



**CHALMERS**  
UNIVERSITY OF TECHNOLOGY

## **A unique AA5 alcohol oxidase fused with a catalytically inactive CE3 domain from the bacterium *Burkholderia pseudomallei***

Downloaded from: <https://research.chalmers.se>, 2024-04-28 03:53 UTC



Citation for the original published paper (version of record):

Mazurkewich, S., Seveso, A., Larsbrink, J. (2023). A unique AA5 alcohol oxidase fused with a catalytically inactive CE3 domain from the bacterium *Burkholderia pseudomallei*. *FEBS Letters*, 597(13): 1779-1791.  
<http://dx.doi.org/10.1002/1873-3468.14632>

N.B. When citing this work, cite the original published paper.

## RESEARCH LETTER

# A unique AA5 alcohol oxidase fused with a catalytically inactive CE3 domain from the bacterium *Burkholderia pseudomallei*

 Scott Mazurkewich , Andrea Seveso and Johan Larsbrink 

Division of Industrial Biotechnology, Department of Life Sciences, Wallenberg Wood Science Center, Chalmers University of Technology, Gothenburg, Sweden

## Correspondence

S. Mazurkewich or J. Larsbrink, Division of Industrial Biotechnology, Department of Life Sciences, Wallenberg Wood Science Center, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden  
 Tel: +46317723843 or +46317723839  
 E-mail: [scott.mazurkewich@chalmers.se](mailto:scott.mazurkewich@chalmers.se) (SM) or [johan.larsbrink@chalmers.se](mailto:johan.larsbrink@chalmers.se) (JL)

Scott Mazurkewich and Andrea Seveso contributed equally to this article

(Received 3 March 2023, revised 19 April 2023, accepted 23 April 2023, available online 11 May 2023)

doi:10.1002/1873-3468.14632

Edited by Peter Brzezinski

Copper radical oxidases (CROs) are redox enzymes able to oxidize alcohols or aldehydes, while only requiring a single copper atom as cofactor. Studied CROs are found in one of two subfamilies within the Auxiliary Activities family 5 (AA5) in the carbohydrate-active enzymes database. We here characterize an AA5 enzyme outside the subfamily classification from the opportunistic bacterial pathogen *Burkholderia pseudomallei*, which curiously was fused to a carbohydrate esterase family 3 domain. The enzyme was shown to be a promiscuous primary alcohol oxidase, with an activity profile similar to enzymes from subfamily 2. The esterase domain was inactive on all tested substrates, and structural predictions revealed this being an effect of crippling substitutions in the expected active site residues.

**Keywords:** alcohol oxidase; Auxiliary Activity family 5; *Burkholderia*; copper-radical oxidase; multidomain enzyme

The use of redox enzymes to replace chemical reactions in industrial settings could reduce hazardous waste or improve process efficiency by harnessing the specificity and regioselectivity of enzymes [1,2]. Copper-radical oxidases (CROs) are attractive oxidizing tools, using oxygen as the final electron acceptor and producing hydrogen peroxide [3], and only requiring a copper atom in the active site [4]. In the Carbohydrate-Active Enzymes database (CAZy, [www.cazy.org](http://www.cazy.org), [5]) CROs are classified into the Auxiliary Activity family 5 (AA5). AA5 was recently divided into two subfamilies; subfamily 1 (AA5\_1) comprises characterized (methyl)-glyoxal oxidases (GLOx, EC1.2.3.15) [6], active on a variety of simple aldehydes,  $\alpha$ -hydroxycarbonyl or  $\alpha$ -dicarbonyl compounds generating the corresponding

carboxylic acids [3,7–9], and galactose oxidases (GalOx, EC1.1.3.9) [10,11]. Subfamily 2 (AA5\_2) has been more studied and several activities have been described: GalOx [12–14], raffinose oxidase (RafOx, EC1.1.3.-) [15], aryl alcohol oxidase (AAO, EC1.1.3.7) [16,17], and general alcohol oxidase (AlcOx, EC1.1.3.13) [18–23]. Some AA5\_2 members oxidize their primary alcohol substrates to only the corresponding aldehydes and not to carboxylic acids [24], making them emerge as attractive biocatalysts within green chemistry and biotechnology.

The first characterized AA5\_2 CRO was the galactose oxidase *FgrGalOx* from *Fusarium graminearum*, which oxidizes the C6-hydroxyl group in D-galactose [13]. This enzyme has been extensively studied and has

## Abbreviations

AA5, Auxiliary Activities family 5; AAO, aryl alcohol oxidase; AlcOx, alcohol oxidase; CRO, copper-radical oxidase; HMF, 5-hydroxymethylfurfural; HMFA, 5-hydroxymethyl-2-furancarboxylate.

seen utilization in applications such as glycoprotein labelling [25,26], development of biosensors [27], mono- and polysaccharide modification [28–30] and even in synthesis of an anti-HIV drug [31]. More recently, two characterized AA5\_2 enzymes from *Colletotrichum graminicola* were found to be essentially inactive on galactose [18]. The first, *CgrAAO*, instead oxidizes aromatic primary alcohols and efficiently converts 5-hydroxymethylfurfural (HMF) to 2,5-diformylfuran, a compound with several uses in biopolymer synthesis [16]. The second, *CgrAlcOx*, showed higher activity on the primary hydroxyl group of diverse aliphatic alcohols, and has been evaluated for production of aldehydes in flavours and fragrances [22]. Since the discovery of the *C. graminicola* enzymes, other AA5\_2 members with similar substrate specificity profiles were characterized from the same genus (*C. tabacum*, *C. destructivum*, *C. orbiculare*, *C. higginsianum*) and from *Pyricularia oryzae* [19,20,23]. Other enzymes from *Fusarium* have shown preference for aryl alcohols and furans in contrast to the well-known *FgrGalOx* [17]. More recently, the combination of phylogenetic and sequence similarity network analysis allowed a broad sampling across different fungal clades and led to characterization of 12 new AA5 enzymes, including GalOx, RafOx, and more AlcOx enzymes [21].

While CAZy has identified putative AA5 members across all kingdoms of life, the lack of biochemically characterized enzymes limits functional predictions, and further exploration of its sequence diversity is needed. To date, the vast majority of characterized AA5 members are from eukaryotes and, apart from a few plant enzymes from *Arabidopsis thaliana* [10], these are predominantly fungal. Bacterial enzymes comprise over a third of AA5 but compared to fungal enzymes, few bacterial CROs have been studied. The first was from *Stigmatella aurantiaca*, and encoded by the gene *fbfB* [32]. It was annotated as a putative GalOx through the identification of several Kelch motifs typical of the  $\beta$ -propeller architecture of CROs, together with conserved metal-binding residues. The enzyme was not characterized *in vitro*, but *in vivo* studies demonstrate a role in the formation of fruiting bodies. With a similar approach, the CRO SCO2837p was identified in *Streptomyces coelicolor*, and shown to play a role in cell wall remodelling associated with hyphal tip growth [33]. The enzyme was also assayed on a range of simple alcohols and had the highest activity on glycolaldehyde. A comparable activity profile was displayed by the *Streptomyces lividans* GlxA (*SlGlxA*), with the highest specificity for glycolaldehyde and only weak activity on glucose and galactose [34]. Notably, *SlGlxA* represents the first structurally determined bacterial CRO.

A bacterium encoding several putative CROs is *Burkholderia pseudomallei*, a soil-dwelling Gram-negative opportunistic pathogen from the Pseudomonadota phylum. Infections by *B. pseudomallei* can lead to Melioidosis, with symptoms ranging from mild to severe pneumonia and possibly fatal septic shock [35,36]. While the species occupies various ecological niches, it is not regarded as a major plant degrader, and research has been focused on understanding its virulence factors and role in pathogenesis [37]. Interestingly, *B. pseudomallei* encodes three putative unstudied AA5 enzymes, one of which being fused to a putative catalytic Carbohydrate Esterase family 3 (CE3) domain. CE3 comprises acetyl xylan esterases that cleave *O*-linked acetyl groups from xylan oligo- and polysaccharides [38–40], making the fusion to an AA5 domain puzzling. Here, we present our characterization of this unusual multidomain enzyme, where the oxidase was shown to oxidize carbohydrates but was more active on longer aliphatic alcohols, furans, and polyols. The CE3 domain was in contrast completely inactive, and from structure predictions we postulate that this is an effect of peculiar substitutions in the positions of the expected catalytic residues. We demonstrate that the AA5 domain, *BpAlcOx*, is active on primary alcohols, and that it appears unable to oxidize the closed-ring conformation of aldoses but rather only oxidizes open-chain forms.

