Discovery of a Fungal Copper Radical Oxidase with High Catalytic Efficiency Towards 5-Hydroxymethylfurfural and Benzyl Alcohols for Bioprocessing

Yann Mathieu[1], Wendy A. Offen[2], Stephanie M. Forget[1,3], Luisa Ciano[3,+], Alexander Holm Viborg[1], Elena Blagova[3], Bernard Henrissat[4,5], Paul H. Walton[3], Gideon J. Davies[3], Harry Brumer[1,2,6,7,]*

[1]Michael Smith Laboratories, University of British Columbia, 2185 East Mall, Vancouver, BC, V6T 1Z4, Canada;

[2]Department of Chemistry, University of York, Heslington, YO10 5DD, York, UK.

[3]Department of Chemistry, University of British Columbia, 2036 Main Mall, Vancouver, BC, V6T 1Z1, Canada;

[4]Architecture et Fonction des Macromolécules Biologiques (AFMB), CNRS, Aix-Marseille University, Marseille, 13288, France;

[5]INRA, USC1408 Architecture et Fonction des Macromolécules Biologiques (AFMB), Marseille, 13288, France;

[6]Department of Biochemistry and Molecular Biology, University of British Columbia, 2350 Health Sciences Mall, Vancouver, BC, V6T 1Z3, Canada;

[7]Department of Botany, University of British Columbia, 3200 University Boulevard, Vancouver, BC, V6T 1Z4, Canada

[+] Current address: School of Chemistry and Photon Science Institute, University of Manchester, Oxford Road, Manchester, M13 9PL, UK

ABSTRACT Alternatives to petroleum-based chemicals are highly sought-after for on-going efforts to reduce the damaging effects of human activity on the environment. Copper radical oxidases from Auxiliary Activity Family 5/Subfamily 2 (AA5_2) are attractive biocatalysts because they oxidize primary alcohols in a chemo-selective manner without complex organic cofactors. However, despite numerous studies on canonical galactose oxidases (GalOx, EC 1.1.3.9) and engineered variants, and the recent discovery of a *Colletotrichum graminicola* copper radical alcohol oxidase (AlcOx, EC 1.1.3.13), the catalytic potentials of very few AA5_2 members have been characterized. Guided by sequence similarity network and phylogenetic analyses, in this study we targeted a distinct paralog from the fungus *C. graminicola* as a representative member of a large uncharacterized subgroup of AA5_2. Through recombinant production and detailed kinetic analysis, we demonstrated that this enzyme is weakly active towards carbohydrates, but efficiently catalyzes the oxidation of aryl alcohols to the corresponding aldehydes. As such, this represents the initial characterization of a demonstrable aryl alcohol oxidase (AAO, EC 1.1.3.7) in AA5, an activity which is classically associated with flavin-dependent glucose-methanol-choline (GMC)

oxidoreductases of Auxiliary Activity Family 3 (AA3). X-ray crystallography revealed a distinct multidomain architecture comprising an N-terminal PAN domain abutting a canonical AA5 seven-bladed propeller catalytic domain. Of direct relevance to biomass processing, the wild-type enzyme exhibits the highest activity on the primary alcohol of 5-hydroxymethylfurfural (HMF), a product of significant interest in the lignocellulosic bio-refinery concept. Thus, the chemoselective oxidation of HMF to 2,5-diformylfuran (DFF) by *C. graminicola* aryl alcohol oxidase (*Cgr*AAO) from AA5 provides a fundamental building block for chemistry via biotechnology.

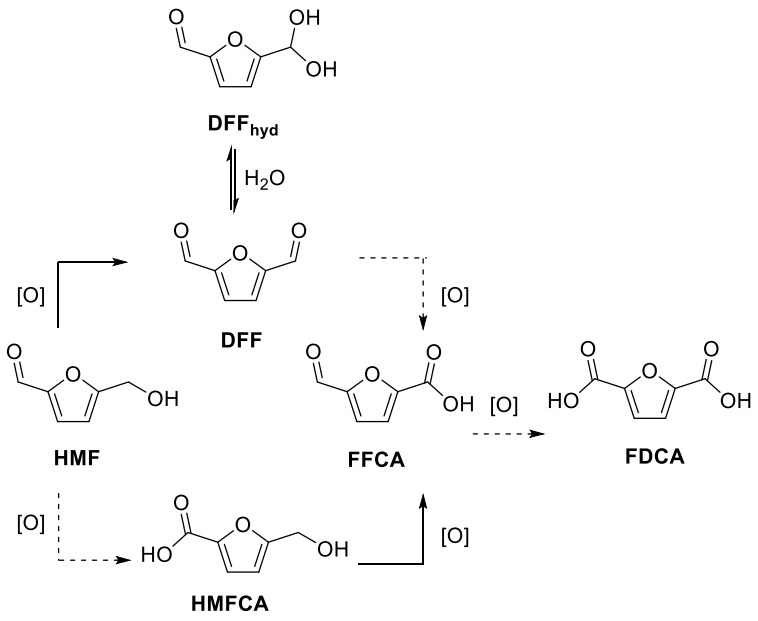
KEYWORDS: oxidoreductases, enzyme kinetics, structural biology, biocatalysis, bioproducts, EC 1.1.3.7, EC 1.1.3.47

INTRODUCTION

The depletion of crude oil reserves reflected in rising prices, coupled with growing concern about environmental impacts, emphasizes the need for the development of sustainable alternatives to petroleum-based fuels and chemicals[1]. Lignocellulosic biomass has received sustained attention as an alternative to fossil petroleum because it constitutes a large renewable resource that does not compete with food crops (*i.e.* starchy plants) as a biomass source[2]. However, the complex chemical and structural composition of lignocellulosic biomass poses a significant challenge in processing the raw material, which is primarily responsible for the high cost of lignocellulose conversion[3].

Among the valuable chemicals that can be obtained from lignocellulosic biomass, 5-hydroxymethylfurfural (HMF), which is derived from the depolymerization and dehydration of cellulose, is viewed as a versatile building block, particularly for the polymer industry, due to its rich chemistry (Scheme 1)[4, 5]. For example, oxidation of the primary alcohol generates 2,5-

diformylfuran (DFF), which has been used in the synthesis of renewable furan-urea polymeric resins[6], antifungal agents[7], pharmaceutical compounds[8] and electroconductors[9]. Further oxidation of DFF generates 2,5-formylfurancarboxylic acid (FFCA), which is an intermediate for the preparation of surfactants, biofuels, resins and other compounds[10] and 2,5-furandicarboxylic acid (FDCA), which has received significant attention as an alternative to fossil-fuel based terephthalic acid for the production of commodity plastics[11]. Likewise, 5-hydroxymethyl-2-furancarboxylic acid (HMFCA) has applications in polymers[12] and therapeutics[13]. Hence, the discovery and application of (bio)catalysts that perform chemoselective redox transformations of HMF to access these renewable building blocks remains an area of significant contemporary interest[14-21].



Scheme 1. Oxidation products of HMF*.

*Reactions catalyzed by *Cgr*AAO are shown using solid arrows.

In nature, to enable access to nutrients, a cache of enzymes which target the different components and linkages of lignocellulosic biomass is produced by microorganisms, especially saprotrophs and phytopathogens[22, 23]. Motivated by industrial interest in monosaccharide production from lignocellulose for fermentation to fuels and chemicals (*e.g.* "bioethanol"), there has been longstanding interest in the characterization and application of diverse polysaccharide hydrolases, especially cellulases[24-27]. More recently, microbial oxidoreductases have gained increased attention for biomass valorization, not only due to their ability to enhance cellulose saccharification, but also for their ability to functionalize polysaccharides, lignin, and small organic molecules[28-33].

The Carbohydrate-Active enZYmes database (CAZy, www.cazy.org) was recently expanded to integrate oxidoreductases associated with biological lignocellulose modification, resulting in the 16 current Auxiliary Activity (AA) families[34]. Among these, AA5 comprises mononuclear Copper Radical Oxidases (CROs), which are further divided into two subfamilies based on protein phylogeny. Subfamily 1 (AA5_1) currently includes two characterized (methyl)glyoxal oxidases (EC 1.2.3.15), which carry out the two electron oxidation of these aldehydes (likely via their hydrates) to carboxylic acids with concomitant reduction of O_2 to H_2O_2[35]. Subfamily 2 (AA5_2) comprises canonical galactose 6-oxidases (GalOx, EC 1.1.3.9) and recently discovered general alcohol oxidases (AlcOx, EC 1.1.3.13), which perform analogously the oxidation of primary alcohols into the corresponding aldehydes[36, 37]. In accordance with their classification in a single CAZyme family, all AA5 members are anticipated to share a conserved tertiary structure, active-site, and overall catalytic mechanism[34, 38].

Many early and contemporary studies of fungal AA5_2 members found these enzymes to be highly specific for the 6-hydroxyl group of D-galactose and terminal D-galactosides in di-, oligo-, and polysaccharides[39-47]. Contrary to this paradigm, we recently showed that two AA5_2 orthologs

from the phytopathogenic fungi *Colletotrichum graminicola* and *Colletotrichum gloeosporioides* have predominant activity on primary aliphatic alcohols and essentially no activity on galactosyl substrates[36]. Encouraged by this observation of a broader, untapped enzymatic diversity within AA5_2 enzymes, in the present study we have used Sequence Similarity Network and Maximum Likelihood phylogenetic analyses to guide the selection of a paralog from *C. graminicola* as a representative member of a large uncharacterized subgroup. Through recombinant expression, enzymology and structural biology, we demonstrate that this enzyme is poorly active towards carbohydrates, but has a high aryl-alcohol oxidase (AAO) activity (EC 1.1.3.7)[48] and HMF oxidase activity (EC 1.1.3.47) [14] that is representative of a new specificity class among characterized AA5_2 members.

RESULTS

*Cgr*AAO is the first representative of a previously uncharacterized group within AA5_2

A manually curated multiple protein sequence alignment of nearly 400 AA5_2 catalytic modules (*i.e.*, with signal peptides and additional domains removed) was used to generate a Maximum Likelihood (ML) phylogenetic tree and a sequence similarity network (SSN) to visualize relationships within this subfamily (Figure 1). Using selected AA5_1 glyoxal oxidases[49, 50] as an outgroup, 27 different subgroups were identified based on tree topology and bootstrap values >75 (100 ML replicates). The corresponding SSN at an alignment score (bitscore) cut-off of 550 was concordant with this phylogeny, and was used to map biochemical data and predicted protein modularity, across the AA5_2 landscape (Figure 1 and Figure S1). Characterized AA5_2 (*Pru*AA5_2A)[44], galactose 6-oxidases (GalOx)[39-42, 45, 47], alcohol oxidases (AlcOx)[36, 37], and

raffinose-specific galactose 6-oxidase (RafOx)[43] segregated unambiguously into distinct subgroups, indicating that analysis of subgroup membership is valuable to guide target selection.

