N., Dang, H., Abdoun, K. & Berlemont, R. Natural diversity of cellulases, xylanases, and chitinases in bacteria. *Biotechnol. Biofuels* **9**, 11. <https://doi.org/10.1186/s13068-016-0538-6> (2016).
16. Sabbadin, F. *et al.* An ancient family of lytic polysaccharide monoxygenases with roles in arthropod development and biomass digestion. *Nat. Commun.* **9**, 12. <https://doi.org/10.1038/s41467-018-03142-x> (2018).
17. Filiatrault-Chastel, C. *et al.* AA16, a new lytic polysaccharide monoxygenase family identified in fungal secretomes. *Biotechnol. Biofuels* <https://doi.org/10.1186/s13068-019-1394-y> (2019).
18. de Maria, P. D. On the use of seawater as reaction media for large-scale applications in biorefineries. *Chemcatchem* **5**, 1643–1648. <https://doi.org/10.1002/cctc.201200877> (2013).
19. Jiang, L. P., Du, P. & Wang, H. Seawater modification of lignocellulosic fibers: comparison of rice husk and rice straw fibers. *Mater. Res. Express* <https://doi.org/10.1088/2053-1591/abe8c4> (2021).
20. Fang, C. J. *et al.* Factors affecting seawater-based pretreatment of lignocellulosic date palm residues. *Bioresour. Technol.* **245**, 540–548. <https://doi.org/10.1016/j.biortech.2017.08.184> (2017).
21. Zhang, X. K., Zhang, W. W., Lei, F. H., Yang, S. J. & Jiang, J. X. Coproduction of xylooligosaccharides and fermentable sugars from sugarcane bagasse by seawater hydrothermal pretreatment. *Bioresour. Technol.* <https://doi.org/10.1016/j.biortech.2020.123385> (2020).
22. Wu, Y. *et al.* Promising seawater hydrothermal combining electro-assisted pretreatment for corn stover valorization within a biorefinery concept. *Bioresour. Technol.* <https://doi.org/10.1016/j.biortech.2022.127066> (2022).
23. Jiang, Z. C., Yi, J., Li, J. M., He, T. & Hu, C. W. Promoting effect of sodium chloride on the solubilization and depolymerization of cellulose from raw biomass materials in water. *ChemSusChem* **8**, 1901–1907. <https://doi.org/10.1002/cssc.201500158> (2015).
24. Jiang, Z. C. *et al.* Sodium chloride-assisted depolymerization of xylo-oligomers to xylose. *ACS Sustain. Chem. Eng.* **6**, 4098. <https://doi.org/10.1021/acssuschemeng.7b04463> (2018).
25. vom Stein, T. *et al.* Salt-assisted organic-acid-catalyzed depolymerization of cellulose. *Green Chem.* **12**, 1844–1849. <https://doi.org/10.1039/c0gc00262c> (2010).
26. Das, L. *et al.* Seawater-based one-pot ionic liquid pretreatment of sorghum for jet fuel production. *Bioresour. Technol. Rep.* **13**, 100622. <https://doi.org/10.1016/j.biteb.2020.100622> (2021).
27. Yuan, W. J., Zhao, X. Q., Ge, X. M. & Bai, F. W. Ethanol fermentation with *Kluyveromyces marxianus* from Jerusalem artichoke grown in salina and irrigated with a mixture of seawater and freshwater. *J. Appl. Microbiol.* **105**, 2076–2083. <https://doi.org/10.1111/j.1365-2672.2008.03903.x> (2008).
28. Zaky, A. S., Greetham, D., Tucker, G. A. & Du, C. Y. The establishment of a marine focused biorefinery for bioethanol production using seawater and a novel marine yeast strain. *Sci. Rep.* <https://doi.org/10.1038/s41598-018-30660-x> (2018).
29. Dev, B., Bakshi, A. & Paramasivan, B. Prospects of utilizing seawater as a reaction medium for pretreatment and saccharification of rice straw. *Chemosphere* <https://doi.org/10.1016/j.chemosphere.2022.133528> (2022).
30. Indira, D. & Jayabalan, R. Saccharification of lignocellulosic biomass using seawater and halotolerant cellulase with potential application in second-generation bioethanol production. *Biomass Convers. Biorefin.* **10**, 639–650. <https://doi.org/10.1007/s13399-019-00468-4> (2020).
31. Bonatto, C. *et al.* Utilization of seawater and wastewater from shrimp production in the fermentation of papaya residues to ethanol. *Bioresour. Technol.* <https://doi.org/10.1016/j.biortech.2020.124501> (2021).