## Materials and methods

### Phylogenetic analysis

All AA5 and CE3 sequences were accessed from CAZy (February 2023) and trimmed to comprise only catalytic domains. Identical sequences and fragments were removed, and remaining sequences for each family aligned using CLUSTAL OMEGA [41]. Phylogenetic trees were constructed using IQ-TREE [42], with default settings, including choice of optimal substitution model (WAG+F+I+G4 for both families) and 1000 ultrafast bootstraps. Trees were visualized using ITOL [43].

### Molecular biology

The Uniprot entry for *BpAlcOx*-CE3 (Q63UA8) is longer at the N-terminus by 10 residues compared to the NCBI record (AIV61401.1). SIGNALP 5.0 [44] predicts the same signal peptide cleavage site in both entries but with greater confidence for the UniProt entry which thus is the more likely reading frame (Fig. S1), and residue numbering refers to the UniProt entry. The gene encoding *BpAlcOx*-CE3, lacking its predicted signal peptide, was codon optimized (Table S1) for expression in *Escherichia coli* and synthesized (Twist

Bioscience, San Francisco, CA, USA) into a pET-28a vector in-frame with an N-terminal hexahistidine tag. Constructs to express the individual *BpAlcOx* or *BpCE3* domains were created by ligating their respective PCR products into modified pET-28a vectors containing an N-terminal hexahistidine tag and a TEV protease cleavage site instead of the native thrombin site. Protein variants were created by site-specific mutagenesis by the QuikChange method [45]. Primer sequences utilized for gene amplifications and mutagenesis are provided in Table S2. All constructs were verified by DNA sequencing.

## Protein production and purification

*Escherichia coli* BL21 ( $\lambda$ DE3) cells harbouring plasmids encoding either *BpAlcOx*-CE3, *BpAlcOx*, or *BpCE3* were grown in lysogeny broth (LB) supplemented with 50  $\mu\text{g}\cdot\text{mL}^{-1}$  neomycin at 37 °C with 200 r.p.m. shaking to an  $\text{OD}_{600} \sim 0.5$ , when expression was induced by addition of 0.2 mM isopropyl  $\beta$ -D-1-thiogalactopyranoside (IPTG), and the cells were incubated at 16 °C overnight, then harvested by centrifugation (5000 g, 10 min), resuspended in 20 mM tris(hydroxymethyl)aminomethane (Tris) buffer, pH 8, containing 250 mM NaCl, and disrupted by sonication. Cell debris was removed by centrifugation (18 000 g, 10 min) and proteins were purified using immobilized metal ion affinity chromatography (IMAC) on an ÄKTA system (Cytiva, Marlborough, MA, USA) with 5 mL HisTrap Excel columns, using 50 mM Tris pH 8 with 250 mM NaCl as binding buffer and elution using a linear gradient of the same buffer containing 250 mM imidazole. Collected proteins were concentrated by ultrafiltration (Amicon Ultra-15; Merck-Millipore, Burlington, MA, USA), loaded onto a HiLoad Superdex 200 16/60 gel filtration column and resolved with an isocratic gradient with the IMAC binding buffer. Protein samples were treated with 1 mM EDTA overnight, 4 °C, to remove exogenous metals, buffer exchanged by passing through a Cytiva HiPrep 26/10 Desalting Column, then treated with 500  $\mu\text{M}$   $\text{CuCl}_2$  overnight, 4 °C, and buffer exchanged into 25 mM Tris pH 8 with 250 mM NaCl. Protein samples were concentrated by ultrafiltration as before and stored at 4 °C. Sodium dodecyl sulfate polyacrylamide gel electrophoresis using Mini-PROTEAN TGX Stain-Free Gels (Bio-Rad, Solna, Sweden) was used to verify protein purity and protein concentrations were determined using a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA).

## Enzyme characterization

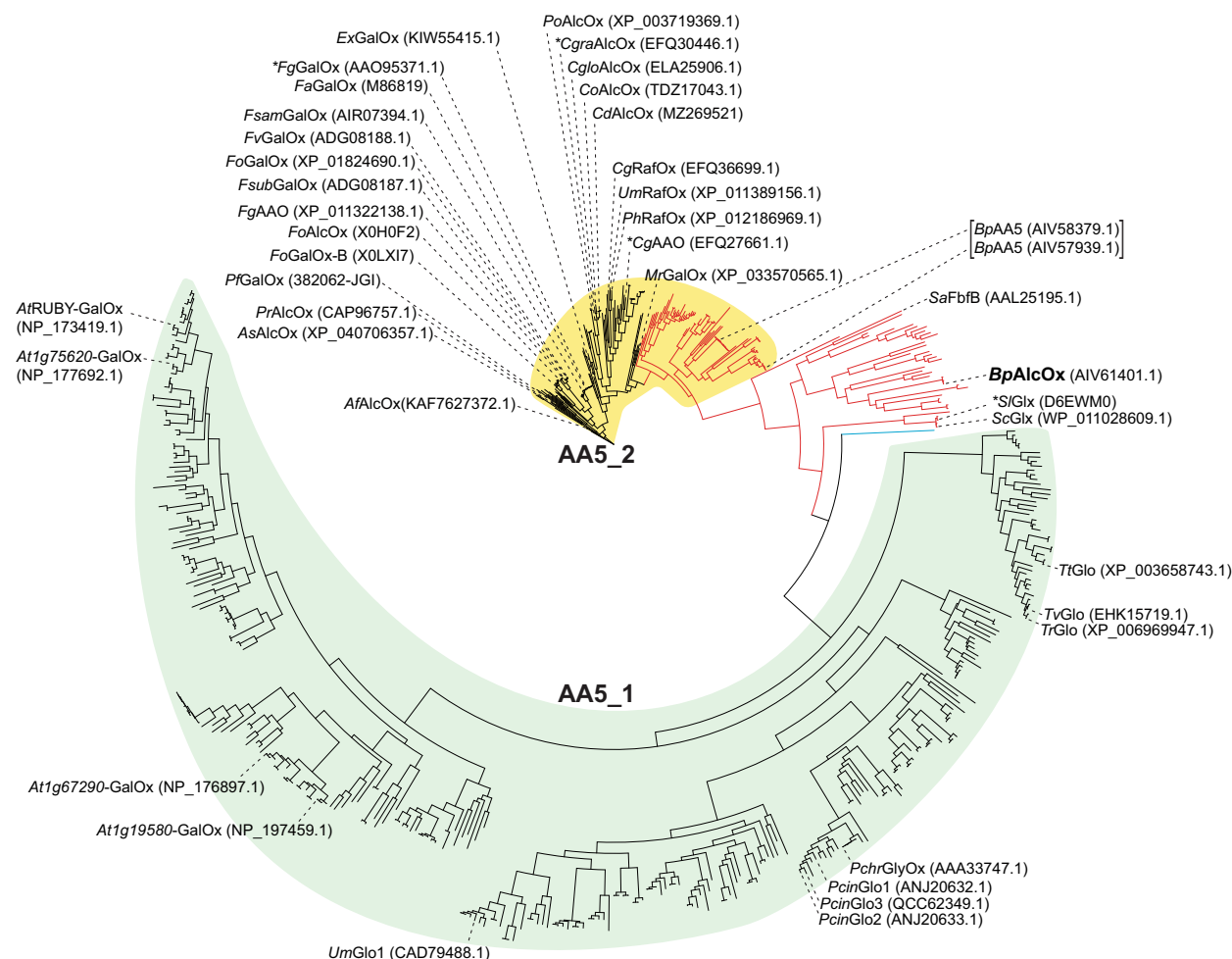
The oxidase activity of *BpAlcOx* was monitored continuously by coupling  $\text{H}_2\text{O}_2$  production to horseradish peroxidase (HRP) oxidation of 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid) (ABTS). Standard assay reactions of 200  $\mu\text{L}$  were carried out in 50 mM sodium citrate pH 7, with 1 mM