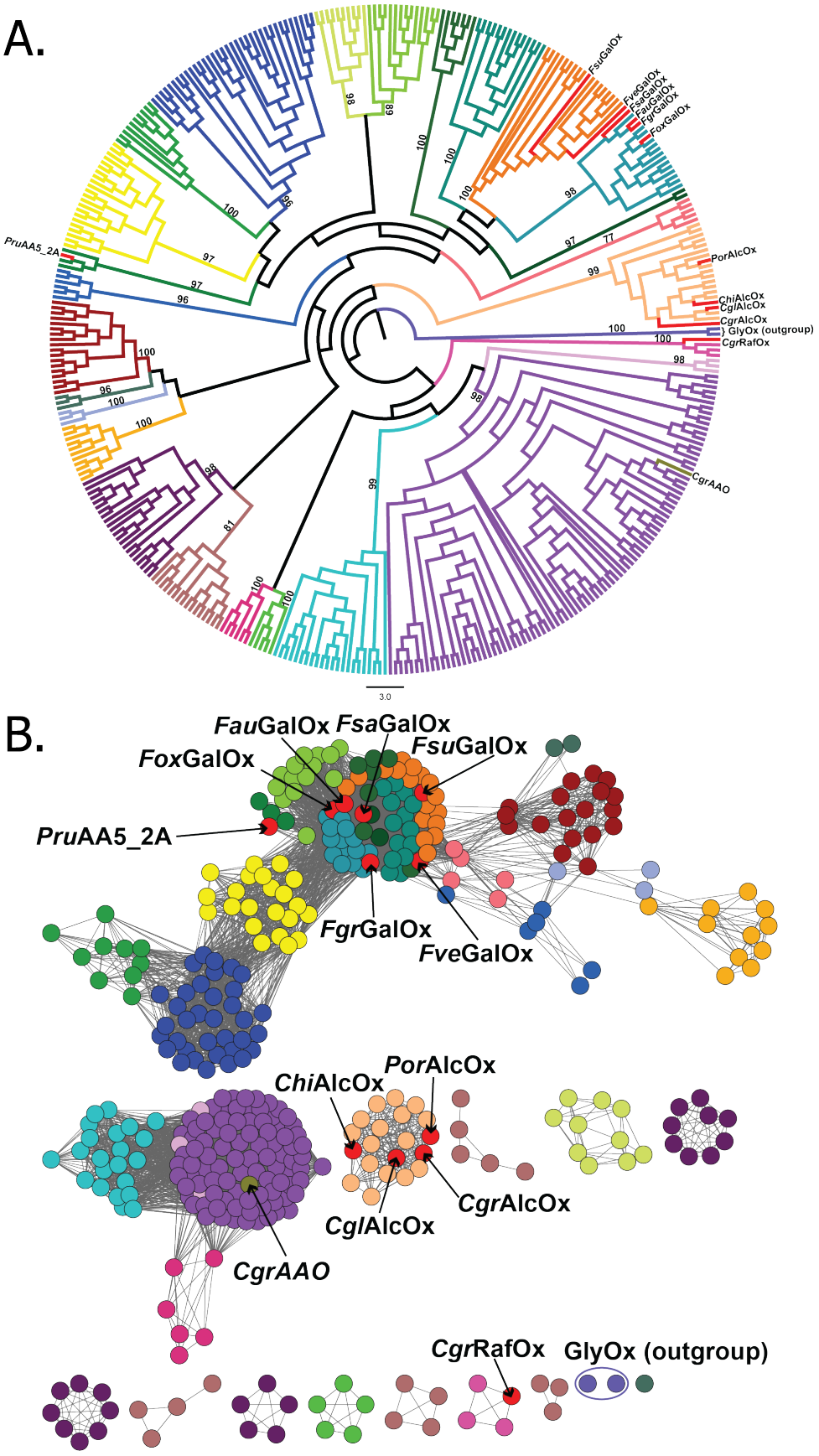

A.
FsuGalOx
FveGalOx
FsaGalOx
FauGalOx
FgrGalOx
FoxGalOx
PruAA5_2A
PorAlcOx
ChiAlcOx
CglAlcOx
CgrAlcOx
GlyOx (outgroup)
CgrRafOx
CgrAAO
3.0
B.
FauGalOx
FsaGalOx
FsuGalOx
FoxGalOx
PruAA5_2A
FgrGalOx
FveGalOx
PorAlcOx
ChiAlcOx
CgrAlcOx
CglAlcOx
CgrAAO
CgrRafOx
GlyOx (outgroup)

Figure 1: Sequence relationships of 392 AA5_2 catalytic modules. (A) Phylogenetic tree with Bootstrap values >75 supporting the 27 subgroups are indicated at each node/branch. (B) Sequence similarity network (SSN). Each node corresponds to one of the 392 catalytic modules used as an input to build the SSN created in Cytoscape with yFiles Organic layout[51]. Edges represent an alignment bitscore threshold of 550 that clusters the sequences into groups and resolves the same monophyletic groups as those observed in (A). For each panel, AA5 members whose biochemical characterization is available are colored in red and indicated as (methyl)glyoxal oxidases (GlyOx)[49, 50] (comprising the outgroup), galactose oxidases (GalOx)[39-42, 45, 47], alcohol oxidases (AlcOx)[36, 37], raffinose oxidase (RafOx)[43], and *Penicillium rubens* Wisconsin 54–1255 AA5_2 oxidase (*Pru*AA5_2A)[44]. The sequence characterized in this study is colored in ochre and indicated as *Cgr*AAO.

The genome of the phytopathogen *Colletotrichum graminicola* M1.001 (Maize anthracnose fungus) encodes three AA5_2 homologs, in addition to one AA5_1 ortholog[52] (see http://www.cazy.org/AA5_eukaryota.html). The loci GLRG_05590 (GenBank EFQ30446.1) and GLRG_11847 (GenBank EFQ36699.1) have been shown previously to encode a general AlcOx[36] and a raffinose/galactose oxidase[43], respectively. The heretofore uncharacterized third *C. graminicola* AA5_2 paralog, encoded by locus GLRG_02805 (GenBank EFQ27661.1) and referred to here as *Cgr*AAO, after detailed biochemical characterization (*vide infra*), is found within a large subgroup comprising 91 sequences without any characterized members (Figure 1). This subgroup was part of a bigger cluster, comprised of four subgroups containing 122 sequences in total, which did not exhibit connections with other clusters in the SSN, possibly indicative of unique enzymatic activity. Indeed, a protein sequence alignment of the catalytic modules of

*Cgr*AAO with characterized AA5_2 members revealed that they share an overall sequence identity slightly lower than 50%, specifically 49% with *Fgr*GalOx[42], 47% with *Cgr*AlcOx[36], 45% with *Cgl*AlcOx[36], 48% with *Cgr*RafOx[43], 42% with *Pru*AA5_2[44], 45% with *Por*AlcOx[37] and 46% with *Chi*AlcOx[37], with notable differences in key active-site residues (Figure S1A, green highlighting).

The modularity of the full-length sequences was mapped onto the SSN that was generated using only the catalytic modules, revealing that modularity correlated strongly with cluster identity (Figure S1B *cf.* Figure 1). Whereas canonical GalOxs contain carbohydrate-binding module family 32 (CBM32) members consistent with galactose recognition[53], members of the AlcOx group lack CBM32 modules altogether. Within this group, full-length sequences contained only the catalytic modules (*i.e.* AlcOx), while many other sequences comprised catalytic modules with one or two Wall Stress-responsive Component (WSC) domains,[54] which have recently been implicated in carbohydrate binding[37]. In contrast, all members of the primary group containing *Cgr*AAO possess an N-terminal PAN_1 domain[55] as do sequences within the small RafOx cluster. Specifically, the 711 amino-acid full-length sequence of *Cgr*AAO comprises a predicted 22 residue signal peptide, a 198 residue PAN_1 domain, and a 491 residue AA5_2 catalytic module.

*Cgr*AAO has a distinct substrate specificity profile

Full-length *Cgr*AAO-WT was produced in *Pichia pastoris* using conditions optimized for the production of other AA5_2 members[36, 56, 57], which yielded 17 mg from 400 mL of buffered complex methanol medium (BMMY) after purification (>95% pure according to SDS-PAGE analysis, Figure S2). Inductively coupled plasma mass spectrometry (ICP-MS) analysis indicated a copper-to-protein ratio of 1.08 ± 0.03, consistent with a fully-loaded mononuclear copper site. The protein was predicted to be N-glycosylated on sites N309, N386 and N644 using the NetNglyc

server (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0). N-glycosylation was confirmed by treatment with PNGaseF or EndoH, and a crystal structure of the *apo* enzyme (PDB ID 6STX, *vide infra*) suggested possible glycosylation on N309, although the corresponding electron density was too low to model reliably. The deglycosylated protein had an electrophoretic mobility consistent with its predicted molecular mass of 77507 Da (Figure S2).

Initial screening using a diverse set of alcohols representing known substrates of AA5 enzymes revealed that *Cgr*AAO was active on galactose and galactosylated oligosaccharides, short chain alkane diols and glycerol, and aryl alcohols with similar specific activity values (Figure 2 and Table S1).

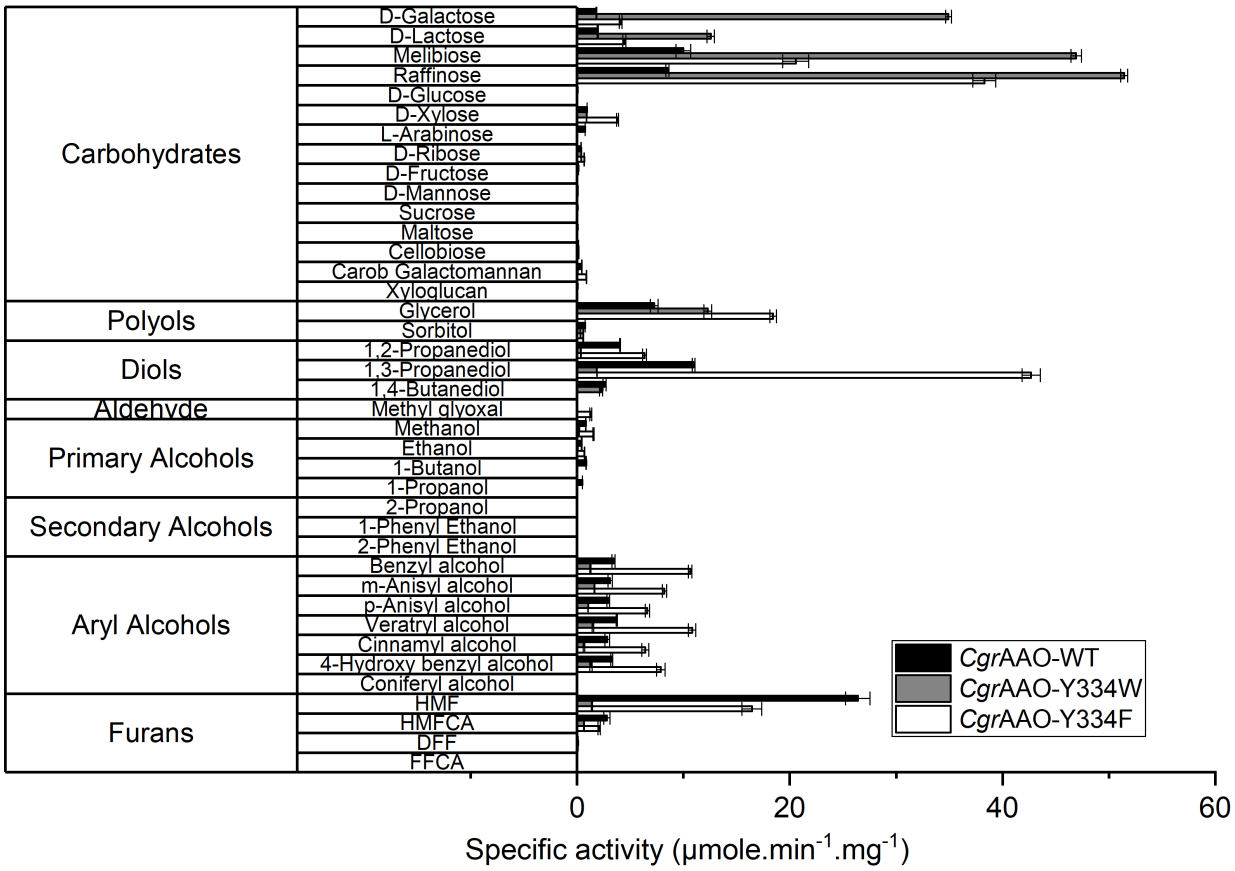


Figure 2: Initial activity screens of *Cgr*AAO-WT and mutants. Measurements were performed in triplicate at 25 °C in 100 mM sodium phosphate buffer pH 7 using the HRP/ABTS assay. Activities were monitored using 300 mM for carbohydrates, polyols, diols and primary alcohols, 2.5 mg.mL^{-1} for galactose containing polysaccharides, 5 mM for methyl glyoxal, aryl alcohols and furans and 10 mM for secondary alcohols. Reactions were started with the addition of 1.3 to 65 μmole of purified enzyme.

This substrate range is reminiscent of both canonical GalOx and AlcOx of AA5[39, 40, 43, 44, 46], although very poor activity on primary alkanols distinguishes *Cgr*AAO from its paralog *Cgr*AlcOx[36]. Striking, however, was the observation of the highest specific activity on HMF (Figure 2 and Table S1, *cf*. Scheme 1), which motivated us to perform detailed kinetic analyses to accurately quantify specificity for the top substrates.

HMF was used to determine the pH-rate profile and temperature stability of *Cgr*AAO because it had the highest specific activity of all substrates. The wild-type enzyme exhibited an optimal activity at pH 7.0 with a bell-shaped profile, similar to *Cgr*AlcOx[36] and *Fgr*GalOx[58]. A rapid drop in activity was observed at acidic pH values, with 50% of activity remaining at pH 5.5 versus pH 7.0, while 50% activity was retained up to pH 10 (Figure S3). After an initial rapid loss of ca. 10% activity at 25 °C, the enzyme was essentially stable over a 22 hour-long experiment. Much greater, rapid activity losses were observed above this temperature (Figure S4), hence all subsequent kinetics analyses were performed at pH 7.0 and 25 °C.