32. Baghel, R. S. Developments in seaweed biorefinery research: A comprehensive review. *Chem. Eng. J.* <https://doi.org/10.1016/j.cej.2022.140177> (2023).
33. Llano, T., Arce, C., Gallart, L. E., Perales, A. & Coz, A. Techno-economic analysis of macroalgae biorefineries: A comparison between ethanol and butanol facilities. *Ferment. Basel* <https://doi.org/10.3390/fermentation9040340> (2023).
34. Moniz, P., Martins, D., Oliveira, A. C., Reis, A. & da Silva, T. L. The biorefinery of the marine microalga *Cryptocodinium cohnii* as a strategy to valorize microalgal oil fractions. *Ferment. Basel* <https://doi.org/10.3390/fermentation8100502> (2022).
35. Dadwal, A., Sharma, S. & Satyanarayana, T. Recombinant cellobiohydrolase of *Myceliophthora thermophila*: Characterization and applicability in cellulose saccharification. *Amb Express* <https://doi.org/10.1186/s13568-021-01311-8> (2021).
36. Baramée, S. *et al.* A novel GH6 cellobiohydrolase from *Paenibacillus curdolanolyticus* B-6 and its synergistic action on cellulose degradation. *Appl. Microbiol. Biotechnol.* **101**, 1175–1188. <https://doi.org/10.1007/s00253-016-7895-8> (2017).
37. Cerda-Mejía, L., Valenzuela, S. V., Frias, C., Diaz, P. & Pastor, F. I. J. A bacterial GH6 cellobiohydrolase with a novel modular structure. *Appl. Microbiol. Biotechnol.* **101**, 2943–2952. <https://doi.org/10.1007/s00253-017-8129-4> (2017).
38. Cai, L. N., Lu, T., Lin, D. Q. & Yao, S. J. Discovery of extremophilic cellobiohydrolases from marine *Aspergillus niger* with computational analysis. *Process Biochem.* **115**, 118–127. <https://doi.org/10.1016/j.procbio.2022.02.016> (2022).
39. Takeda, M. *et al.* Metagenomic mining and structure-function studies of a hyper-thermostable cellobiohydrolase from hot spring sediment. *Commun. Biol.* <https://doi.org/10.1038/s42003-022-03195-1> (2022).
40. Bonugli-Santos, R. C. *et al.* Marine-derived fungi: Diversity of enzymes and biotechnological applications. *Front. Microbiol.* <https://doi.org/10.3389/fmicb.2015.00269> (2015).
41. Sinha, R. & Khare, S. K. Protective role of salt in catalysis and maintaining structure of halophilic proteins against denaturation. *Front. Microbiol.* **5**, 6. <https://doi.org/10.3389/fmicb.2014.00165> (2014).
42. Pramanik, S. *et al.* An engineered cellobiohydrolase I for sustainable degradation of lignocellulosic biomass. *Biotechnol. Bioeng.* **118**, 4014–4027. <https://doi.org/10.1002/bit.27877> (2021).
43. Contreras, F. *et al.* Engineering robust cellulases for tailored lignocellulosic degradation cocktails. *Int. J. Mol. Sci.* <https://doi.org/10.3390/ijms21051589> (2020).
44. Leadbeater, D. R. *et al.* Mechanistic strategies of microbial communities regulating lignocellulose deconstruction in a UK salt marsh. *Microbiome* <https://doi.org/10.1186/s40168-020-00964-0> (2021).
45. Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786. <https://doi.org/10.1038/nmeth.1701> (2011).
46. Lever, M. Carbohydrate determination with 4-hydroxybenzoic acid hydrazide (PAHBAH)—effect of bismuth on reaction. *Anal. Biochem.* **81**, 21–27. [https://doi.org/10.1016/0003-2697\(77\)90594-2](https://doi.org/10.1016/0003-2697(77)90594-2) (1977).
47. Gomez, L. D., Bristow, J. K., Statham, E. R. & McQueen-Mason, S. J. Analysis of saccharification in *Brachypodium distachyon* stems under mild conditions of hydrolysis. *Biotechnol. Biofuels* <https://doi.org/10.1186/1754-6834-1-15> (2008).

48. Herlet, J. *et al.* A new method to evaluate temperature vs. pH activity profiles for biotechnological relevant enzymes. *Biotechnol. Biofuels* **10**, 12. <https://doi.org/10.1186/s13068-017-0923-9> (2017).