ABTS (Sigma Aldrich, Solna, Sweden) and 1  $\mu\text{g}$  HRP (Thermo Scientific). Kinetic measurements were performed in at least duplicate at 22 °C and followed at 420 nm using the extinction coefficient of oxidized ABTS:  $3.6 \times 10^4 \text{ M}^{-1}\cdot\text{cm}^{-1}$ . Substrates were from either Fisher Scientific, Sigma Aldrich, or BioSynth (Compton, England) and were dissolved in either water or DMSO, in the case of aromatic substrates and the long-chain alkyl alcohols hexanol and nonanol. DMSO concentrations up to 10% (v/v) did not affect oxidase activity with ethanol and thus all assays where substrates were dissolved in DMSO kept the solvent concentrations  $\leq 10\%$  (v/v). Beyond the substrates tested in Table S3, 2-propanol and the following aldehydes were assayed but gave no detectable activity over 1 h: ethanal, propanal, butanal, pentanal, hexanal, and benzaldehyde. The pH dependency was determined with 250 mM 1,3-propanediol (saturating conditions) in a three-component buffer (25 mM acetic acid, 25 mM 2-(*N*-morpholino)ethanesulfonic acid, and 50 mM Tris-HCl), covering pH 4.5–9.5 [46]. Acetyl esterase activity was assayed using 4-nitrophenyl acetate (*p*NP-Ac; Sigma Aldrich), in sodium phosphate pH 7.5, with *p*NP release monitored at 405 nm using the extinction coefficient  $13.3 \text{ mM}^{-1}\cdot\text{cm}^{-1}$ . Nonlinear data were fitted to either the Michaelis–Menten or substrate inhibition equations using GRAPHPAD PRISM (GraphPad, Boston, MA, USA) or ORIGINPRO (OriginLab, Northampton, MA, USA). For unsaturable reactions, linear regression of velocity over substrate concentration was used to determine  $k_{\text{cat}}/K_{\text{m}}$  values.

## Results

### Bioinformatic analyses

The fusion AA5 and CE3 proteins appear unusual, and further sequence analysis showed that homologous sequences are limited to *B. pseudomallei* strains and one instance from *B. savannae* (91% sequence identity). The individual domains of *BpAlcOx*-CE3 were then analysed, excluding *Burkholderiales*, and the closest homologs to the AA5 domain were found in other Pseudomonadota species, namely *Luteimonas panaciterrae* (48.5% seq. id.) and species of *Lysobacter* ( $\leq 46.5\%$  seq. id.), while for the CE3 domain, top homologs were from various *Streptomyces* species from Actinomycetota, with sequence identities of  $\leq 42\%$ . A phylogenetic tree of AA5 was constructed to visualize whether *BpAlcOx* clusters with previously characterized enzymes (simplified in Fig. 1, and with all entries and bootstrap values in Fig. S2). Following the subfamily annotations in CAZy, *BpAlcOx* is found in a clade between the two subfamilies and not close to previously characterized enzymes. It is found in a larger bacterial clade where *S. aurantiaca* FbfB is also positioned, to which



**Fig. 1.** Phylogenetic tree of AA5. Bacterial enzymes are indicated by red branches, eukaryotic in black, and the sole archaeal in blue. AA5\_1 sequences are shown in green shading and AA5\_2 in yellow. Characterized members are indicated with their names and respective Genbank (or Uniprot) identifiers as in [68], and structurally determined enzymes are further marked with asterisks. *BpAlcOx* is highlighted in bold, and the two additional uncharacterized AA5 modules from *Burkholderia pseudomallei* are shown within square brackets.

*BpAlcOx* has 29% sequence identity. The other two AA5 enzymes of *B. pseudomallei* in contrast belong to AA5\_2, though both are also located in an unexplored subfamily clade (Fig. 1; Fig. S2). A similar phylogenetic analysis of the enzyme's CE3 domain was less informative (Fig. S3), as few CE3 members have to date been characterized [5], and *BpCE3* is not closely related to any of these.

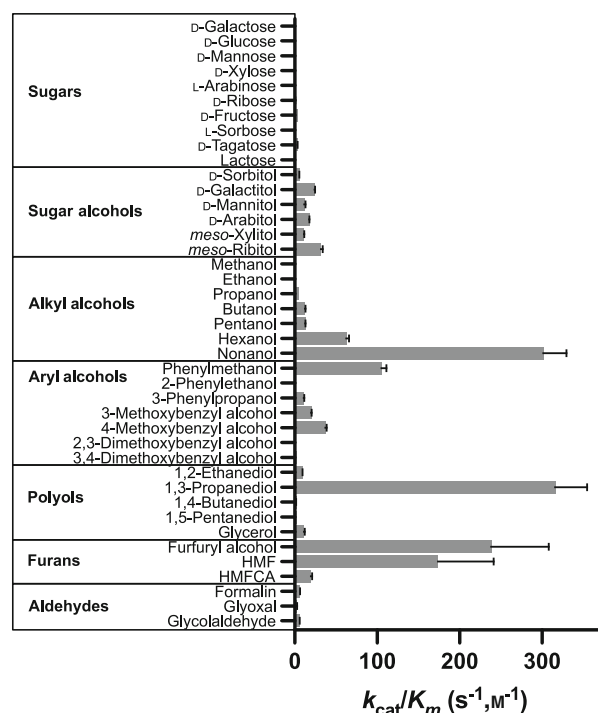
### Biochemical characterization

We successfully recombinantly produced both the full-length and individual domains of *BpAlcOx*-CE3. The activity of the isolated AA5 catalytic domain ( $11.84 \pm 0.862 \text{ s}^{-1}$ ) with a saturating amount of 1,3-propanediol (250 mM), one of its preferred substrates,

was very similar to that of the full-length protein ( $8.579 \pm 0.296 \text{ s}^{-1}$ ), with the small differences likely attributable to experimental error, and thus the isolated catalytic domain was utilized for in-depth characterizations. The pH dependence on the CRO activity with 1,3-propanediol was bell-shaped with a maximal activity between pH 6–8 (Fig. S4), and pH 7 was used as the standard condition for subsequent assays using a wide range of alcohols known to be CRO substrates (Fig. 2).

As a potential GalOx, *BpAlcOx* was assayed on galactose and other carbohydrates. The enzyme oxidized D-galactose and other monosaccharides (Table S3), though these reactions could not be saturated and the specificity constant for galactose was minimal ( $0.2179 \pm 0.00514 \text{ M}^{-1} \cdot \text{s}^{-1}$ ) compared to





**Fig. 2.** Specificity constants of *BpAlcOx* with a range of compounds. Assays were completed in sodium citrate at pH 7 at 25 °C as described in the [Materials and methods](#). Kinetic parameters for the substrates where saturating conditions could be obtained are listed in Table 1, and the parameters for all the compounds tested here are supplied in Table S3.

previously characterized AA5\_2 GalOx enzymes such as *FgrGalOx* ( $11\,900 \pm 1320\,M^{-1}s^{-1}$  [47]). Notably, *BpAlcOx* has comparable activity with D-galactose, D-glucose, D-mannose, L-arabinose, and D-xylose, where the latter lacks a primary alcohol in the pyranose form leading to a hypothesis that the enzyme acts on non-pyranose forms of these sugars. Aldoses can exist in pyranose-, furanose- or open-chain forms, but usually the latter represents only a small percentage in solution. For example, in solution glucose is found approximately in 99.6 : 0.4 : < 0.1% ratios as pyranose : furanose : open-chain [48,49]. The equilibria of some monosaccharides tested here are supplied in Table S4. To explore the hypothesis, we tested a range of open-chain sugar alcohols and observed a considerable increase in specificity constant relative to the respective monosaccharides and, notably, observed saturation kinetics for mannitol and ribitol (Table 1; Fig. S5). No activity was observed on fucose, an aldose lacking a primary alcohol in either the open or cyclic forms, nor on the non-reducing disaccharides sucrose and trehalose, which also lack open chain forms. Minimal activity was however observed on

lactose, where the reducing-end glucose moiety can exist in an open configuration, indicating that the inactivity on non-reducing disaccharides is not caused by the carbohydrate length. *BpAlcOx* was inactive on the polysaccharides glucomannan, galactomannan, and galactoglucomannan, consistent with requirement for non-cyclic aldose substrates. Interestingly, the specificity constant of *BpAlcOx* for the ketoses D-fructose and D-tagatose was  $\geq 10$ -fold higher than that of the aldoses, and *BpAlcOx* could be saturated with D-tagatose ( $K_m$  66 mM). However, the ketose L-sorbose was a poor substrate with a specificity constant similar to those for aldoses.