Superseding the initial specific activity data, Michaelis-Menten kinetic analyses revealed that galactose and the terminal-galactose-containing saccharides lactose (Galβ(1,4)-Glc), melibiose (Galα(1,6)-Glc), and raffinose (Galα(1,6')-sucrose) were in fact poor substrates for *Cgr*AAO, with

k_{cat}/K_m values in the range 10 - 70 $M^{-1}s^{-1}$, as a result of very high K_m values (0.2 – 1.0 M, Table 1). Indeed, D-galactose did not show saturation behavior even at substrate concentrations up to 2 M (Figure S5), which prevented the determination of individual kinetic parameters. Nonetheless, we were able to detect conversion of raffinose to the aldehyde, corresponding hydrate, and eventually the acid, at high enzyme loading and extended incubation by MALDI-TOF-MS (Figure S6). The enzyme was likewise poorly active on glycerol, 1,2-propanediol, and 1,3-propanediol (Table 1).

Table 1. Substrate specificity of *Cgr*AAO-WT and its variants*

Substrate	*Cgr*AAO-WT			*Cgr*AAO-Y334F		
	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}.s^{-1}$)	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m ($M^{-1}.s^{-1}$)
Carbohydrates						
Galactose	n.d.§	n.d.§	13.1 ± 0.8	n.d.§	n.d.§	16 ± 1
Lactose	999 ± 83	23.4 ± 2.2	23.5 ± 1.2	575 ± 43	18.7 ± 0.7	32.5 ± 2.7
Melibiose	314 ± 15	12.3 ± 0.2	39.2 ± 1.7	n.d.§	n.d.§	67 ± 4
Raffinose	291 ± 21	20.8 ± 0.6	71.5 ± 5.6	523 ± 35	91 ± 3	174 ± 13
Polyols						
Glycerol	605 ± 34	58.7 ± 0.9	97 ± 5.6	1480 ± 60	104 ± 2	70.3 ± 3.2
Diols						
1,2 Propanediol	1047 ± 64	17.4 ± 0.4	16.6 ± 1.1	n.d.§	n.d.§	87 ± 3
1,3 Propanediol	384 ± 15	32.8 ± 0.4	85.4 ± 3.5	1708 ± 46	212 ± 3	124.1 ± 3.8
Aryl alcohols						
Benzyl alcohol	27 ± 0.9	54.5 ± 0.6	(2.02 ± 0.07) x 10^3	61 ± 3	226 ± 5	(3.7 ± 0.2) x 10^3
p-anisyl alcohol	24 ± 1.3	48 ± 1	(2 ± 0.12) x 10^3	88 ± 5	218 ± 8	(2.47 ± 0.17) x 10^3
m-anisyl alcohol	21 ± 0.8	140 ± 2	(6.6 ± 0.3) x 10^3	52.3 ± 1.8	289 ± 5	(5.5 ± 0.2) x 10^3
Veratryl alcohol	n.d.§	n.d.§	(4.7 + 0.2) x 10^3	n.d.§	n.d.§	(4.4 ± 0.1) x 10^3
4-Hydroxy benzyl alcohol	4.5 ± 0.6	20.1 ± 1.2	(4.5 + 0.6) x 10^3	19 ± 3.5	33 ± 3	(1.74 + 0.36) x 10^3
Furans						
HMF	6.5 ± 0.3	126 ± 1.5	(1.94 ± 0.09) x 10^4	14 ± 0.5	201 ± 2	(1.44 ± 0.05) x 10^4
HMFCA	26.9 ± 3	28.3 ± 1.3	(1.1 ± 0.1) x 10^3	27 ± 1.4	15.4 ± 0.3	570 ± 32

*The measurements were performed by varying the substrate concentration from 2.5 mM to 1.5 M for carbohy for polyols and diols, 100 μM to 200 mM for aryl alcohols and 100 μM to 100 mM for furans. The apparent steady were determined in 100 mM phosphate buffer, pH 7, at 25 °C using the colorimetric HRP/ABTS assay using 1.3 t enzyme. The affinity constant for substrate, K_m, and turnover number, k_{cat}, were calculated by nonlinear regressio Menten equation. Data represent means ± standard deviations (n = 3).

§ Individual k_{cat} and K_m values not determinable; k_{cat}/K_m values obtained from slope of linear v_0 versus [S] plots

In contrast, benzyl alcohol and the derivatives *p*-anisyl alcohol (4-methoxybenzyl alcohol), *m*-anisyl alcohol (3-methoxybenzyl alcohol), veratryl alcohol (3,4-methoxybenzyl alcohol), and 4-hydroxybenzyl alcohol all exhibited k_{cat}/K_m values of the order of 10^3 $M^{-1}s^{-1}$, with K_m values in the range 4 -27 mM. Among all substrates, both good and poor, k_{cat} values were in the range 12 – 140 s^{-1}, which implicates substrate binding, as approximated by K_m, as the primary driver of specificity (Table 1). Interestingly, substitution of the benzyl alcohol aromatic ring had little effect on the oxidation kinetics of aryl alcohols: 4-methoxylation and 3-methoxylation had no effect on K_m, whereas 4-hydroxylation decreased K_m by 6-fold compared to benzyl alcohol. As a result, catalytic efficiencies remained in the same order of magnitude, with similar values for *p*-anisyl and benzyl alcohol while they slightly increased for *m*-anisyl and 4-hydroxybenzyl alcohols, respectively. Despite not being able to obtain reliable independent kinetic constants for veratryl alcohol due to substrate inhibition, estimation of k_{cat}/K_m at low substrate concentrations indicated that dimethoxylation likewise did not affect catalysis significantly. Independent product formation and substrate depletion experiments, respectively, using an optimized Purpald® [4-Amino-3-hydrazino-5-mercapto-1,2,4-triazole] microplate assay indicated that HRP (2.3 μM) was not able to oxidize these aryl-alcohols or their corresponding aldehydes over a 15 minutes reaction time (Figure S7), thus ruling out any potential interference by the co-enzyme on the kinetic data. Likewise, HRP has no activity on HMF and its derivatives [59].

The observed specificity profile is notably distinct from that of *Cgr*AlcOx and *Cgl*AlcOx, which are only able to oxidase the parent compound, but not substituted benzyl alcohols, which is a surprise given that small molecule models of galactose oxidase are known to oxidize benzyl alcohols readily[60]. However, the catalytic efficiency of *Cgr*AAO towards benzyl alcohol is ca. two

orders-of-magnitude less than *Cgr*AlcOx and *Cgl*AlcOx, primarily due to differences in K_m values[36]. Comparison with two recently characterized AA5_2 AlcOx, *Por*AlcOx and *Chi*AlcOx, is not possible due to a lack of specific kinetic data on benzyl alcohols[37]. Also distinct from *Cgr*AlcOx and *Cgl*AlcOx, cinnamyl alcohol was a poor substrate for *Cgr*AAO; despite a competent specific activity (Table S1), we were unable to obtain kinetic constants for this substrate due to chromophore interference with the ABTS-HRP coupled assay at concentrations above 2.5 mM.

In accord with our initial substrate screen, Michaelis-Menten analysis revealed HMF to be the best substrate for *Cgr*AAO, with a specificity constant of 2 x 10^4 $M^{-1}.s^{-1}$. This value is ca. 3-fold higher than the best aryl alcohol derivative, *m*-anisyl alcohol, and is the result of a correspondingly lower K_m value in light of a similarly high k_{cat} value (Table 1). In the context of sustained interest in HMF valorization, we subsequently tested activity on the diverse oxidation products of HMF, namely HMFCA, DFF, and FFCA (Scheme 1). *Cgr*AAO was only able to catalyze the oxidation of HMFCA, albeit with a k_{cat}/K_m value one order-of-magnitude lower than for the parent compound (Table 1).

To validate these kinetics results, the products from individual reactions of *Cgr*AAO with HMF, HMFCA and DFF were analyzed using proton nuclear magnetic resonance (^{1}H-NMR) spectroscopy. Control reactions in which HMF, HMFCA or DFF were incubated with HRP and catalase in the absence of *Cgr*AAO-WT showed no conversion after 16 h (data not shown). In contrast, 20 mM HMF was quantitatively converted to DFF after 16 h by incubation with *Cgr*AAO-WT, HRP and catalase, with no observable subsequent oxidation products (Figure 3).

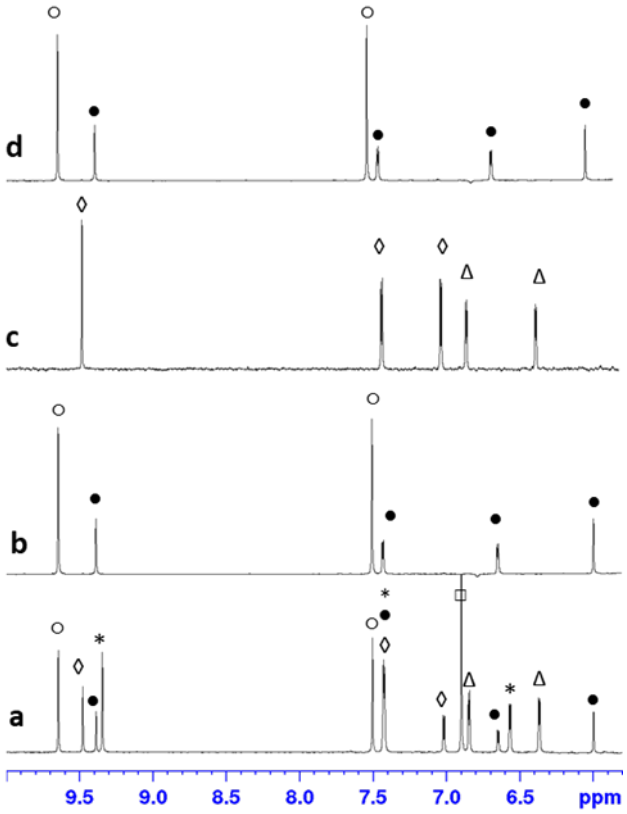


Figure 3. *Cgr*AAO HMF oxidation product analysis. ^{1}H NMR spectra (400 MHz, 1:9 D_2O:phosphate buffer, 100 mM, pH 7); (a) Standards mixture at 5 mM of HMF (*), DFF (O) and its hydrate form DFF_{hyd}(●), HMFCA (Δ), FFCA (◊) and FDCA (□). Reaction product profiles after 16 h incubation with *Cgr*AAO in the presence of catalase and HRP for (b) HMF at 20 mM showing full conversion to DFF, (c) HMFCA at 10 mM showing 46 % conversion to FFCA, and (d) DFF at 10 mM showing no conversion.

DFF was observed to exist in equilibrium with the monohydrated form (DFF_{hyd}, Figure 1), with a characteristic chemical shift of 5.99 ppm, corresponding to the C-H of the aldehyde hydrate, in a 1:1 ratio with the dialdehyde. Analogously, 10 mM HMFCA was oxidized to FFCA with 46 % conversion after incubation for 16 h (Figure 3), consistent with the slower initial rate kinetics

measured for this substrate (Table 1). Thus, the carboxylic acid of HMFCA was less well tolerated than the aldehyde of HMF by *Cgr*AAO-WT. The lack of further oxidation of both DFF and FFCA demonstrates that *Cgr*AAO-WT is highly specific for primary alcohol oxidation and does not act on aldehydes nor their hydrates (Figure 3).

Three-dimensional structure and active site geometry

To illuminate the structural features responsible for the distinct specificity profile of *Cgr*AAO, we combined X-ray crystallography and electron paramagnetic resonance (EPR) spectroscopy. The tertiary structure of *Cgr*AAO-WT was solved by molecular replacement using the *apo* form of *Cgr*AlcOx (PDB entry 5C86[36]) as a model (Figure 4A and B).