49. Kurasin, M. & Valjamae, P. Processivity of cellobiohydrolases is limited by the substrate. *J. Biol. Chem.* **286**, 169–177. <https://doi.org/10.1074/jbc.M110.161059> (2011).
50. Kern, M. *et al.* Structural characterization of a unique marine animal family 7 cellobiohydrolase suggests a mechanism of cellulase salt tolerance. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 10189–10194. <https://doi.org/10.1073/pnas.1301502110> (2013).
51. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583. <https://doi.org/10.1038/s41586-021-03819-2> (2021).
52. Pettersen, E. F. *et al.* UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Sci.* **30**, 70–82. <https://doi.org/10.1002/pro.3943> (2021).
53. Jones, E., Oliphant, T. & Peterson, P. SciPy: Open source scientific tools for Python. (2001).
54. Pedregosa, F. *et al.* Scikit-learn: Machine learning in Python. *Journal of machine Learning research* **12**, 2825–2830 (2011).
55. Fox, J. M., Levine, S. E., Clark, D. S. & Blanch, H. W. Initial- and processive-cut products reveal cellobiohydrolase rate limitations and the role of companion enzymes. *Biochemistry* **51**, 442–452. <https://doi.org/10.1021/bi2011543> (2012).
56. Uchiyama, T. *et al.* Convergent evolution of processivity in bacterial and fungal cellulases. *Proc. Natl. Acad. Sci. U. S. A.* **117**, 19896–19903. <https://doi.org/10.1073/pnas.2011366117> (2020).
57. Hrmova, M. & Scherdt, J. G. Molecular mechanisms of processive glycoside hydrolases underline catalytic pragmatism. *Biochem. Soc. Trans.* **51**, 1387–1403. <https://doi.org/10.1042/bst20230136> (2023).
58. Jalak, J., Kurasin, M., Teugias, H. & Valjamae, P. Endo-exo synergism in cellulose hydrolysis revisited. *J. Biol. Chem.* **287**, 28802–28815. <https://doi.org/10.1074/jbc.M112.381624> (2012).
59. Vuong, T. V. & Wilson, D. B. Processivity, synergism, and substrate specificity of *Thermobifida fusca* Cel6B. *Appl. Environ. Microbiol.* **75**, 6655–6661. <https://doi.org/10.1128/aem.01260-09> (2009).
60. ReverbelLeroy, C., Pages, S., Belaich, A., Belaich, J. P. & Tardif, C. The processive endocellulase CelF, a major component of the *Clostridium cellulolyticum* cellulosome: Purification and characterization of the recombinant form. *J. Bacteriol.* **179**, 46–52. <https://doi.org/10.1128/jb.179.1.46-52.1997> (1997).
61. Parsiegla, G. *et al.* The crystal structure of the processive endocellulase CelF of *Clostridium cellulolyticum* in complex with a thioglucosaccharide inhibitor at 2.0 Å resolution. *Embo J.* **17**, 5551–5562. <https://doi.org/10.1093/emboj/17.19.5551> (1998).
62. Zhang, S., Irwin, D. C. & Wilson, D. B. Site-directed mutation of noncatalytic residues of *Thermobifida fusca* exocellulase Cel6B. *Eur. J. Biochem.* **267**, 3101–3115. <https://doi.org/10.1046/j.1432-1327.2000.01315.x> (2000).
63. Watson, B. J., Zhang, H. T., Longmire, A. G., Moon, Y. H. & Hutcheson, S. W. Processive endoglucanases mediate degradation of cellulose by *Saccharophagus degradans*. *J. Bacteriol.* **191**, 5697–5705. <https://doi.org/10.1128/jb.00481-09> (2009).
64. von Ossowski, I. *et al.* Engineering the exo-loop of *Trichoderma reesei* cellobiohydrolase, Cel7A: A comparison with *Phanerochaete chrysosporium* Cel7D. *J. Mol. Biol.* **333**, 817–829. [https://doi.org/10.1016/s0022-2836\(03\)00881-7](https://doi.org/10.1016/s0022-2836(03)00881-7) (2003).
65. Reyes-Ortiz, V. *et al.* Addition of a carbohydrate-binding module enhances cellulase penetration into cellulose substrates. *Biotechnol. Biofuels* **6**, 13. <https://doi.org/10.1186/1754-6834-6-93> (2013).