*BpAlcOx* was then assayed on a range of primary alkyl alcohols but could again not be saturated with any tested substrate. Assays revealed an increased specificity for longer alkyl chain length with the specificity constant for nonanol being  $\sim 100$ - or  $\sim 10$ -fold greater than for ethanol or hexanol, respectively, and over 1000-fold higher than for galactose (Fig. 2; Table S3). The enzyme was active on aryl alcohols where, interestingly, increased chain lengths of the alkyl group negatively affected the activity, with the specificity constant for phenylmethanol being  $\sim 1000$ - and  $\sim 10$ -fold greater than that for 2-phenylethanol and 3-phenylpropanol, respectively, and apparently the presence of methoxy substituents on the aryl group diminishes the activity. *BpAlcOx* was active also on the aldehydes glyoxal and formaldehyde but lacked detectable activity on benzyl aldehyde or other tested alkyl aldehydes. In aqueous solutions, aldehydes exist in equilibrium with their hydrated geminal diol forms and the equilibrium is shifted towards aldehyde functionality with increasing chain length, but for both the small glyoxal and formaldehyde compounds, the hydrated forms are predominant in solution [50,51]. Given the significant activity observed with longer chain alcohol substrates, and the lack of comparable activity with their respective aldehydes, we postulate that the hydrated geminal diol forms of glyoxal and formaldehyde may be the actual substrates instead of their aldehyde forms as similarly proposed for AA5 enzymes [9]. The enzyme was active with glycolaldehyde, a substrate shown to be preferred by the previously characterized bacterial AA5 member SCO2837p from *S. coelicolor* [33]. Although SCO2837p preferred this small compound as a substrate, it did not saturate the enzyme. *BpAlcOx* could in contrast be saturated by glycolaldehyde leading to the lowest  $K_m$  value ( $13.89 \pm 1.48\,mM$ ) of all substrates tested but the turnover rate ( $0.07737 \pm 0.00253\,s^{-1}$ ) was more than 100-fold lower than some other saturating compounds, such as 1,3-propanediol described in more detail

**Table 1.** Kinetic parameters of *BpAlcOx* with alcohol and aldehyde substrates. Substrates which led to saturating conditions are presented here, and the parameters for all the compounds tested in this study are supplied in Table S3. Assays were completed in sodium citrate at pH 7 at 25 °C as described in the Materials and methods. No activity was detected with 2-propanol, ethanal, propanal, butanal, pentanal, hexanal, or benzaldehyde. ND, not detected.

| Substrate        | $K_m$ (mM)   | $K_{si}$ (mM) | $k_{cat}$ (s <sup>-1</sup> ) | $k_{cat}/K_m$ (M <sup>-1</sup> ·s <sup>-1</sup> ) |
|------------------|--------------|---------------|------------------------------|---|
| meso-Ribitol     | 72.58 ± 4.43 | ND            | 2.312 ± 0.0516               | 31.86 ± 2.07                                      |
| D-Mannitol       | 99.66 ± 10.2 | ND            | 1.151 ± 0.0632               | 11.55 ± 1.34                                      |
| D-Tagatose       | 67.08 ± 3.34 | ND            | 0.2205 ± 0.00341             | 3.288 ± 0.171                                     |
| 1,3-Propanediol  | 25.03 ± 2.87 | ND            | 7.909 ± 0.332                | 316.0 ± 38.5                                      |
| Glycerol         | 211.7 ± 16.7 | ND            | 2.355 ± 0.0882               | 11.12 ± 0.971                                     |
| Furfuryl alcohol | 24.46 ± 5.76 | 37.27 ± 9.92  | 5.829 ± 1.02                 | 238.3 ± 69.8                                      |
| HMF              | 74.06 ± 22.2 | 39.07 ± 13.6  | 12.86 ± 3.18                 | 173.7 ± 67.5                                      |
| HMFA             | 58.98 ± 3.51 | ND            | 1.159 ± 0.0238               | 19.66 ± 1.24                                      |
| Glycolaldehyde   | 13.89 ± 1.48 | ND            | 0.07737 ± 0.00253            | 5.570 ± 0.621                                     |
| Glyoxal          | 70.49 ± 5.55 | ND            | 0.1762 ± 0.00552             | 2.499 ± 0.212                                     |

below, leading to a low specificity constant similar to glyoxal.

The recent characterization of a unique AA5\_2 enzyme from *C. graminicola* (*CgrAAO*) highlighted the specificity of the oxidase towards furan substrates [16,18]. *BpAlcOx* was also active on furan alcohols and had saturation kinetics leading to amongst the highest observed specificity constants (Fig. 2, Table 1). However, compared to *CgrAAO*, the turnover numbers ( $k_{cat}$ ) for *BpAlcOx* were at least 10-fold lower. Substrate inhibition was observed for the *BpAlcOx* catalysed reactions with both furfuryl alcohol and HMF with inhibitory constant ( $K_{si}$ ) values close to the observed  $K_m$  values. To the best of our knowledge, substrate inhibition has not been observed in AA5 with furan substrates. Notably, no substrate inhibition was observed with 5-hydroxymethyl-2-furancarboxylate (HMFA) indicating that the carboxylate restricts the unproductive binding modes leading to the substrate inhibition observed with the planar cyclic molecule. *BpAlcOx* was active on HMFA but inactive on the corresponding aldehyde, 5-formyl-2-furancarboxylate, and lacked activity with furfural, further supporting the previous observation of *BpAlcOx* not oxidizing aldehydes.

*BpAlcOx* was active on diols and glycerol, with the greatest activity on 1,3-propanediol (Fig. 2, Table 1). Interestingly, saturation kinetics was observed for glycerol and 1,3-propanediol but not any other diol. The  $K_m$  value for 1,3-propanediol was similar to that for furfuryl alcohol (25 mM) though a higher  $k_{cat}$  for the former resulted in the highest observed  $k_{cat}/K_m$  value. The specificity constant for 1,3-propanediol is ~100-fold higher than that of 1-propanol and ~30-fold higher than that of glycerol, primarily resulting from a ~10-fold increase in  $K_m$  for the latter. Activity on

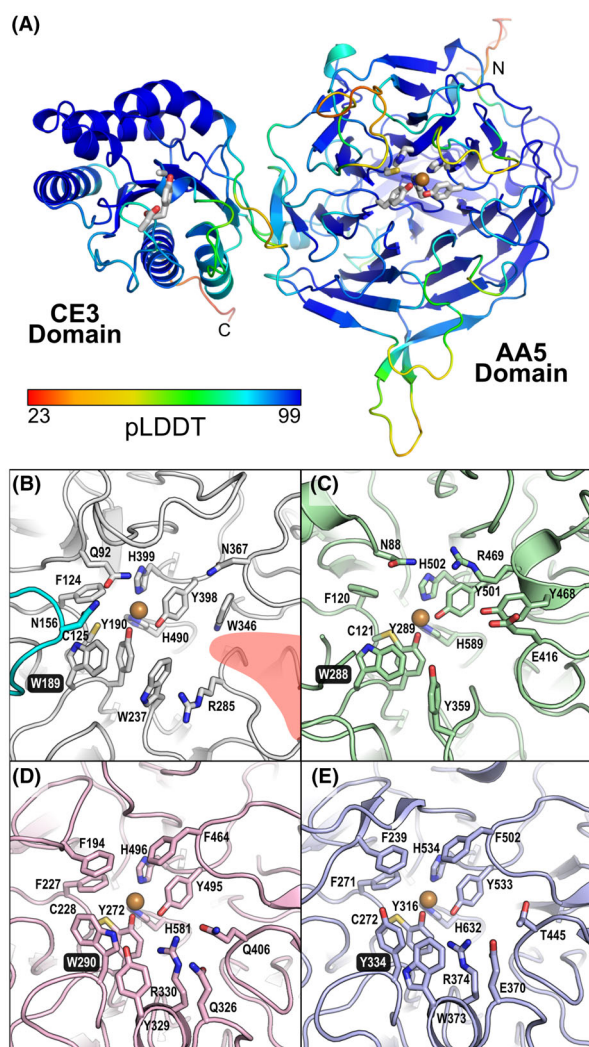
diols and glycerol has been observed for other AA5\_2 members such as *CgrAAO* [18], but the clear specificity for 3-carbon units of *BpAlcOx* is unique within AA5 thus far.

The CE3 domain was produced as an isolated protein like the AA5 domain, but neither the full-length protein nor the isolated CE3 domain had detectable esterase activity. As discussed below, structural predictions suggest this is not an experimental artefact but an effect of the lack of an expected functional catalytic triad.