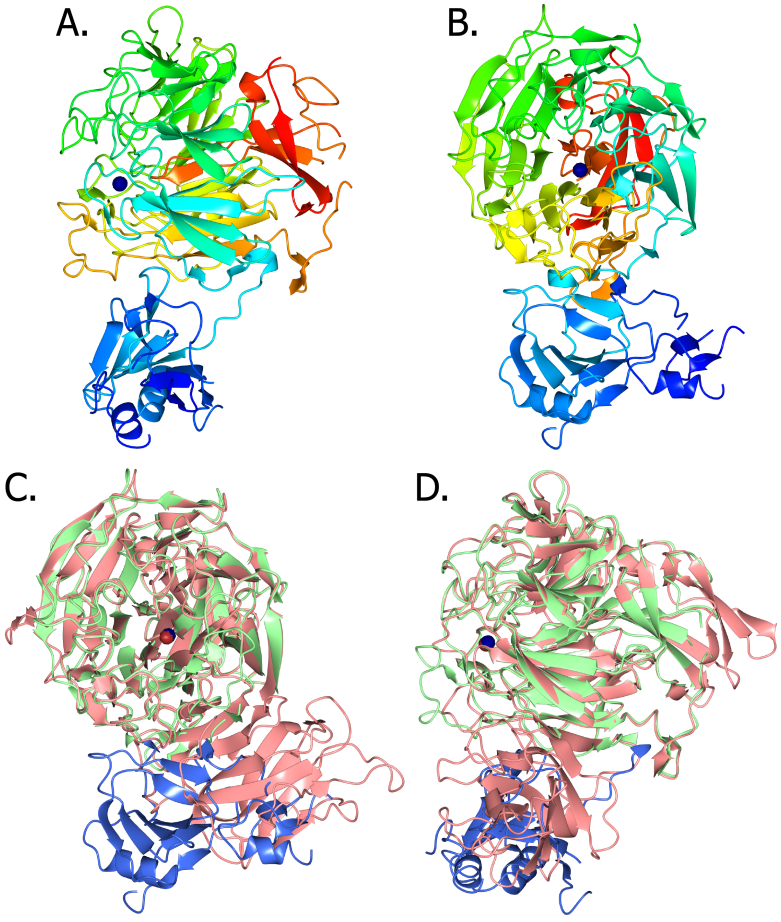


Figure 4 Tertiary structure of *Cgr*AAO. A) Overall fold (color blend from blue to red), **B)** Orthogonal view. The Cu ion is showed as a dark blue sphere. **C)** Overlay of *Fgr*GalOx (pink, 1GOF) on *Cgr*AAO-WT, **D)** Orthogonal view. PAN_1 domain in blue and catalytic domain in light green. Figures made using CCP4MG[61].

As anticipated from the primary structure, *Cgr*AAO-WT comprises three domains. The central, seven-bladed beta-propeller domain and the C-terminal, immunoglobulin-like beta-sandwich together constitute the AA5_2 catalytic module (Val199 to Leu689). This arrangement is directly homologous the AA5_2 modules of *Cgr*AlcOx[36] and *Fgr*GalOx[42]. Only residues Thr10-Thr21, Leu5-Val61, Leu91-Thr116, and Glu119-Ala198 of the N-terminal domain could be modelled in the original wild-type structure (PDB ID 6RYV), although this was later increased to cover Thr10-Tyr28, Leu35-Lys72 and Asp84-Ala198 in a second, higher resolution *apo-Cgr*AAO-WT structure (PDB ID 6STX) resulting from an attempt to generate a ligand complex by co-crystallization with the soft-Lewis-base substrate analog benzylmercaptan. The N-terminal domain contains conserved disulfide bridges and an overall fold typical of a PAN_1 domain[55].

Specifically, the N-terminal domain contains disulfide links between Cys residues 51 and 57, 145 and 167, and 149 and 155, as well as a unique disulfide link between Cys 104 and 123. This domain also contains the only two α-helices in the structure, at the exterior surface of the protein; the shorter one lies near a 2-stranded β-sheet near the N-terminus, and the longer one has a 6-stranded β-sheet twisted around it. An additional 3-stranded β-sheet is sandwiched between the latter sheet and the central beta-propeller domain. *Cgr*AlcOx lacks a corresponding N-terminal domain altogether[36], while *Fgr*GalOx contains a CBM32 module in this position that is not homologous to the PAN_1 domain (Figure 4C and D). The functional role of this domain in

*Cgr*AAO is currently unclear, but PAN_1 domains are known to mediate protein-protein and protein-carbohydrate interactions[55].

As in the other two available structures of AA5_2 oxidases[36, 42], the seven individual blades of the β-propeller fold are comprised of four anti-parallel β-strands (Figure 4). The active-site copper ion is located on the solvent-accessible side opposite to the C-terminal domain, and orthogonal to the N-terminal domain. The C-terminal domain is comprised mainly of loops at the peripheral edge of the protein and β-strands, which partly form a beta-sandwich, with gaps due to residue disorder/mobility between Thr608 and Thr621, and Val645 and Thr655 in *Cgr*AAO-WT (PDB ID 6RYV), which could be modelled in the subsequent *apo*-*Cgr*AAO-WT structure (PDB ID 6STX). A loop comprised of residues Leu627 to Gln638 extends from between two β-strands of the C-terminal domain through the center of the β-propeller, thereby orienting His632, of which NE2 is coordinated (2.2 Å) to the copper ion on the opposite side of the protein. Despite absence of the copper ion in the *apo*-*Cgr*AAO-WT structure, the loop and His632 lie in a very similar orientation. This arrangement is very similar to that in *Fgr*GalOx and *Cgr*AlcOx, for which the equivalent histidine residues are His581 and His423, respectively.

In addition to His632, the catalytic copper ion is coordinated by the NE2 atom of His534 (2.1 Å) and the OH group of Tyr316 (2.3 Å) that forms the unusual but characteristic thioether linkage to Cys292 (first observed in the structure of *Fgr*GalOx[42] and later in *Cgr*AlcOx[36]) (Figure 5).

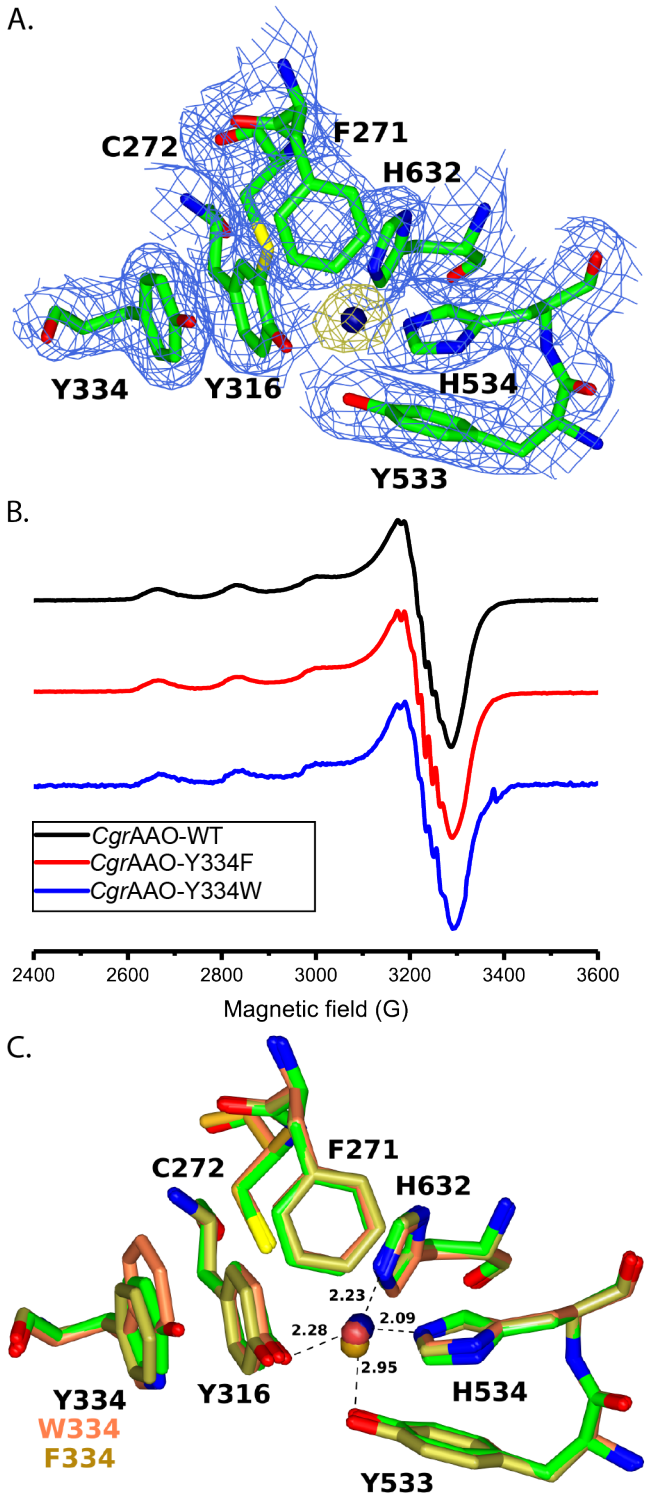


Figure 5. Active site of *Cgr*AAO-WT, *Cgr*AAO-Y334W and *Cgr*AAO-Y334F. A) Observed electron density map of *Cgr*AAO-WT, maximum likelihood weighted (REFMAC) $2F_{obs}$-F_{calc} map

contoured at 0.0172 $e^-/Å^3$ (light blue) and anomalous difference density map contoured at 0.0394 $e^-/Å^3$ (gold). **B)** Continuous wave X-band EPR spectra (9.30 GHz, 165 K) of *Cgr*AAO-WT, -Y334F and –Y334W in 100 mM Na phosphate buffer pH 7.0 with 10% *v/v* glycerol. **C)** *Cgr*AAO-WT, -Y334F and –Y334W active site residues overlay with WT interactions with copper. C atoms for WT in green, Y334F in gold, Y334W in coral, labels for WT in black.

Additional coordination is provided by the OH group of Tyr533 (3.0 Å), which lies opposite the Tyr316-Cys292 moiety, and there is a water molecule 3.4 Å from the metal ion that forms a hydrogen bond to Tyr316. Overall, this primary coordination sphere is directly homologous to that observed in *Fgr*GalOx and *Cgr*AlcOx (Figure 5). Also characteristic of AA5_2, an aromatic stacking interaction is observed over the thioether bond, here from Tyr334. This second-shell interaction is similar to that in *Fgr*GalOx and *Cgr*AlcOx, except that a tryptophan (Trp290) and a phenylalanine (Phe138), respectively, are found in place of the tyrosine (see Active-site Mutagenesis section below). Additional aromatic and hydrophobic groups line the active site: four phenylalanine residues (Phe239, Phe271, Phe479, and Phe502), Tyr444, Val501, and Trp373. Notably, the Phe271 ring forms an edge-to-face interaction with Tyr316-Cys292 moiety (Figure 5).

The electron paramagnetic resonance (EPR) spectrum *Cgr*AAO-WT (Figure 5) showed that, in the resting state, the active-site is in its semi-reduced form, with the Cu ion in the +2 oxidation state and a neutral (*i.e.* 1-electron reduced) Cys-Tyr ligand, identical to previous reports on *Fgr*GalOx, *Cgr*AlcOx, and *Cgl*AlcOx[36, 62-65]. The spectrum indicated a singly occupied molecular orbital (SOMO) with mostly $d(x^2-y^2)$ character and a typical Type 2 copper site according to the Peisach and Blumberg classification[66] (g_z = 2.27 – 2.28, A_z = 530 MHz, Table S2). The spectrum

also showed well-resolved superhyperfine (SHF) coupling to the coordinated nitrogen atoms, which could be simulated by addition of two N atoms with coupling in the order of 45 MHz (Table S2 and Figure S8).

Active-site mutagenesis

To explore the effect of differences in the second-shell stacking residue on substrate specificity and catalysis in light of known natural variants of AA5_2 members, we produced the site-directed mutants *Cgr*AAO-Y334F (mimicking *Cgr*AlcOx) and *Cgr*AAO-Y334W (mimicking *Fgr*GalOx). The recombinant production of both was similar to the wild-type enzyme and yielded 18 mg of *Cgr*AAO-Y334W and 11 mg of *Cgr*AAO-Y334F after purification (Figure S2). SDS-PAGE analysis following PNGaseF or EndoH treatment also indicated a similar glycosylation pattern to the wild-type and molecular masses consistent with those calculated from the primary structure (Figure S2). ICP-MS analysis indicated that both variants were fully loaded with copper, on the basis of copper-to-protein ratios of 1.07 ± 0.02 and 1.08 ± 0.02 *Cgr*AAO-Y334F and *Cgr*AAO-Y334W, respectively.

Structural analysis by crystallography and EPR confirmed the overall fidelity of protein folding and revealed a strikingly similar active-site structure. Indeed, superposition of both site-directed mutants with the wild-type enzyme shows that all first-shell coordinating residues are positioned identically and that the aromatic rings of Tyr334, Phe334, and Trp334 are co-planar and stacked homologously with the Tyr-Cys moiety (Figure 5). Apart from the obvious sidechain substitution, the only major active-site differences between the variants concern associated water. As discussed above, the hydroxyl group of Tyr334 in *Cgr*AAO-WT forms a hydrogen bond to a water molecule, which is also hydrogen bonded to OG of Ser289 and O of Ser240. As expected, this hydrogen

bonding is disrupted in *Cgr*AAO-Y334F. There are also no hydrogen bonding interactions for NE1 of Trp334 in *Cgr*AAO-Y334W, just as there are none for Trp290 in *Fgr*GalOx (Figure 5). EPR spectra of *Cgr*AAO-Y334W and *Cgr*AAO-Y334F (Figure 5) were likewise highly similar to the wild-type (g_z = 2.27 – 2.28, with A_z = 530 MHz for *Cgr*AAO-WT and *Cgr*AAO-Y334F, and A_z = 515 MHz for *Cgr*AAO-Y334W, Table S2). The slight but significant difference in the spin-Hamiltonian parameters of the Y334W mutant is in accordance with the ability of this second-shell "stacking residue" to modulate the redox-active Tyr316 residue's interaction with the copper ion, and thus the reactivity of the system.