66. Nakamura, A. *et al.* Domain architecture divergence leads to functional divergence in binding and catalytic domains of bacterial and fungal cellobiohydrolases. *J. Biol. Chem.* **295**, 14606–14617. <https://doi.org/10.1074/jbc.RA120.014792> (2020).
67. Horn, S. J. *et al.* Costs and benefits of processivity in enzymatic degradation of recalcitrant polysaccharides. *Proc. Natl. Acad. Sci. U. S. A.* **103**, 18089–18094. <https://doi.org/10.1073/pnas.0608909103> (2006).
68. Cantarel, B. L. *et al.* The carbohydrate-active enzymes database (CAZy): An expert resource for glycogenomics. *Nucleic Acids Res.* **37**, D233–D238. <https://doi.org/10.1093/nar/gkn663> (2009).
69. Billard, H., Faraj, A., Lopes Ferreira, N., Menir, S. & Heiss-Blanquet, S. Optimization of a synthetic mixture composed of major *Trichoderma reesei* enzymes for the hydrolysis of steam-exploded wheat straw. *Biotechnol. Biofuels* **5**, 13. <https://doi.org/10.1186/1754-6834-5-9> (2012).
70. Zhang, T. *et al.* Identification of a haloalkaliphilic and thermostable cellulase with improved ionic liquid tolerance. *Green Chem.* **13**, 2083–2090. <https://doi.org/10.1039/c1gc15193b> (2011).
71. Asha, B. M. & Sakthivel, N. Production, purification and characterization of a new cellulase from *Bacillus subtilis* that exhibit halophilic, alkalophilic and solvent-tolerant properties. *Ann. Microbiol.* **64**, 1839–1848. <https://doi.org/10.1007/s13213-014-0835-x> (2014).
72. Cai, L. N., Xu, S. N., Lu, T., Lin, D. Q. & Yao, S. J. Salt-tolerant mechanism of marine *Aspergillus niger* cellulase cocktail and improvement of its activity. *Chin. J. Chem. Eng.* **28**, 1120–1128. <https://doi.org/10.1016/j.cjche.2019.11.012> (2020).
73. Baumann, H., Wallace, R. B., Tagliaferri, T. & Gobler, C. J. Large natural pH, CO₂ and O₂ fluctuations in a temperate tidal salt marsh on diel, seasonal, and interannual time scales. *Estuaries Coasts* **38**, 220–231. <https://doi.org/10.1007/s12237-014-9800-y> (2015).
74. Prabmark, K. *et al.* Enhancement of catalytic activity and alkaline stability of cellobiohydrolase by structure-based protein engineering. *3 Biotech* <https://doi.org/10.1007/s13205-022-03339-4> (2022).
75. Yin, Y. R. *et al.* Heterologous expression and characterization of a novel halotolerant, thermostable, and alkali-stable GH6 endoglucanase from *Thermobifida halotolerans*. *Biotechnol. Lett.* **37**, 857–862. <https://doi.org/10.1007/s10529-014-1742-8> (2015).
76. Paul, S., Bag, S. K., Das, S., Harvill, E. T. & Dutta, C. Molecular signature of hypersaline adaptation: Insights from genome and proteome composition of halophilic prokaryotes. *Genome Biol.* **9**, 19. <https://doi.org/10.1186/gb-2008-9-4-r70> (2008).
77. Jaenicke, R. Protein stability and molecular adaptation to extreme conditions. *Eur. J. Biochem.* **202**, 715–728. <https://doi.org/10.1111/j.1432-1033.1991.tb16426.x> (1991).
78. Amoozegar, M. A., Siroosi, M., Atashgahi, S., Smidt, H. & Ventosa, A. Systematics of haloarchaea and biotechnological potential of their hydrolytic enzymes. *Microbiol.-Sgm* **163**, 623–645. <https://doi.org/10.1099/mic.0.000463> (2017).
79. Sidar, A. *et al.* Carbohydrate binding modules: Diversity of domain architecture in amylases and cellulases from filamentous microorganisms. *Front. Bioeng. Biotechnol.* <https://doi.org/10.3389/fbioe.2020.00871> (2020).
80. Lehmann, C. *et al.* Reengineering CelA2 cellulase for hydrolysis in aqueous solutions of deep eutectic solvents and concentrated seawater. *Green Chem.* **14**, 2719–2726. <https://doi.org/10.1039/c2gc35790a> (2012).
81. Grande, P. M. & de Maria, P. D. Enzymatic hydrolysis of microcrystalline cellulose in concentrated seawater. *Bioresour. Technol.* **104**, 799–802. <https://doi.org/10.1016/j.biortech.2011.10.071> (2012).