### Structural predictions and analyses

Attempts to crystallize either the full-length enzyme or the isolated AA5 domain were unsuccessful, and we instead predicted its structure using ALPHAFOLD2 [52,53], ESMFOLD [54], and transform-restrained Rosetta [55–58]. These all gave reasonably high confidence models for both the AA5 and CE3 domains, with the main differences being the orientation of the domains relative to each other (Fig. S6), and we chose the ALPHAFOLD2 model for further analyses (found on UniProt, accession Q63UA8). As anticipated by the primary sequence, the overall structure contains two distinct domains: an N-terminal  $\beta$ -propeller integrated with an immunoglobulin-like domain consistent with structurally determined AA5 members followed by a short linker of ~12 residues to a C-terminal SGNH-hydrolase domain consistent with CE3 enzymes (Fig. 3A).

*BpAlcOx* does not contain an additional domain N-terminally linked to the catalytic domain, like the CBM32 domain in the *F. graminicola* GalOx [17,59] or the PAN\_1 domain in *CgrAAO* [16,60]. However, after the predicted signal peptide cleavage site, starting



**Fig. 3.** Predicted structure of *BpAlcOx*-CE3 and comparison of its predicted AA5 active site formation with homologs. (A) The overall structure of the *BpAlcOx*-CE3 protein as predicted by ALPHAFOLD [52,53] coloured relative to the pLDDT confidence values in the colour bar. Note that the core of the protein domains is modelled with relatively high confidence but the termini and loops around the AA5 domain are of lower confidence. Putative metal binding residues in the AA5 domain are shown in grey sticks and the copper ion is placed relative to *CgrAAO* ([16], PDB: 6RYV) derived by structural alignment by the DALI server [61]. Residues corresponding to the position of the catalytic triad in the CE3 SGNH-hydrolase domain are also shown as grey sticks. The active sites of (B) the *BpAlcOx* domain as predicted by ALPHAFOLD [52,53], (C) GlxA from *Streptomyces lividans* (PDB: 4UNM), (D) *FgrGalOx* from *Fusarium graminearum* (PDB: 2EIE), and (E) *CgrAAO* from *Colletotrichum graminicola* (PDB: 6RYV). The copper ion shown in *BpAlcOx* was generated by the copper ion observed in *CgrAAO*. The secondary shell aromatic residue stacking on top of the thioether linked tyrosine is highlighted in white text on black background in all structures. Note that in *BpAlcOx* and GlxA this second shell aromatic residue precedes the thioether modified tyrosine in primary sequence whereas in *FgrGalOx* and *CgrAAO* the residue originates from a different loop. The long loop only present in *BpAlcOx* is shown in cyan and the open ridge in *BpAlcOx* created by the absence of a conserved loop amongst AA5\_2 members is highlighted in red.

at residue 35, there is a ~ 19 residue region preceding the  $\beta$ -propeller domain in *BpAlcOx*. The region, comprising many glycine and small side chain residues, is unstructured in all predictions albeit with low confidence, but being located on the opposite side of the AA5 active site it is unlikely to contribute to activity. A search of the PDB with the predicted *BpAlcOx* domain (residues 55–546) using DALI [61] identified *CgrAAO* of the three structurally determined AA5 members as the most closely related ( $C\alpha$  root mean square deviation 2.0 Å; 32% structure-based sequence identity). The overall topology of the *BpAlcOx* AA5 domain is similar to both *CgrAAO* and *FgrGalOx*, with a solvent accessible active site cleft at the surface of the proteins (Fig. 3B–E). The residues coordinating the copper ion in *CgrAAO* and *FgrGalOx* are fully conserved in *BpAlcOx*, as is the cysteine (Cys125 in *BpAlcOx*) shown to form the unique thioether linkage

with the metal coordinating tyrosine residue in CROs (Fig. 3B–E; Fig. S7) [18,62]. Characterized AA5 members contain a secondary shell aromatic residue which stacks on top of the thioether linked tyrosine residue and the identity of the residue is important for defining GalOx activity in certain CROs [6,16,18,63]. Interestingly, *BpAlcOx* has a tryptophan (Trp189) in the equivalent position, but it originates from the same loop as the tyrosine of the thioester linkage in contrast to *CgrAAO*, *CgrAlcOx*, and *FgrGalOx* where it originates from a loop between  $\beta$ -strands of the  $\beta$ -propeller, and a glycine is found in the equivalent position as Trp189 in *BpAlcOx*. This differently originating aromatic residue is also observed in GlxA from *S. lividans*, a putative GalOx that lacks GalOx activity *in vitro* but its deletion, which leads to a loss of glycan accumulation at hyphal tips [34], showcases that substrate specificity determinants are complex within the family and the presence of a tryptophan at this location alone is insufficient to define specificity. Sequence identity between *BpAlcOx* differs compared to studied enzymes beyond the conserved metal binding core, but residues lining the active site typically have functional similarities that likely contribute to similar substrate specificities. A noticeable distinct feature of *BpAlcOx* is the presence of a ~ 23-residue loop inserted between two  $\beta$ -strands of the  $\beta$ -propeller on one side of the



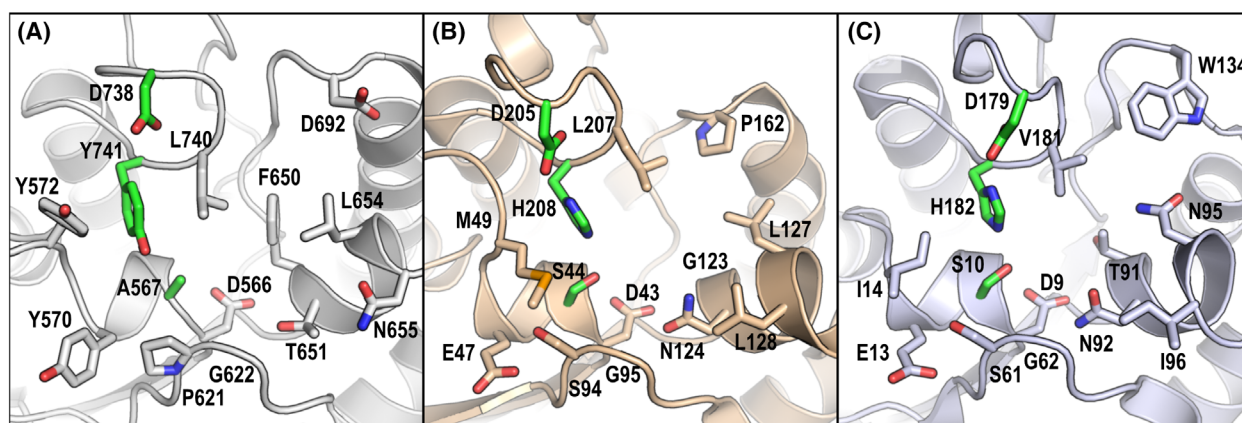
active site, and a shorter loop region between  $\beta$ -strands on the opposite side of the cleft. The shorter loop region leads to a loss of putative substrate-interacting residues found in other enzymes and may create a more open cleft along that face. The position of the long loop region is modelled with low confidence, and while its position relative to the active site remains unknown, it is comprised of hydrophilic residues which could possibly contribute to substrate specificity.

The predicted C-terminal SGNH-hydrolase domain contains a 5-stranded  $\beta$ -sheet packed by helices on both faces similar to structurally determined CE3 enzymes (Fig. 3A). The DALI server [61] was utilized to identify proteins similar in structure to the *BpCE3* and revealed the two structural determined CE3 members, the N-terminal catalytic module of *CtCes3* from *Acetivibrio thermocellus* (formerly *Clostridium thermocellum*) (PDB: 2VPT) [40] and *TcAXE* from *Talaromyces cellulolyticus* (PDB: 5B51, 5B5S) [64], as close homologs ( $C\alpha$  root mean square deviation 1.5–2.0, structure-based sequence identity 27–32%). CE3 enzymes utilize the classical Asp/Glu-His-Ser catalytic triad common amongst SGNH hydrolases which is found in a small solvent exposed cleft. Interestingly, although being very similar in overall structure to homologous enzymes, *BpCE3* lacks both the catalytic histidine and serine, which are instead predicted to be a tyrosine (Tyr741) and alanine (Ala567), respectively (Fig. 4). As the catalytic triad residues are necessary for catalysis [65] we propose that the lack of esterase activity is due to these drastic changes to crucial active site residues. Substitution of the residues to create the

expected catalytic triad by site-directed mutagenesis (A567S and Y741H) did not restore esterase activity towards *p*NP-acetate, a common esterase substrate, indicating that additional molecular determinants facilitating activity are absent in *BpCE3*.