To wit, substitution of Tyr334 with phenylalanine, as in *Cgr*AlcOx, or tryptophan, as in *Fgr*GalOx, resulted in distinct substrate specificities, as determined by specific activity measurements on the same panel of substrates used to characterize the wild-type enzyme (Figure 2). Briefly, *Cgr*AAO-Y334W exhibited increased specific activity toward all carbohydrates for which activity was detected for *Cgr*AAO-WT, and decreased activity towards diols, aryl alcohols and furans. *Cgr*AAO-Y334F, on the other hand, exhibited specific activities that were comparable to the wild-type for carbohydrates, diols, aryl alcohols, HMF, and HMFCA. In light of these results, melibiose and HMF were used to perform pH-rate profile and temperature stability measurements with *Cgr*AAO-Y334W and *Cgr*AAO-Y334F, respectively. Both of these variants showed little deviation from the behavior of the wild-type enzyme (Figure S3 and S4). Subsequently, full Michaelis-Menten kinetic analyses were performed for selected substrates at pH 7.0 and 25 °C for direct comparison with *Cgr*AAO-WT (Table 1 and Figure S5).

The conservative substitution of tyrosine for phenylalanine-334 resulted in essentially no change in catalytic efficiency for all compounds, with K_m, k_{cat}, and k_{cat}/K_m values differing less than a factor of two versus *Cgr*AAO-WT (Table 1). Possible exceptions were the poor substrates melibiose,

raffinose, and 1,2-propanediol, although it should be noted that the high apparent K_m values associated with these compounds warrant cautious interpretation of the fitted kinetic parameters. Analogous to the wild-type enzyme, product analysis indicated that *Cgr*AAO-Y334F produced only DFF from HMF, albeit with a reduced conversion (75%) over a similar reaction period (Figure S9). Taken together, the data indicate that the distinct substrate specificity profile exhibited by *Cgr*AAO-WT vis-à-vis the previously characterized *Cgr*AlcOx (and *Cgl*AlcOx)[36] do not arise solely from a single active-site substitution, but collectively from many differences in their primary structures (Figure S1).

In contrast, substitution of tyrosine with tryptophan, to produce the *Fgr*GalOx-like variant *Cgr*AAO-Y334W, resulted in a large, concordant shift in overall specificity. Strikingly, activity for all aryl alcohols was significantly reduced, while activity on galactose-containing carbohydrates was increased. On one hand, large increases in K_m values resulted in an inability to achieve substrate saturation of *Cgr*AAO-Y334W with aryl alcohols, despite apparent k_{cat}/K_m values (determined from linear fits to the data, Figure S5) exhibiting reductions of only 2- to 5-fold. On the other hand, increases in galactosyl specificity of *Cgr*AAO-Y334W (10- to 25-fold in the cases of the free sugar and the alpha-linked galactosides melibiose and raffinose) resulted from the combination of reduced K_m values and increased k_{cat} values. Beta-linked lactose had a more modest increase in k_{cat}/K_m in the *Cgr*AAO-Y334W mutant (4-fold), primarily the result of a corresponding reduction in K_m. The ability to achieve substrate saturation was also lost in 1,2- and 1,3-propane diols, resulting in significantly lower apparent k_{cat}/K_m values than for the wild-type enzyme (Figure S5). Yet, glycerol oxidation was essentially unaffected.

Also notable, an order-of-magnitude decrease in catalytic efficiency was observed for both HMF and HMFCA. In the case of HMF, this was primarily a K_m effect; k_{cat} was essentially unaffected.

In contrast, the Y334W mutation reduced the k_{cat} value for HMFCA, with K_m essentially unchanged. Although HMF remained the best substrate for *Cgr*AAO-Y334W, the k_{cat}/K_m value was only marginally higher on this substrate than for melibiose and raffinose, indicating a significant shift toward carbohydrate specificity. Product analysis revealed DFF as the only product of HMF oxidation, but with even lower conversion (48%) than for the Y334F variant, concordant with the correspondingly lower activity of *Cgr*AAO-Y334W on this compound (Figure S9).

DISCUSSION

The oxidation of alcohols is of particular biotechnological interest[67-69] and over the last years, oxidation of the bio-based platform compound HMF has gained considerable attention due to its rich chemistry and potentially sustainable production[70]. While several chemical methods have been developed for the specific oxidation of HMF to DFF[71-74], they have individual drawbacks including the use of expensive metals (Pt, V or Ru) or corrosive chemicals (bromide), harsh reaction conditions (elevated temperatures and/or pressures), and/or wastes associated with catalyst separation and product purification. Spurred by an interest in developing "greener" processes, the past five years have witnessed a significant growth in interest in the use of biocatalysts to generate HMF oxidation products with high selectivity under mild reaction conditions[18].

In particular, a diversity of enzymes have been investigated for the conversion of HMF to carboxylic acid derivatives, including early studies on a *Caldariomyces fumago* (fungal) chloroperoxidase[75, 76], and more recent work on a range of bacterial and fungal flavin-dependent Glucose-methanol-choline (GMC) oxidoreductases/aryl-alcohol oxidases[14, 17, 76], an *Escherichia coli* xanthine oxidase[77], three fungal laccases plus (2,2,6,6-tetramethylpiperidin-1-yl)oxidanyl (TEMPO)[77], and three *Pycnoporus cinnabarinus* (fungal) AA5_1 glyoxal oxidases[59]. Of particular

relevance to the present study, several enzymes have been shown to produce the likewise valuable dialdehyde DFF as a terminal product by selective, limited oxidation, including yeast alcohol oxidases[77], a *Myceliophthora thermophila* (fungal) AA5_1 glyoxal oxidase[78], and wild-type and mutant forms of the archetypal *Fusarium graminicola*/*Fusarium austroamericanum* AA5_2 galactose oxidase[21, 77, 79].

In this context, *Cgr*AAO is unique as a biocatalyst since it is ostensibly the first wild-type AA5 member with predominant aryl-alcohol oxidase (EC 1.1.3.7) and HMF oxidase activity (EC 1.1.3.47). Until now, this activity has only been associated with AA3 members, in particular flavin-dependent GMC oxidoreductases[34, 80, 81]. Although the oxidation of benzylic and other aryl alcohols is comparatively efficient in the case of mutant *Fgr*GalOx variants[79, 82] it is generally poor for wild-type AA5 GalOx enzymes[44, 63, 83]. And while benzyl alcohol is itself a competent substrate for the paralog *Cgr*AlcOx, ring-substituted congeners are not tolerated[36]. In this sense, the substrate scope of *Cgr*AAO (Table 1) recapitulates that of the archetypal aryl-alcohol oxidase[48], is distinct vis-à-vis other AA5_2 members, and is correlated with its unique subgroup membership (Figure 1). Where necessary, the enzyme can be referred to as "*Cgr*AAO (AA5_2)" to clearly distinguish it from AA3 AAOs that may be eventually characterized from this and other *Collectotrichum* species (http://www.colletotrichum.org/genomics/)[52].

A further aspect of the uniqueness of *Cgr*AAO is apparent upon consideration of its ability to oxidize HMF and HMFCA, versus those enzymes introduced above for which detailed kinetic parameters are available (Table S3). *Cgr*AAO exhibits a k_{cat}/K_m value for HMF two orders of magnitude higher than those of the three *Pycnoporus cinnabarinus* (fungal) glyoxal oxidases[59] and one order of magnitude higher than that of the *Myceliophthora thermophila* (fungal) glyoxal oxidase[78] from AA5_1. Likewise, the k_{cat}/K_m value for HMF of *Cgr*AAO is two orders of

magnitude higher than that of a *Pleurotus eryngii* (fungal) flavin-dependent GMC oxidoreductase from AA3, with which incomplete (ca. 60%) conversion was additionally observed after 25 h[17]. *Cgr*AAO also outpaces the eponymous *Methylovorus* sp. (bacterial) HMF oxidase, an AA3 flavin-dependent GMC oxidoreductase[14] by a factor of 3 in k_{cat}/K_m for HMF, due to a ca. one order of magnitude higher k_{cat} value (Table S3). *Cgr*AAO is likewise superior to the *Methylovorus* sp. HMF oxidase and the *Pleurotus eryngii* (fungal) oxidase in the oxidation of HMFCA. As such, we had originally considered naming *Cgr*AAO *Cgr*HMFOx, which was ultimately not justified in light of the marginal selectivity difference over aryl alcohols (Table 1).

Notably, the majority of the enzymes discussed here convert HMF to higher oxidation products[14, 17, 59, 75-77], whereas *Cgr*AAO stops cleanly at DFF (Figure 3). Specifically, 270 nM of *Cgr*AAO-WT in the presence of HRP and catalase was able to fully convert 20 mM HMF into DFF within 16 h in the absence of side products (Figure 3). Although weak oxidase activity was detected for DFF during initial screening experiments (Table S1), the addition of ca. 10-fold more enzyme necessary to quantify activity (2 μM for DFF versus 270 nM for HMF) nonetheless resulted in a 26400-fold difference in specific activity. Intriguingly, while previous studies on the oxidation of aryl alcohols by a *Fgr*GalOx mutant have linked a higher degree of product aldehyde hydrate formation to increased subsequent conversion to carboxylic acids[82], DFF was not a good substrate for *Cgr*AAO despite an equilibrium favoring a high proportion (50%) of the monohydrate (Figure 3 cf. Scheme 1). Commensurate with the initial-rate kinetics results (Table 1), *Cgr*AAO was not able to fully convert the primary alcohol of HMFCA to generate to FFCA using the same conditions as for complete HMF oxidation (46% of HMFCA was oxidized to FFCA in 16 h, Figure 3).

To our knowledge, only the wild-type *F. graminearum* (equivalent to *F. austroamericanum* and *Dactylium dendroides*) GalOx is capable of similar, high yielding, selective oxidation of HMF to

DFF[21, 77]. The recently characterized *M. thermophila* AA5_1 glyoxal oxidase, which has similar k_{cat}/K_m values for HMF and glyoxal, could achieve only 56% conversion of 1 mM HMF to DFF over 24 h ([E] = 500 nM)[78]. Likewise, diverse yeast alcohol oxidases produced 16-41% conversion to DFF as the exclusive product over 72 h[77]. It should be noted that even in the successful GalOx system devised in the *tour-de-force* study of Qin *et al*., incomplete oxidation posed a particular challenge, requiring the use of an additional, exotic process step (deep eutectic solvent extraction) to isolate the desired DFF[77]. From a bioprocess perspective, an additional advantage of AA5 copper radical oxidases is a lack of dependence on complex organic cofactors for catalysis.

Finally, we were pleased to observe that we could modulate the activity profile of *Cgr*AAO by mutation of the key second-shell active-site residue, Tyr334, to the homologous tryptophan in GalOxs, to achieve a corresponding increase in galactosyl activity. Previous extensive protein engineering and docking studies of *Fgr*GalOx have highlighted a series of residues composing the active-site pocket above the copper center that influences substrate binding and/or turnover[84-87] (Figure S1A, green highlighting). Of these, Trp290 plays a pivotal role in *Fgr*GalOx, as it contributes to the stability of the active site tyrosine radical through dual electronic and steric effects associated with the stacking of the indole ring with the crosslinked tyrosine-cysteine copper ligand[64]. Trp290 is also within hydrogen bonding distance with the oxygen atom located on hydroxyl group attached to the carbon 5 of D-galactose and forms an face-to-face aromatic interaction with Tyr329 (*Fgr*GalOx numbering), exhibiting an interplanar angle of 26°[88].

Our analysis of *Cgr*AAO-Tyr334 variants emphasized the crucial role of this second-coordination sphere residue. Indeed, the presence of a phenol group (WT) or a benzene ring (Y334F) seems to preferentially direct the oxidative activity of *Cgr*AAO towards aryl alcohols and furans rather than carbohydrates. On the other hand, the presence of an indole ring modulates the

preference of *Cgr*AAO towards carbohydrates as this variant exhibited better catalytic efficiencies towards these compounds (Table 1). Indeed, a measurable, albeit small, shift in the spin Hamiltonian parameters for the *Cgr*AAO-Y334W mutant can be observed from simulation of the experimental data (Table S2), consistent with a remote electronic effect on the copper center.