82. Indira, D., Sharmila, D., Balasubramanian, P., Thirugnanam, A. & Jayabalan, R. Utilization of sea water based media for the production and characterization of cellulase by *Fusarium subglutinans* MTCC 11891. *Biocat. Agri. Biot.* **7**, 187–192. <https://doi.org/10.1016/j.bcab.2016.06.006> (2016).
83. Bano, A. *et al.* Purification and characterization of cellulase from obligate halophilic *Aspergillus flavus* (TISTR 3637) and its prospects for bioethanol production. *Appl. Biochem. Biotechnol.* **189**, 1327–1337. <https://doi.org/10.1007/s12010-019-03086-y> (2019).
84. Pasin, T. M. *et al.* A halotolerant endo-1,4-β-xylanase from *Aspergillus clavatus* with potential application for agroindustrial residues saccharification. *Appl. Biochem. Biotechnol.* **191**, 1111–1126. <https://doi.org/10.1007/s12010-020-03232-x> (2020).
85. Zhao, B., Al Rasheed, H., Ali, I. & Hu, S. L. Efficient enzymatic saccharification of alkaline and ionic liquid-pretreated bamboo by highly active extremozymes produced by the co-culture of two halophilic fungi. *Bioresour. Technol.* <https://doi.org/10.1016/j.biortech.2020.124115> (2021).

86. Arakawa, T., Yamaguchi, R., Tokunaga, H. & Tokunaga, M. Unique features of halophilic proteins. *Curr. Prot. Pept. Sci.* **18**, 65–71. <https://doi.org/10.2174/1389203717666160617111140> (2016).
87. Mevarech, M., Frolow, F. & Gloss, L. M. Halophilic enzymes: Proteins with a grain of salt. *Biophys. Chem.* **86**, 155–164. [https://doi.org/10.1016/s0301-4622\(00\)00126-5](https://doi.org/10.1016/s0301-4622(00)00126-5) (2000).
88. Karan, R., Capes, M. D. & DasSarma, S. Function and biotechnology of extremophilic enzymes in low water activity. *Aquat. Biosyst.* **8**, 1–15 (2012).
89. Mu, Y. H. *et al.* Surface charge engineering of 6-glucosidase using rational design improves catalytic capacity and ionic liquid tolerance. *J. Mol. Liquids* <https://doi.org/10.1016/j.molliq.2022.120577> (2022).
90. Qiu, J. J., Han, R. & Wang, C. Microbial halophilic lipases: A review. *J. Basic Microbiol.* **61**, 594–602. <https://doi.org/10.1002/jobm.202100107> (2021).
91. Mokashe, N., Chaudhari, B. & Patil, U. Operative utility of salt-stable proteases of halophilic and halotolerant bacteria in the biotechnology sector. *Int. J. Biol. Macromol.* **117**, 493–522. <https://doi.org/10.1016/j.ijbiomac.2018.05.217> (2018).
92. Kasirajan, L. & Maupin-Furlow, J. A. Halophilic archaea and their potential to generate renewable fuels and chemicals. *Biotechnol. Bioeng.* **118**, 1066–1090. <https://doi.org/10.1002/bit.27639> (2021).
93. Zhao, J. Q., Guo, C., Zhang, L. & Tian, C. G. Biochemical and functional characterization of a novel thermoacidophilic, heat and halo-ionic liquids tolerant endo- β -1,4-glucanase from saline-alkaline lake soil microbial metagenomic DNA. *Int. J. Biol. Macromol.* **118**, 1035–1044. <https://doi.org/10.1016/j.ijbiomac.2018.06.141> (2018).
94. Sinha, S. K., Datta, M. & Datta, S. A glucose tolerant β -glucosidase from *Thermomicrobium roseum* that can hydrolyze biomass in seawater. *Green Chem.* **23**, 7299–7311. <https://doi.org/10.1039/d1gc01357b> (2021).

Author contributions

D.R.L. designed and conducted experiments and analyzed the data, wrote and edited the manuscript. N.C.B. conceived the idea, provided expertise and edited the manuscript.

Funding

DRL was supported by a studentship from the BBSRC Doctoral Training Program (BB/J014443/1) and DRL and NCB acknowledges funding from the Biotechnology and Biological Sciences Research Council (BBSRC), UK (BB/W000695/1).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to D.R.L. or N.C.B.

Reprints and permissions information is available at 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