## Discussion

Interest in the utilization of enzymes for green chemistry has been increasing as often enzymes are specific for both substrates and products and can be competitive alternatives for complex and costly chemical reactions. Oxidative reactions are particularly valuable in chemical synthesis processes and investigations into AA5 enzymes have uncovered a wide range of industrially relevant activities [1,4]. *BpAlcOx* is promiscuous for primary alcohol substrates, where its specificity increases with alkyl chain length and its inactivity on secondary alcohols such as isopropanol is similar to many AA5\_2 enzymes. While *BpAlcOx* oxidizes monosaccharides, our results indicate a lack of activity on aldoses in their closed-ring forms, and oxidation only of open-chain configurations. The enzyme was most specific for nonanol, 1,3-propanediol, and furan alcohol, and notably catalysed only alcohol and not aldehyde oxidation. Some members of AA5\_2, such as *CgrAAO*, have been noted to oxidize primary alcohols only to the corresponding aldehydes and not to the fully oxidized carboxylic acid products [16]. Possibly, this limited oxidation by *BpAlcOx* and similar enzymes could be exploited to convert alcohols into aldehydes for functionalization by e.g. imine and oxime click chemistry [66,67].



**Fig. 4.** Comparison of the active site formation of predicted *BpCE3* domain and homologs. The active site of (A) the *BpCE3* domain as predicted by ALPHAFOLD [52,53], (B) the N-terminal catalytic CE3 module of *CtCes3* from *Acetivibrio thermocellus* (PDB: 2VPT), and (C) the CE3 *TcAE206* from *Talaromyces cellulolyticus* (PDB: 5B5S). The catalytic triad in each, and the associated residues predicted in *BpCE3*, are coloured green.

Protein structure models of *BpAlcOx*-CE3 were generated and although the AA5 domain has less than 35% sequence identity to structurally determined AA5 members, the residues supporting metal coordination are conserved, suggesting an appropriate model. Several studies have previously aimed to elucidate the molecular determinants governing substrate specificity in AA5 members [4,68]. However, phylogenetic analyses have revealed a diverse family regarding primary sequences, resulting in enzymes with broad substrate specificities and defined determinants for specificity remain elusive, and consequently also functional predictions. The oxidase domain of *BpAlcOx*-CE3 is in a distinct and unexplored clade within AA5, and the closest studied homolog was not studied as a purified enzyme but through gene deletion and phenotype observation [32]. Similar to the AA5\_2 member *CgrAAO*, *BpAlcOx* has some specificity towards furan compounds. However, the high  $K_m$  values and low turnover rates lead to specificity constants > 75 times lower for *BpAlcOx* than *CgrAAO* suggesting the sequence divergence has led to a distinct function for *BpAlcOx* and members of the clade.

As for all AA5 enzymes, the native substrate(s) of *BpAlcOx*-CE3 and its biological role is unknown, but the signal peptide suggests it being exported out of the cytoplasm. The homologous AA5\_2 GlxA from *S. lividans* is found in membrane fractions and comprises a predicted N-terminal transmembrane domain which likely aids in this association [34]. *BpAlcOx*-CE3 contains a 12-residue N-terminal region between the predicted signal peptide cleavage site and the AA5 domain, and while sequence analysis does not suggest this to be a transmembrane helix, it is a possibility. *BpAlcOx*-CE3 could localize to either the periplasmic space or extracellularly where it would encounter a range of potential substrates to act upon. Possibly, this enzyme could be involved in *B. pseudomallei* infections, similar to the phytopathogenic fungi, *Colletotrichum orbiculare* and *Magnaporthe oryzae*, where deletion of AlcOx-encoding genes resulted in decreased pathogenicity and penetration ability [23]. A role for CROs to produce  $H_2O_2$  as a co-substrate for other enzymes has been postulated for lignin degradation [3] and in fungal morphogenesis [69,70] possibly indicating that the biological role for *BpAlcOx* could be to produce  $H_2O_2$  for some as of yet unknown cellular process. Our large substrate screen has shown many unique features of the enzyme which could help elucidate a preferred substrate(s) *in vivo*. For example, *BpAlcOx* has enhanced activity with ribitol compared to pentanol and 1,5-pentanediol, and saturation kinetics observed with ribitol where the  $K_m$  of

$72.3 \pm 4.43$  mM indicates that the enzyme has specificity for hydroxylated substrates. Similarly, the enzyme prefers 1,3-propanediol over glycerol and considerably more over 1-propanol which again supports the idea of a preference for hydroxylated substrates. The comparatively low  $K_m$  for 1,3-propanediol and glycolaldehyde could possibly help point towards a preferred substrate skeleton for the enzyme. The  $K_m$  was similarly low for furfuryl alcohol, but also considerable substrate inhibition was observed with an inhibitory constant similar to its  $K_m$ . Inhibition was absent for 5-hydroxymethyl-2-furancarboxylic acid indicating that substitutions on the planar compound confer specificity. It is interesting to speculate whether the activities with furan alcohols indicate a preference for furanose compounds. At equilibrium in solution, most monosaccharides tested here exist as furanoses only in small proportions and given the diversity of sugar isomers in solution it is difficult to ascribe the minimal activities of *BpAlcOx* observed here to a defined isomeric form. However, the lack of activity on sucrose could indicate a furanose preference being a faulty hypothesis. In the cellular environment, certain anomeric configurations could however be more common and serve as substrates for *BpAlcOx* and possibly other homologous family members. Gene deletions of the GlxA enzymes from *S. coelicolor* [71] and *S. lividans* [34] lead to a loss of glycan build-up at hyphal tips. A role in glycan modification also for *BpAlcOx*-CE3 could be possible, and potentially pursued with genetic studies, as well as determination of cellular localization and potential enzyme activities with cell wall components. Whether this enzyme has a role in Melioidosis or other pathogenicity is not known, but a deeper understanding of the biochemical properties of *BpAlcOx*-CE3 could lay a foundation for future therapeutic developments.

In summary, our investigation of the unusual AA5-CE3 fusion protein from *B. pseudomallei* revealed a primary alcohol oxidase domain with specificity similar to some AA5\_2 members, and an inactive CE3 domain. *BpAlcOx* represents one of only a few characterized bacterial CROs, and while it was not as active as some fungal counterparts, its biochemical characterization here could still provide useful information in the development of new CRO-based biotechnological applications. The biological role for *BpAlcOx*-CE3 is unclear, and the reason for combining an AlcOx and an esterase domain remains unresolved, and though physically linking CEs has previously been shown to improve the catalytic potential of such multicatalytic enzymes [72], a similar additive effect for *BpAlcOx*-CE3 is unlikely given the inactive CE3 domain.

## Acknowledgements

This work was supported by the Knut and Alice Wallenberg Foundation, *via* the Wallenberg Wood Science Center. We would like to thank Dr Sara Da Costa in the Department of Chemistry and Chemical Engineering at Chalmers University of Technology for supplying some of the substrates used in this study.

## Author contributions

SM and JL conceived the project. SM and AS planned and completed the experimental work, analysed the data, and wrote the manuscript. JL directed the study and critically appraised manuscript drafts. All authors read and approved the final manuscript.

## Peer review

The peer review history for this article is available at <https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/1873-3468.14632>.

## Data accessibility

The data that support the findings of this study are available in the main text and [Supporting Information](#) of this article.