In notable contrast, mutagenesis of the structurally homologous phenylalanine residue in *Cgr*AlcOx to a tryptophan did not increase GalOx activity, but rather decreased benzyl alcohol specific activity by half and abolished activity for alkanols[36]. The "reverse" mutation of the native tryptophan in *Fgr*GalOx to phenylalanine led to a large reduction in GalOx activity (primarily due to a K_m effect) without significant increase in 2-methylene-1,3-propanediol oxidase activity[64]. Unfortunately, activities toward aryl alcohols and HMF were not performed in these early, seminal studies, thereby precluding a comprehensive like-for-like comparison. Likewise, the paralogous *Colletotrichum graminicola* raffinose/galactose oxidase, which bears a second-shell tyrosine identical to *Cgr*AAO (Figure S1A) was not assessed for the oxidation of aryl alcohols[43]. Thus, a lack of systematic measurements against a comprehensive library of potential substrates precludes precise definition of protein structure-function relationships that dictate compound class specificity among AA5_2 members. In light of the phylogenetic diversity of AA5 and the limited structural enzymology on a handful of members, we anticipate that further exploration of this family will yield additional unique biocatalysts for green chemical transformations.

MATERIALS AND METHODS

Sequence alignment, phylogenetic and functional analysis

A total of 212 bacterial and 43 fungal AA5_2 sequences were collected from the CAZy database in December 2017[34]. Sequences of characterized AA5_1 members were also retrieved from the

CAZy database. In addition, full-length sequences of *Fgr*GalOx (Uniprot P0CS93) and *Cgr*AlcOx (Genbank EFQ30446), as well as their corresponding catalytic domains, were used as templates for BLAST analysis, retrieving 455 sequences, of Ascomycota and Basidiomycota translated genomes in the Mycocosm portal for fungal genomics (https://genome.jgi.doe.gov/programs/fungi/index.jsf)[89, 90]. Where present, signal peptides and additional modules, such as carbohydrate-binding modules, were removed to isolate the catalytic modules for subsequent analyses. Catalytic modules sharing 100% identity were down-sampled to one sequence to eliminate redundancy. A multiple sequence alignment was created with MAFFT v7.402 using the L-INS-i algorithm[91], on The CIPRES Science Gateway[92] (www.phylo.org). Using this alignment, catalytic modules having deletions/substitutions at key active site residues, namely Cys272 and Tyr316 (which form the cross-linked thioether-tyrosyl cofactor) and other copper coordinating residues of *Fgr*GalOx[42], were removed. Moreover, obviously erroneous sequences generated by incorrect splicing predictions were removed through identification of unusually long deletions/insertions.

Employing this curated multiple sequence alignment, comprising 392 AA5 catalytic modules, a maximum likelihood phylogenetic tree was produced using RAxML v.8, with 100 bootstrap replications,[93] on The CIPRES Science Gateway portal. The defined monophyletic groups were supported by bootstrap values >75 and the tree was visualized using FigTree. SSNs were generated by computing BLASTP[94] all-versus-all local alignments of the 392 curated AA5_2 catalytic domains using SSNpipe (https://github.com/ahvdk/SSNpipe), which generated the *E*-value, bit score, alignment length, sequence identity and sequence similarity for all sequence pairs. The data was filtered using a bit score threshold between 500 to 600, with increments of 25, to generate the final SSNs. A bit score threshold of 550, clustering the sequences into groups, which resolves the

same monophyletic groups as those observed in the phylogenetic tree, was retained. The SSNs were visualized with Cytoscape using yFiles Organic Layout[51] and coloring of each node was based on monophyletic groups inferred from the phylogenetic tree.

DNA cloning, strain production, and site-directed mutagenesis

cDNA encoding wild-type *Cgr*AAO (*Cgr*AAO-WT, Genbank ID EFQ27661) without its predicted peptide signal (MVRSCAYKAIAAASLLSQLASA)[95] was commercially synthesized (Integrated DNA Technologies, Coralville, USA) into pUCIDT-AMP and subcloned into the *P. pastoris* expression vector *pPicZα-A* (Invitrogen) flush with the *S. cerevisiae* α-factor signal peptide coding sequence on the N-terminal part and the *c-myc* epitope followed by a His_6 tag on the C-terminal part. *Cgr*AAO-pUCIDT-AMP and *pPicZα-A* were separately digested using EcoRI and XbaI restriction enzymes (New England Biolabs) for 2 h at 37 °C. The resulting products were then separated on a 1% w/v agarose gel and the band corresponding to the digested plasmid and the coding gene were excised and purified using a gel/PCR DNA fragments kit (Geneaid, New Taipei City, Taiwan).

The digested products were combined in a 3:1 insert to vector ratio and ligated using T4 DNA ligase (Lucigen, Middleton, WI, USA) overnight at 4 °C. After inactivation of the ligase at 70 °C for 20 minutes, the ligated product was transformed into chemically competent *E. coli DH5α* using the heat-shock method. Transformants were grown overnight at 37 °C on a Luria-Bertani-Low-Salt (LBLS) agar plate and selected using 25 $\mu g.mL^{-1}$ Zeocin (Invitrogen). Surviving colonies were picked and grown overnight at 37 °C in 5 mL of Luria Bertani Low Salt (LBLS) medium containing 25 $\mu g.mL^{-1}$ Zeocin. Plasmids were then extracted using a commercial miniprep-kit (Geneaid, New Taipei City, Taiwan). The presence of the insert was confirmed by restriction

enzyme digestion followed by agarose (1% w/v) gel electrophoresis and the correct insertion into the corresponding vector was verified by DNA sequencing.

Transformation of *pEFQ27661* into *P. pastoris* X33 was performed by digesting 20 μg of plasmid DNA for 2 h at 37 °C with PmeI (New England Biolabs). After enzyme inactivation at 65 °C for 10 minutes, the linearized plasmid was purified using a gel/PCR DNA fragments kit (Geneaid, New Taipei City, Taiwan). The resulting digested plasmid (2.26 μg) was transformed into electrocompetent *P. pastoris* X33 cells prepared on the same day[96]. Transformants were grown on yeast extract peptone dextrose (YPD) agar plates for 3 days at 30 °C and selected against 100 or 500 µg.mL^{-1} Zeocin. Several colonies were streaked on YPD agar plates containing either 100 or 500 μg.mL^{-1} Zeocin and grown for 3 days at 30 °C to select clones with multicopy insertions. Nine colonies selected from 100 or 500 μg.mL^{-1} Zeocin plates were inoculated into 50 mL sterile conical tubes containing 10 mL of buffered complex glycerol medium (BMGY) and grown in a shaking incubator overnight at 30 °C and 250 r.p.m. Cells were then pelleted by centrifugation at 3000 r.p.m. for 10 minutes, the BMGY medium was discarded and replaced with 2 mL of BMMY containing 1% (v/v) methanol. The cultures were shaken at 250 r.p.m. over 4 days at either 25 °C or 16 °C with regular feeding of 1% methanol every 24h to ensure continued protein expression. Secreted proteins were separated from the cells by centrifugation at 4700 r.p.m. for 10 minutes and protein production was monitored using SDS-PAGE. The clone yielding the highest amount of protein was retained for large-scale production.

Point variants of EFQ27661 were generated using the Q5® Site-Directed Mutagenesis Kit (New England Biolabs) according to the supplier's protocol. Primer sequences and their corresponding nucleotide replacements are listed in Table S4. The PCR products were then treated with a kinase-ligase-DpnI enzyme mix for 5 minutes at room temperature prior to their transformation into

competent *E. coli DH5α* using the heat shock method. Positive transformants were selected on LBLS plates containing 25 μg.mL^{-1} Zeocin at 37 °C overnight. Plasmids were isolated from the colonies, and those with the correct mutation were identified by sequencing. Each plasmid was then linearized with PmeI and transformed into *P. pastoris* X33 by electroporation, and positive transformants were selected as described above.

Protein production and purification

Single colonies of *P. pastoris* X33 expressing the gene of interest were individually streaked on a YPD agar plate containing 500 μg.mL^{-1} Zeocin and grown for 2 days at 30 °C. A single colony was then used to inoculate 10 mL of YPDand grown at 30 °C shaken at 200 r.p.m. for 8 h. Biomass production was initiated by addition of the 10 mL culture to 1 L of BMGY medium in a 4 L baffled flask with a foam cap, which was shaken at 250 r.p.m at 30 °C until the OD_{600} reached 5-6. Cells were harvested by centrifugation at 3700 rpm for 10 minutes at room temperature under sterile conditions. The cell pellet was quickly resuspended in 400 mL of BMMY supplemented with 0.5 mM of $CuSO_4$ and 1 % methanol (v/v) in a sterile 2 L flask with a foam cap. The cultures were shaken at 250 r.p.m. at 16 °C, or room temperature depending on the protein, for 4 days with regular feeding of methanol every 24 h to ensure continuing protein expression. After 4 days, proteins in the extracellular medium were separated from cells by centrifugation at 4700 r.p.m. for 15 minutes at 4 °C and used for subsequent purification.

Media containing the protein of interest were first raised to pH 7.4 by addition of 5 M NaOH with stirring, cleared by filtration through a sterile 0.45 μm polyethersulfone membrane (Millipore, Massachusetts, USA) and concentrated 10-fold by ultrafiltration through a polyether-sulfone membrane of 10 kDa cut-off (Vivaflow crossflow cassette, Sartorius, Les Ulis, France). The

concentrated protein sample was loaded onto a 5 mL pre-packed Ni-NTA (GE Healthcare, Velizy-Villacoublay, France) column pre-equilibrated with five column volumes of binding buffer (20 mM sodium phosphate buffer, 500 mM NaCl, 10 mM imidazole at pH 7.4). The column was then washed with 50 mL of the same binding buffer and proteins were eluted using a linear gradient from 2 to 100 % of 500 mM imidazole in 20 mM sodium phosphate buffer, 500 mM NaCl at pH 7.4 at 5 mL.min^{-1} and collected in 1 mL fractions. The fractions of interest were pooled and further purified and desalted using size exclusion chromatography on a 16 mm x 600 mm column of Superdex 75 pre-equilibrated with 100 mM sodium phosphate at pH 7.0 at 1 mL.min^{-1} and collected in 1 mL fractions. The fractions of interest were pooled and concentrated using Vivaspin® centrifugal concentrators with a 30 kDa cut-off (Sartorius, Göttingen, Germany) at 4 °C. Finally, the proteins were aliquoted, flash frozen in liquid nitrogen and stored at -20 °C. SDS-PAGE were performed using pre-cast 4-20% (w/v) polyacrylamide gels in the presence of 2% (w/v) SDS under reducing conditions; proteins were visualized with Coomassie blue R-250. Protein concentrations were determined by measuring A_{280} using extinction coefficients of 107760 M^{-1}.cm^{-1}, 111770 M^{-1}.cm^{-1} and 106270 M^{-1}.cm^{-1} for *Cgr*AAO-WT, -Y334W, and -Y334F respectively, which were calculated using the ProtParam tool on the ExPASy server (http://web.expasy.org/protparam/).

Recombinant proteins were N-deglycosylated by treatment with endoglycosidase H_f (New England Biolabs) or N-glycosidase F (PNGase F, New England Biolabs) according to the manufacturer's protocol. Briefly, after heat denaturation at 100 °C for 10 minutes, 10 μg of pure enzyme were incubated with either endo-H_f or PNGase F for 1 h at 37 °C. Deglycosylation efficiency was analyzed on pre-cast 4-20% (w/v) polyacrylamide gradient gels in the presence of 2% (w/v) SDS under reducing conditions.

Enzyme kinetic analysis

Oxidation of all substrates was determined using a colorimetric HRP-ABTS coupled assay essentially as previously described[36]. Upon substrate oxidation, AA5_2 CROs consume 1 equivalent of O_2 and concomitantly generate 1 equivalent of H_2O_2, which is used in turn by HRP to catalyze ABTS oxidation (ε_{420} = 36000 $M^{-1}.cm^{-1}$). One unit of AA5_2 activity was defined therefore as the amount of enzyme necessary for the oxidation of 2 μmol of ABTS per minute, corresponding to the consumption of 1 μmol of O_2 per minute.