## References

- Hollmann F, Arends IWCE, Buehler K, Schallmey A and Bühler B (2011) Enzyme-mediated oxidations for the chemist. *Green Chem* **13**, 226–265.
- Wahart AJ, Staniland J, Miller GJ and Cosgrove SC (2022) Oxidase enzymes as sustainable oxidation catalysts. *R Soc Open Sci* **9**, 211572.
- Kersten P and Cullen D (2014) Copper radical oxidases and related extracellular oxidoreductases of wood-decay Agaricomycetes. *Fungal Genet Biol* **72**, 124–130.
- Koschorreck K, Alpdagtas S and Urlacher VB (2022) Copper-radical oxidases: a diverse group of biocatalysts with distinct properties and a broad range of biotechnological applications. *Eng Microbiol* **2**, 100037.
- Drula E, Garron ML, Dogan S, Lombard V, Henrissat B and Terrapon N (2022) The carbohydrate-active enzyme database: functions and literature. *Nucleic Acids Res* **50**, D571–D577.
- Daou M and Faulds CB (2017) Glyoxal oxidases: their nature and properties. *World J Microbiol Biotechnol* **33**, 87.
- Kersten PJ and Kirk TK (1987) Involvement of a new enzyme, glyoxal oxidase, in extracellular H<sub>2</sub>O<sub>2</sub> production by *Phanerochaete chrysosporium*. *J Bacteriol* **169**, 2195–2201.
- Daou M, Piumi F, Cullen D, Record E and Faulds CB (2016) Heterologous production and characterization of two glyoxal oxidases from *Pycnoporus cinnabarinus*. *Appl Environ Microbiol* **82**, 4867–4875.
- Whittaker MM, Kersten PJ, Nakamura N, Sanders-Loehr J, Schweizer ES and Whittaker JW (1996) Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. *J Biol Chem* **271**, 681–687.
- Šola K, Gilchrist EJ, Ropartz D, Wang L, Feussner I, Mansfield SD, Ralet M-C and Haughn GW (2019) RUBY, a putative galactose oxidase, influences pectin properties and promotes cell-to-cell adhesion in the seed coat epidermis of *Arabidopsis*. *Plant Cell* **31**, 809–831.
- Šola K, Dean GH, Li Y, Lohmann J, Movahedan M, Gilchrist EJ, Adams KL and Haughn GW (2021) Expression patterns and functional characterization of *Arabidopsis* galactose oxidase-like genes suggest specialized roles for galactose oxidases in plants. *Plant Cell Physiol* **62**, 1927–1943.
- Cooper JA, Smith W, Bacila M and Medina H (1959) Galactose oxidase from *Polyporus circinatus*, Fr. *J Biol Chem* **234**, 445–448.
- Avigad G, Amaral D, Asensio C and Horecker BL (1962) The D-galactose oxidase of *Polyporus circinatus*. *J Biol Chem* **237**, 2736–2743.
- Whittaker MM and Whittaker JW (2001) Catalytic reaction profile for alcohol oxidation by galactose oxidase. *Biochemistry* **40**, 7140–7148.
- Andberg M, Møllerup F, Parikka K, Koutaniemi S, Boer H, Juvonen M, Master E, Tenkanen M and Kruus K (2017) A novel *Colletotrichum graminicola* raffinose oxidase in the AA5 family. *Appl Environ Microbiol* **83**, e01383-17.
- Mathieu Y, Offen WA, Forget SM, Ciano L, Viborg AH, Blagova E, Henrissat B, Walton PH, Davies GJ and Brumer H (2020) Discovery of a fungal copper radical oxidase with high catalytic efficiency toward 5-hydroxymethylfurfural and benzyl alcohols for bioprocessing. *ACS Catal* **10**, 3042–3058.
- Cleveland M, Lafond M, Xia FR, Chung R, Mulyk P, Hein JE and Brumer H (2021) Two *Fusarium* copper radical oxidases with high activity on aryl alcohols. *Biotechnol Biofuels* **14**, 138.
- Yin DT, Urresti S, Lafond M, Johnston EM, Derikvand F, Ciano L, Berrin JG, Henrissat B, Walton PH, Davies GJ *et al.* (2015) Structure-function characterization reveals new catalytic diversity in the galactose oxidase and glyoxal oxidase family. *Nat Commun* **6**, 10197.
- Oide S, Tanaka Y, Watanabe A and Inui M (2019) Carbohydrate-binding property of a cell wall integrity and stress response component (WSC) domain of an alcohol oxidase from the rice blast pathogen *Pyricularia oryzae*. *Enzyme Microb Technol* **125**, 13–20.

- 20 Ribeaucourt D, Saker S, Navarro D, Bissaro B, Drula E, Correia LO, Haon M, Grisel S, Lapalu N, Henrissat B *et al.* (2021) Identification of copper-containing oxidoreductases in the secretomes of three *Colletotrichum* species with a focus on copper radical oxidases for the biocatalytic production of fatty aldehydes. *Appl Environ Microbiol* **87**, e0152621.
- 21 Cleveland ME, Mathieu Y, Ribeaucourt D, Haon M, Mulyk P, Hein JE, Lafond M, Berrin JG and Brumer H (2021) A survey of substrate specificity among auxiliary activity family 5 copper radical oxidases. *Cell Mol Life Sci* **78**, 8187–8208.
- 22 Ribeaucourt D, Bissaro B, Guallar V, Yemloul M, Haon M, Grisel S, Alphand V, Brumer H, Lambert F, Berrin J-G *et al.* (2021) Comprehensive insights into the production of long chain aliphatic aldehydes using a copper-radical alcohol oxidase as biocatalyst. *ACS Sustain Chem Eng* **9**, 4411–4421.
- 23 Bissaro B, Kodama S, Nishiuchi T, Díaz-Rovira AM, Hage H, Ribeaucourt D, Haon M, Grisel S, Simaan AJ, Beisson F *et al.* (2022) Tandem metalloenzymes gate plant cell entry by pathogenic fungi. *Sci Adv* **8**, eade9982.
- 24 Whittaker JW (2003) Free radical catalysis by galactose oxidase. *Chem Rev* **103**, 2347–2364.
- 25 Ramya TN, Weerapana E, Cravatt BF and Paulson JC (2013) Glycoproteomics enabled by tagging sialic acid- or galactose-terminated glycans. *Glycobiology* **23**, 211–221.
- 26 Roberts GP and Gupta SK (1965) Use of galactose oxidase in the histochemical examination of mucus-secreting cells. *Nature* **207**, 425–426.
- 27 Monosik R, Stredansky M, Tkac J and Sturdik E (2012) Application of enzyme biosensors in analysis of food and beverages. *Food Anal Methods* **5**, 40–53.
- 28 Schoevaart R and Kieboom T (2004) Application of galactose oxidase in chemoenzymatic one-pot cascade reactions without intermediate recovery steps. *Top Catal* **27**, 3–9.
- 29 Mikkonen KS, Parikka K, Suuronen J-P, Ghafar A, Serimaa R and Tenkanen M (2014) Enzymatic oxidation as a potential new route to produce polysaccharide aerogels. *RSC Adv* **4**, 11884–11892.
- 30 Leppänen AS, Xu C, Parikka K, Eklund P, Sjöholm R, Brumer H, Tenkanen M and Willför S (2014) Targeted allylation and propargylation of galactose-containing polysaccharides in water. *Carbohydr Polym* **100**, 46–54.
- 31 Huffman MA, Fryszkowska A, Alvizo O, Borra-Garske M, Campos KR, Canada KA, Devine PN, Duan D, Forstater JH, Grosser ST *et al.* (2019) Design of an in vitro biocatalytic cascade for the manufacture of islatravir. *Science* **366**, 1255–1259.
- 32 Silakowski B, Ehret H and Schairer HU (1998) fbfB, a gene encoding a putative galactose oxidase, is involved in *Stigmatella aurantiaca* fruiting body formation. *J Bacteriol* **180**, 1241–1247.
- 33 Whittaker MM and Whittaker JW (2006) *Streptomyces coelicolor* oxidase (SCO2837p): a new free radical metalloenzyme secreted by *Streptomyces coelicolor* A3 (2). *Arch Biochem Biophys* **452**, 108–118.
- 34 Chaplin AK, Petrus MLC, Mangiameli G, Hough MA, Svistunenko DA, Nicholls P, Claessen D, Vijgenboom E and Worrall JAR (2015) GlxA is a new structural member of the radical copper oxidase family and is required for glycan deposition at hyphal tips and morphogenesis of *Streptomyces lividans*. *Biochem J* **469**, 433–444.
- 35 Schweizer HP (2012) Mechanisms of antibiotic resistance in *Burkholderia pseudomallei*: implications for treatment of melioidosis. *Future Microbiol* **7**, 1389–1399.
- 36 Wiersinga WJ, Virk HS, Torres AG, Currie BJ, Peacock SJ, Dance DAB and Limmathurotsakul D (2018) Melioidosis. *Nat Rev Dis Primers* **4**, 17107.
- 37 Bzdyl NM, Moran CL, Bendo J and Sarkar-Tyson M (2022) Pathogenicity and virulence of *Burkholderia pseudomallei*. *Virulence* **13**, 2139063.
- 38 Dalrymple BP, Cybinski DH, Layton I, McSweeney CS, Xue GP, Swadling YJ and Lowry JB (1997) Three *Neocallimastix patriciarum* esterases associated with the degradation of complex polysaccharides are members of a new family of hydrolases. *Microbiology (Reading)* **143** (Pt 8), 2605–2614.
- 39 Aurilia V, Martin JC, McCrae SI, Scott KP, Rincon MT and Flint HJ (2000) Three multidomain esterases from the cellulolytic rumen anaerobe *Ruminococcus flavefaciens* 17 that carry divergent dockerin sequences. *Microbiology (Reading)* **146** (Pt 6), 1391–1397.
- 40 Correia MA, Prates JA, Brás J, Fontes CM, Newman JA, Lewis RJ, Gilbert HJ and Flint JE (2008) Crystal structure of a cellulosomal family 3 carbohydrate esterase from *Clostridium thermocellum* provides insights into the mechanism of substrate recognition. *J Mol Biol* **379**, 64–72.
- 41 Madeira F, Pearce M, Tivey ARN, Basutkar P, Lee J, Edbali O, Madhusoodanan N, Kolesnikov A and Lopez R (2022) Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic Acids Research* **50**, W276–W279.
- 42 Trifinopoulos J, Nguyen LT, von Haeseler A and Minh BQ (2016) W-IQ-TREE: a fast online phylogenetic tool for maximum likelihood analysis. *Nucleic Acids Res* **44**, W232–W235.
- 43 Letunic I and Bork P (2021) Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* **49**, W293–w296.
- 44 Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G and Nielsen H (2019) SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nat Biotechnol* **37**, 420–423.



- 45 Liu H and Naismith JH (2008) An efficient one-step site-directed deletion, insertion, single and multiple-site plasmid mutagenesis protocol. *BMC Biotechnol* **8**, 91.
- 46 Ellis KJ and Morrison JF (1982) Buffers of constant ionic strength for studying pH-dependent processes. *Methods Enzymol* **87**, 405–426.
- 47 Deacon SE, Mahmoud K, Spooner RK, Firbank SJ, Knowles PF, Phillips SE and McPherson MJ (2004) Enhanced fructose oxidase activity in a galactose oxidase variant. *Chembiochem* **5**, 972–979.
- 48 Drew KN, Zajicek J, Bondo G, Bose B and Serianni AS (1998) <sup>13</sup>C-labeled aldopentoses: detection and quantitation of cyclic and acyclic forms by heteronuclear 1D and 2D NMR spectroscopy. *Carbohydr Res* **307**, 199–209.
- 49 Zhu Y, Zajicek J and Serianni AS (2001) Acyclic forms of [1-<sup>13</sup>C]aldohexoses in aqueous solution: quantitation by <sup>13</sup>C NMR and deuterium isotope effects on tautomeric equilibria. *J Org Chem* **66**, 6244–6251.
- 50 Zhao R, Lee AKY, Soong R, Simpson AJ and Abbott JPD (2013) Formation of aqueous-phase  $\alpha$ -hydroxyhydroperoxides ( $\alpha$ -HHP): potential atmospheric impacts. *Atmos Chem Phys* **13**, 5857–5872.
- 51 Kua J, Galloway MM, Millage KD, Avila JE and De Haan DO (2013) Glycolaldehyde monomer and oligomer equilibria in aqueous solution: comparing computational chemistry and NMR data. *J Phys Chem A* **117**, 2997–3008.
- 52 Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589.
- 53 Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A *et al.* (2022) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* **50**, D439–D444.
- 54 Lin Z, Akin H, Rao R, Hie B, Zhu Z, Lu W, Smetanin N, Verkuil R, Kabeli O, Shmueli Y *et al.* (2022) Evolutionary-scale prediction of atomic level protein structure with a language model. *bioRxiv*. doi: [10.1101/2022.07.20.500902](https://doi.org/10.1101/2022.07.20.500902)
- 55 Yang J, Anishchenko I, Park H, Peng Z, Ovchinnikov S and Baker D (2020) Improved protein structure prediction using predicted interresidue orientations. *Proc Natl Acad Sci USA* **117**, 1496–1503.
- 56 Su H, Wang W, Du Z, Peng Z, Gao SH, Cheng MM and Yang J (2021) Improved protein structure prediction using a new multi-scale network and homologous templates. *Adv Sci (Weinh)* **8**, e2102592.
- 57 Du Z, Su H, Wang W, Ye L, Wei H, Peng Z, Anishchenko I, Baker D and Yang J (2021) The trRosetta server for fast and accurate protein structure prediction. *Nat Protoc* **16**, 5634–5651.
- 58 Wang W, Peng Z and Yang J (2022) Single-sequence protein structure prediction using supervised transformer protein language models. *Nat Comput Sci* **2**, 804–814.
- 59 Abbott DW, Eirín-López JM and Boraston AB (2008) Insight into ligand diversity and novel biological roles for family 32 carbohydrate-binding modules. *Mol Biol Evol* **25**, 155–167.
- 60 Tordai H, Bánai L and Patthy L (1999) The PAN module: the N-terminal domains of plasminogen and hepatocyte growth factor are homologous with the apple domains of the prekallikrein family and with a novel domain found in numerous nematode proteins. *FEBS Lett* **461**, 63–67.
- 61 Holm L (2022) Dali server: structural unification of protein families. *Nucleic Acids Res* **50**, W210–W215.
- 62 Ito N, Phillips SE, Stevens C, Ogel ZB, McPherson MJ, Keen JN, Yadav KD and Knowles PF (1991) Novel thioether bond revealed by a 1.7 Å crystal structure of galactose oxidase. *Nature* **350**, 87–90.
- 63 Rogers MS, Tyler EM, Akyumani N, Kurtis CR, Spooner RK, Deacon SE, Tamber S, Firbank SJ, Mahmoud K, Knowles PF *et al.* (2007) The stacking tryptophan of galactose oxidase: a second-coordination sphere residue that has profound effects on tyrosyl radical behavior and enzyme catalysis. *Biochemistry* **46**, 4606–4618.
- 64 Watanabe M, Fukada H, Inoue H and Ishikawa K (2015) Crystal structure of an acetyltransferase from *Talaromyces cellulolyticus* and the importance of a disulfide bond near the active site. *FEBS Lett* **589**, 1200–1206.
- 65 Mølgaard A, Kauppinen S and Larsen S (2000) Rhamnogalacturonan acetyltransferase elucidates the structure and function of a new family of hydrolases. *Structure* **8**, 373–383.
- 66 Collins J, Xiao Z, Müllner M and Connal LA (2016) The emergence of oxime click chemistry and its utility in polymer science. *Polym Chem* **7**, 3812–3826.
- 67 Cosgrove SC, Matthey AP, Riese M, Chapman MR, Birmingham WR, Blacker AJ, Kapur N, Turner NJ and Flitsch SL (2019) Biocatalytic oxidation in continuous flow for the generation of carbohydrate dialdehydes. *ACS Catal* **9**, 11658–11662.
- 68 Fong JK and Brumer H (2022) Copper radical oxidases: galactose oxidase, glyoxal oxidase, and beyond! *Essays Biochem* **67**, 597–613.
- 69 Leuthner B, Aichinger C, Oehmen E, Koopmann E, Müller O, Müller P, Kahmann R, Böcker M and Schreier PH (2005) A H<sub>2</sub>O<sub>2</sub>-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis*. *Mol Genet Genomics* **272**, 639–650.

- 70 Crutcher FK, Moran-Diez ME, Krieger IV and Kenerley CM (2019) Effects on hyphal morphology and development by the putative copper radical oxidase *glx1* in *Trichoderma virens* suggest a novel role as a cell wall associated enzyme. *Fungal Genet Biol* **131**, 103245.
- 71 Liman R, Facey PD, van Keulen G, Dyson PJ and Del Sol R (2013) A laterally acquired galactose oxidase-like gene is required for aerial development during osmotic stress in *Streptomyces coelicolor*. *PLoS One* **8**, e54112.
- 72 Kmezik C, Bonzom C, Olsson L, Mazurkewich S and Larsbrink J (2020) Multimodular fused acetyl-feruloyl esterases from soil and gut Bacteroidetes improve xylanase depolymerization of recalcitrant biomass. *Biotechnol Biofuels* **13**, 60.

## Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**Fig. S1.** Signal peptide prediction by SignalP 5.0 of *BpAlcOx-CE3* found in NCBI and UniProt.

**Fig. S2.** Phylogenetic tree of AA5 with bootstrap values.

**Fig. S3.** Phylogenetic tree of CE3.

**Fig. S4.** Dependence of pH on the *BpAlcOx* activity with 1,3-propanediol.

**Fig. S5.** Plots of *BpAA5* activity with substrates for which saturation kinetics were observed.

**Fig. S6.** Comparison of *BpAlcOx-CE3* models generated by structure predictions.

**Fig. S7.** Multiple sequence alignment of select AA5 members.

**Table S1.** Sequence of codon-optimized *BpAlcOx-CE3* construct utilized.

**Table S2.** Primer sequences utilized for cloning and mutagenesis.

**Table S3.** Kinetic parameters for all the compounds tested.

**Table S4.** Equilibria of some of the monosaccharides utilized in this study.