Each assay was performed at 25 °C in 100 mM sodium phosphate buffer at pH 7 with 0.49 mM ABTS, 2.3 μM horseradish peroxidase (HRP), and different substrate concentrations: from 2.5 mM to 2 M for carbohydrates, 10 mM to 5 M for polyols and diols, 100 μM to 200 mM for benzyl alcohols and 100 μM to 100 mM for furans. Reactions were started with the addition of 1.3 to 13 μmole of purified enzyme and steady-state parameters were determined using Origin software on three replicates of each measurement point. The initial reaction rates at different substrate concentrations were fitted to the Michaelis-Menten equation for the determination of K_m and k_{cat} values.

To determine the effect of temperature on activity, aliquots of purified recombinant enzymes were incubated at 25 °C to 50 °C in 100 mM phosphate buffer at pH 7. Samples were taken-out at different time intervals for assays with 50 mM HMF for *Cgr*AAO-WT and Y334F or 500 mM melibiose for *Cgr*AAO-Y334W. For pH-rate profiles, aliquots of purified enzyme were incubated in 100 mM citrate-phosphate buffer from pH 2.5 to 7.0, in 100 mM phosphate buffer from pH 7.0 to 8.0 and in 100 mM glycine-NaOH buffer from pH 8.0 to 11. Each measurement was made in triplicate.

Enzyme product analysis

Raffinose (10 mM) was dissolved in water (final volume 500 μL) and treated with *Cgr*AAO-WT (100 μg), HRP (1 U/mg substrate), and catalase (115 U/mg substrate). Reactions were incubated for 16 hours at 25 °C. A negative control reaction was performed on the same reaction mix without *Cgr*AAO. At time points 0 h, 2 h, 4 h, 8 h and 16 h an aliquot of 100 μL was taken and enzymes were removed by ultrafiltration (10 kDa cut-off Centricon-Millipore, Billerica, MA, USA). Matrix-assisted laser desorption ionization coupled to time-of-flight mass spectrometry (MALDI-TOF MS) was performed on a Bruker Daltonics Autoflex System (Billerica, MA, U.S.A.). The matrix, 2,5-dihydroxy benzoic acid (DHB), was dissolved in 50% methanol in water to a concentration of 25 mg/ml. Oligosaccharide samples were mixed 1 : 1 (v/v) with the matrix solution. Five microliters of this solution were placed on a Bruker MTP 384 ground steel MALDI plate and left to air dry for 30 min prior to analysis.

Reactions containing 10 or 20 mM putative substrate (HMF, DFF or HMFCA) and 1 $mg.mL^{-1}$ of both catalase and HRP were initiated by the addition of 270 nM purified enzyme in a final volume of 1 mL (100 mM phosphate buffer, pH 7.0). For each reaction, a negative control was performed with identical reaction components omitting the purified enzyme. Reactions were incubated at room temperature for 16 h, at which time enzymes were removed by ultrafiltration (10 kDa cut-off Centricon-Millipore, Billerica, MA, USA). D_2O was added to the filtrate to a final composition of 10% (v/v). NMR experiments were run on a Bruker AV III HD 400 MHz spectrometer equipped with a BBFO smart probe. ^1H NMR spectra were collected with water suppression (4.7 ppm) using a standard pre-saturation pulse sequence. Chemical shifts were calibrated to external TMSI (0 ppm). Standards for each HMF, DFF, HMFCA, FFCA and FDCA

were used to identify distinct chemical shifts for each molecule. The integration value for relevant peak areas were used to determine percent conversions.

Possible activity of the co-enzyme HRP on relevant aromatic alcohols and aldehydes was assessed by aldehyde detection using Purpald® [4-Amino-3-hydrazino-5-mercapto-1,2,4-triazole] in microplate format, according to the protocol developed by Quesenberry and Lee[97]. Briefly, 100 μL of sample reaction was mixed with 100 μL of 34 mM Purpald® freshly prepared in 2 N NaOH and incubated at room temperature for 20 min. Then, 100 μL of 33 mM $NaIO_4$ freshly prepared in 0.2 N NaOH were added and A_{550} was read on a Tecan M200 plate reader. Standard curves between 100 μM and 20 mM for each aromatic aldehyde gave a linear response ($r^2 > 0.99$) with a conservative limit-of-detection of 50 μM. Sample reactions with HRP were made by incubating 50 mM aromatic alcohols or aromatic aldehydes and 10 mM H_2O_2 with or without 2.3 μM HRP for 15 minutes. Then, the enzyme was removed by ultrafiltration (10 kDa cut-off Centricon-Millipore, Billerica, MA, USA) before aldehyde detection.

Electron Paramagnetic Resonance (EPR) spectroscopy

Continuous wave (cw) X-band frozen solution EPR spectra of 0.19, 0.15 and 0.21 mM solutions of Cu(II)-*Cgr*AAO-WT, -Y334W and –Y334F, respectively, in 100 mM sodium phosphate buffer pH 7.0 were recorded, with and without 10% (v/v) glycerol, on a Bruker micro EMX spectrometer operating at ~9.30 GHz, with modulation amplitude of 4 G, modulation frequency of 100 kHz and microwave power of 10.02 mW at 165 K. Addition of glycerol did not affect the appearance of the spectra (Figure S8a). Spectral simulations were carried out using EasySpin 5.2.24 integrated within MatLab[98]. It was assumed that *g* and A tensors were axially coincident. Accurate determination of the perpendicular region parameters was not possible from X-band data alone, although it was

noted that satisfactory simulations could be achieved with the set of values reported in Table S2. Furthermore, addition of two coupled nitrogen atoms to the spin system was necessary to simulate the superhyperfine coupling observed in the experimental data. Raw EPR data are available on request through the University of York Library, Research Data York service (DOI: 10.15124/9caf3050-588d-4eac-b787-0b3a59027c51)

ICP-MS

Protein samples of *Cgr*AAO-WT (95 μM), -Y334W (115 μM), and –Y334F (88 μM) were digested in a total volume of 1 mL concentrated HNO_3 (70%) in a Teflon vial at 130 °C until total evaporation. Once dried, digested proteins were dissolved in 10 mL of 1% HNO_3 with 45Sc (20 ppb) as an internal standard. ICP-MS was performed using a NexION 300D (Perkin Elmer) equipped with a SC-2 DX autosampler, DXi-FAST micro-peristaltic pump, a cyclonic spray chamber, a triple cone interface, a quadrupole ion deflector and Universal Cell Technology. Calibration was performed using IV-Stock-4 calibration standard (Inorganic Ventures). All elements were run in reaction mode (using Dynamic Reaction Cell technology) using ammonia as a reaction gas to remove potential polyatomic interferences. The detection limit for ^{63}Cu was determined to be 0.104 ppb.

Crystallization

*Cgr*AAO-WT and the two variants Y334W and Y334F were deglycosylated by EndoH treatment and crystallized at 15-40 $mg.mL^{-1}$, with 0.1 mM copper (II) chloride added, using the sitting drop vapor diffusion technique. For *Cgr*AAO-WT, crystals were grown with a 1:1 ratio of protein:well solution in the drop, with a well solution comprised of 0.2 M ammonium sulfate, 0.1M sodium

acetate, pH 4.6, 30% (w/v) polyethylene glycol (PEG) 2000 monomethyl ether. *Cgr*AAO-WT crystals were used to make a seed stock for crystallizing the mutants, in a drop with a volume ratio of 3:1:2 protein: seed stock: well solution over 26% (w/v) PEG 4000, 0.1M sodium acetate, pH 6.0 for *Cgr*AAO-Y334W, and 25% (w/v) PEG 3350, 0.1 M sodium acetate, pH 4.9 for *Cgr*AAO-Y334F. *Cgr*AAO-WT was also co-crystallized with 10 – 20 mM benzylmercaptan in an attempt to obtain a structure of a complex. The crystals were harvested via a cryoprotectant solution comprised of the well solution supplemented with 20% (v/v) glycerol, into liquid nitrogen using Hampton Research™ CryoLoops.

Data collection and structure refinement

X-ray data (Table 2) were collected at the Diamond Light Source on beamline io4 for the wild type crystal, and beamline io4-1 for the variants and the *apo*-enzyme from the attempted co-crystallization with benzylmercaptan. Data were processed using *DIALS*[99] for *Cgr*AAO-WT, *Cgr*AAO-Y334W, and *apo-Cgr*AAO-WT; and XDS[100] for *Cgr*AAO-Y334F; followed by data reduction using AIMLESS[101]. The *Cgr*AAO-WT structure was solved by molecular replacement using MOLREP[102], using PDB entry 5C86[36] as a model, and built using BUCCANEER[103]. The structure was refined using iterative cycles of REFMAC[104] followed by manual building in COOT[105]. The mutant and *apo*-enzyme structures were solved analogously, using the structure of *Cgr*AAO-WT as a model.

Table 2. Data collection and refinement statistics

	***Cgr*AAO-WT**	***Cgr*AAO-Y334F**	***Cgr*AAO-Y334W**	***apo-Cgr*AAO-WT**
Data collection				
Space group	*I23*	*I23*	*I23*	*I23*
Cell dimensions				
a, *b*, *c* (Å)	172.1, 172.1, 172.1	171.4, 171.4, 171.4	172.0, 172.0, 172.0	172.3, 172.3, 172.3
α, β, γ (°)	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0	90.0, 90.0, 90.0
Resolution (Å)	60.92-2.30 (2.38-2.30)*	121.20-2.60 (2.72-2.60)	121.65-2.65, (2.78-2.65)	121.85-1.90 (1.94-1.90)
R_{sym} or R_{merge}	0.107 (1.033)	0.174 (3.321)	0.215 (2.692)	0.143 (2.646)
R_{pim}	0.023 (0.216)	0.039 (0.750)	0.045 (0.581)	0.03 (0.549)
$CC_{1/2}$	0.999 (0.612)	0.998 (0.587)	0.998 (0.535)	0.999 (0.631)
$I / \sigma I$	19.2 (3.3)	13.2 (1.2)	12.0 (1.3)	13.9 (1.5)
Completeness (%)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)	100.0 (100.0)
Redundancy	24.4 (24.8)	21.4 (21.7)	24.2 (23.3)	24.4 (25.1)
Refinement				
No. reflections	35829	24557	23451	63444
R_{work} / R_{free}	0.24/0.29	0.22/0.28	0.22/0.28	0.18/0.22
No. atoms				
Protein	4527	4545	4587	4997
Ligand/ion	36	11	22	32
Water	51	12	11	334
B-factors (Å^2)				
Protein	56	63	59	38
Ligand/ion	82	85	69	49
Water	41	51	48	42
R.m.s. deviations				
Bond lengths (Å)	0.009	0.009	0.008	0.010
Bond angles (°)	1.750	1.693	1.666	1.639
Ramachandran plot residues				
In most favorable regions (%)	93.6	94.4	95.9	96.4
In allowed regions (%)	4.9	4.4	3.4	3.0
PDB code	6RYV	6RYW	6RYX	6STX

* Number in parentheses is value for highest resolution shell.

ASSOCIATED CONTENT

Supporting Information.

The Supporting Information is available free of charge.

PCR primers for mutagenesis, EPR spin Hamiltonian parameters, initial activity screens values, comparison of catalytic parameters, sequence analysis, SDS-PAGE gels, pH-rate profiles, temperature stability, enzyme initial-rate kinetics graphs, and product analysis or raffinose oxidation by MALDI-TOF, aldehyde detection bu Purpald, EPR spectra, NMR spectra of HMF oxidation by *Cgr*CRAAO variants (PDF).

AUTHOR INFO

Corresponding Author

*Harry Brumer: brumer@msl.ubc.ca

ORCID

0000-0002-0101-862X

Author contributions

Y.M. performed bioinformatics, cloning, mutagenesis, recombinant protein production, biochemistry, enzyme kinetic and product analysis, all corresponding data analysis, and drafted the manuscript. W.A.O. solved, refined, and interpreted protein crystal structures. S.M.F. developed methodology and performed product analysis by NMR. L.C. and P.H.W. conceived and carried out the EPR spectroscopy study. A.H.V assisted with phylogeny and sequence similarity

network analysis. E.B. crystallized proteins and tested crystal diffraction. B.H. curated and assisted with sequence data collation. B.H., P.H.W., G.J.D., and H.B. coordinated research, analyzed data, and revised the manuscript.

Funding Sources

Work in Vancouver was supported by grants to H.B. from the Natural Sciences and Engineering Research Council of Canada (STPGP 493781–16, "FUNTASTIC - Fungal copper radical oxidases as new biocatalysts for the valorization of biomass carbohydrates and alcohols") and Genome Canada/Genome BC/Ontario Genomics/Genome Quebec (project #10405, "SYNBIOMICS - Functional genomics and techno-economic models for advanced biopolymer synthesis", www.synbiomics.ca). G.J.D, P.H.W. and L.C. thank the UK Biotechnology and Biological Sciences Research Council for funding (BB/L021633/1). GJD is supported by the Royal Society Ken Murray Research Professorship. We thank Diamond Light Source for granting access to beamlines IO4 and IO4-1 (proposals mx-13587 and mx-18598 respectively) for data collection, and for onsite assistance from their staff.

ABBREVIATIONS

AA5_2, Auxiliary Activity Family 5/Subfamily 2; GalOx, Galactose oxidase; AlcOx, Alcohol oxidase; HMF, 5-hydroxymethylfurfural; DFF, 2,5-diformylfuran; FFCA, 2,5-formylfurancarboxylic acid; FDCA, 2,5-furandicarboxylic acid; HMFCA, 5-hydroxymethyl-2-furancarboxylic acid; CAZy, Carbohydrate-Active enZYmes; AA, Auxiliary Activity; CRO, Copper Radical Oxidase; AAO, aryl-alcohol oxidase; SSN, sequence similarity network; ML, Maximum Likelihood; RafOx, Raffinose oxidase; CBM32, carbohydrate-binding module family 32; WSC, Wall Stress-responsive Component; WT, wild type; BMMY, Buffered complex

methanol medium; SDS-PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis, ICP-MS, Inductively coupled plasma mass spectrometry; PNGaseF, Peptide -*N*-Glycosidase F; EndoH, Endoglycosidase H; HRP, horseradish peroxidase; ABTS, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid); NMR, nuclear magnetic resonance; EPR, electron paramagnetic resonance; SOMO, singly occupied molecular orbital; SHF, superhyperfine; GMC, Glucose-methanol-choline; TEMPO, (2,2,6,6-tetramethylpiperidin-1-yl)oxidanyl; FAD, flavin adenine dinucleotide; NAD(P), Nicotinamide adenine dinucleotide (phosphate); LBLS, Luria Bertani Low Salt; YPD, yeast extract peptone dextrose; BMGY, Buffered complex glycerol medium; MALDI-TOF MS, Matrix-assisted laser desorption ionization coupled to time-of-flight mass spectrometry

REFERENCES

1. Bozell, J. J., Connecting Biomass and Petroleum Processing with a Chemical Bridge. *Science* **2010,** *329*, 522-523.

2. Dornburg, V.; Hermann, B. G.; Patel, M. K., Scenario Projections for Future Market Potentials of Biobased Bulk Chemicals. *Environ. Sci. & Technol.* **2008,** *42*, 2261-2267.

3. Himmel, M. E.; Ding, S.-Y.; Johnson, D. K.; Adney, W. S.; Nimlos, M. R.; Brady, J. W.; Foust, T. D., Biomass Recalcitrance: Engineering Plants and Enzymes for Biofuels Production. *Science* **2007,** *315*, 804-807.

4. Zhang, Y.; Zhang, J.; Su, D., 5-Hydroxymethylfurfural: A Key Intermediate for Efficient Biomass Conversion. *J. Energy Chem.* **2015,** *24*, 548-551.

5. Rosatella, A. A.; Simeonov, S. P.; Frade, R. F.; Afonso, C. A., 5-Hydroxymethylfurfural (HMF) as a Building Block Platform: Biological Properties, Synthesis and Synthetic Applications. *Green Chem*. **2011,** *13*, 754-793.

6. Amarasekara, A. S.; Green, D.; Williams, L. D., Renewable Resources Based Polymers: Synthesis and Characterization of 2, 5-Diformylfuran–Urea Resin. *Eur. Polym. J*. **2009,** *45*, 595-598.

7. Del Poeta, M.; Schell, W. A.; Dykstra, C. C.; Jones, S.; Tidwell, R. R.; Czarny, A.; Bajic, M.; Bajic, M.; Kumar, A.; Boykin, D., Structure-in Vitro Activity Relationships of Pentamidine Analogues and Dication-Substituted Bis-Benzimidazoles as New Antifungal Agents. *Antimicrob. Agents Chemother*. **1998,** *42*, 2495-2502.

8. Hopkins, K. T.; Wilson, W. D.; Bender, B. C.; McCurdy, D. R.; Hall, J. E.; Tidwell, R. R.; Kumar, A.; Bajic, M.; Boykin, D. W., Extended Aromatic Furan Amidino Derivatives as Anti-Pneumocystis Carinii Agents. *J. Med. Chem*. **1998,** *41*, 3872-3878.

9. Benahmed-Gasmi, A.; Frere, P.; Jubault, M.; Gorgues, A.; Cousseau, J.; Garrigues, B., 2, 5-Bis (1, 4-dithiafulven-6-yl) Substituted Furans, Thiophenes and N-methyl Pyrroles as Precursors for Organic Metals. *Synth. Met*. **1993,** *56*, 1751-1755.

10. Moreau, C.; Belgacem, M. N.; Gandini, A., Recent Catalytic Advances in the Chemistry of Substituted Furans from Carbohydrates and in the Ensuing Polymers. *Top. Catal*. **2004,** *27*, 11-30.

11. Zhang, Z.; Deng, K., Recent Advances in the Catalytic Synthesis of 2, 5-Furandicarboxylic Acid and its Derivatives. *ACS Catal*. **2015,** *5*, 6529-6544.

12. Hirai, H., Oligomers from Hydroxymethylfurancarboxylic Acid. *J. Macromol. Sci., Chem.* **1984,** *21*, 1165-1179.

13. Braisted, A. C.; Oslob, J. D.; Delano, W. L.; Hyde, J.; McDowell, R. S.; Waal, N.; Yu, C.; Arkin, M. R.; Raimundo, B. C., Discovery of a Potent Small Molecule IL-2 Inhibitor Through Fragment Assembly. *J. Am. Chem. Soc.* **2003,** *125*, 3714-3715.

14. Dijkman, W. P.; Fraaije, M. W., Discovery and Characterization of a 5-Hydroxymethylfurfural Oxidase from Methylovorus sp. Strain MP688. *Appl. Environ. Microbiol.* **2014,** *80*, 1082-1090.

15. Dijkman, W. P.; Groothuis, D. E.; Fraaije, M. W., Enzyme-Catalyzed Oxidation of 5-Hydroxymethylfurfural to Furan-2, 5-Dicarboxylic Acid. *Angew. Chem., Int. Ed.* **2014,** *53*, 6515-6518.

16. Karich, A.; Kleeberg, S.; Ullrich, R.; Hofrichter, M., Enzymatic Preparation of 2, 5-Furandicarboxylic Acid (FDCA)—a Substitute of Terephthalic Acid—by the Joined Action of Three Fungal Enzymes. *Microorganisms* **2018,** *6*, 5.

17. Carro, J.; Ferreira, P.; Rodríguez, L.; Prieto, A.; Serrano, A.; Balcells, B.; Ardá, A.; Jiménez-Barbero, J.; Gutiérrez, A.; Ullrich, R., 5-Hydroxymethylfurfural Conversion by Fungal Aryl-Alcohol Oxidase and Unspecific Peroxygenase. *FEBS J.* **2015,** *282*, 3218-3229.

18. Domínguez de María, P.; Guajardo, N., Biocatalytic Valorization of Furans: Opportunities for Inherently Unstable Substrates. *ChemSusChem* **2017,** *10*, 4123-4134.

19. Hu, L.; He, A.; Liu, X.; Xia, J.; Xu, J.; Zhou, S.; Xu, J., Biocatalytic Transformation of 5-Hydroxymethylfurfural Into High-Value Derivatives: Recent Advances and Future Aspects. *ACS Sustainable Chem. Eng.* **2018,** *6*, 15915-15935.

20. Carro, J.; Fernández-Fueyo, E.; Fernández-Alonso, C.; Ullrich, R.; Hofrichter, M.; Alcalde, M.; Ferreira, P.; Martínez, A. T., Self-Sustained Enzymatic Cascade for the Production of 2,5-Furandicarboxylic Acid from 5-Methoxymethylfurfural. *Biotechnol. Biofuels* **2018,** *11*, 86.

21. Kalum, L.; Morant, M. D.; Lund, H.; Jensen, J.; Lapainaite, I.; Soerensen, N. H.; Pedersen, S.; Ostergaard, L. H.; Xu, F., Enzymatic Oxidation of 5-Hydroxymethylfurfural and Derivatives Thereof. *Patent WO2014015256 A3* **2015**.

22. Dashtban, M.; Schraft, H.; Qin, W., Fungal Bioconversion of Lignocellulosic Residues; Opportunities & Perspectives. *Int. J. Biol. Sci.* **2009,** *5*, 578-595.

23. King, B. C.; Waxman, K. D.; Nenni, N. V.; Walker, L. P.; Bergstrom, G. C.; Gibson, D. M., Arsenal of Plant Cell Wall Degrading Enzymes Reflects Host Preference Among Plant Pathogenic Fungi. *Biotechnol. Biofuels* **2011,** *4*, 4.

24. Divne, C.; Stahlberg, J.; Reinikainen, T.; Ruohonen, L.; Pettersson, G.; Knowles, J.; Teeri, T. T.; Jones, T. A., The Three-Dimensional Crystal Structure of the Catalytic Core of Cellobiohydrolase I from Trichoderma reesei. *Science* **1994,** *265*, 524-528.

25. Eveleigh, D., Cellulase: a Perspective. *Philos. Trans. R. Soc., A* **1987,** *321*, 435-447.

26. Himmel, M. E.; Ruth, M. F.; Wyman, C. E., Cellulase for Commodity Products from Cellulosic Biomass. *Curr. Opin. Biotechnol.* **1999,** *10*, 358-364.

27. Coughlan, M. P.; Folan, M. A., Cellulose and Cellulases: Food for Thought, Food for the Future? *Int. J. Biochem.* **1979,** *10*, 103-108.

28. Martinez, A. T.; Ruiz-Duenas, F. J.; Camarero, S.; Serrano, A.; Linde, D.; Lund, H.; Vind, J.; Tovborg, M.; Herold-Majumdar, O. M.; Hofrichter, M., Oxidoreductases on Their Way to Industrial Biotransformations. *Biotechnol. Adv.* **2017,** *35*, 815-831.

29. Marjamaa, K.; Kruus, K., Enzyme Biotechnology in Degradation and Modification of Plant Cell wall Polymers. *Physiol. Plant.* **2018,** *164*, 106-118.

30. Yalpani, M.; Hall, L. D., Some Chemical and Analytical Aspects of Polysaccharide Modifications. II. A High-Yielding, Specific Method for the Chemical Derivatization of Galactose-Containing Polysaccharides: Oxidation With Galactose Oxidase Followed by Reductive Amination. *J. Polym. Sci., Polym. Chem. Ed.* **1982,** *20*, 3399-3420.

31. Parikka, K.; Master, E.; Tenkanen, M., Oxidation With Galactose Oxidase: Multifunctional Enzymatic Catalysis. *J. Mol. Catal. B: Enzym.* **2015,** *120*, 47-59.

32. Xu, C.; Spadiut, O.; Araújo, A. C.; Nakhai, A.; Brumer, H., Chemo-Enzymatic Assembly of Clickable Cellulose Surfaces via Multivalent Polysaccharides. *ChemSusChem* **2012,** *5*, 661-665.

33. Derikvand, F.; Yin, D. T.; Barrett, R.; Brumer, H., Cellulose-Based Biosensors for Esterase Detection. *Anal. Chem.* **2016,** *88*, 2989-2993.

34. Levasseur, A.; Drula, E.; Lombard, V.; Coutinho, P. M.; Henrissat, B., Expansion of the Enzymatic Repertoire of the CAZy Database to Integrate Auxiliary Redox Enzymes. *Biotechnol. Biofuels* **2013,** *6*, 41.

35. Kersten, P.; Cullen, D., Copper Radical Oxidases and Related Extracellular Oxidoreductases of Wood-Decay Agaricomycetes. *Fungal Genet. Biol.* **2014,** *72*, 124-130.

36. Yin, D.; Urresti, S.; Lafond, M.; Johnston, E. M.; Derikvand, F.; Ciano, L.; Berrin, J.-G.; Henrissat, B.; Walton, P. H.; Davies, G. J.; Brumer, H., Structure–Function Characterization Reveals New Catalytic Diversity in the Galactose Oxidase and Glyoxal Oxidase Family. *Nat. Commun.* **2015,** *6*, 10197.

37. Oide, S.; Tanaka, Y.; Watanabe, A.; Inui, M., 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.* **2019,** *125*, 13-20.

38. Davies, G. J.; Sinnott, M., Sorting the Diverse: The Sequence Based Classifications of Carbohydrateactive Enzymes. *Biochemist* **2008,** *30*, 26-32.

39. Paukner, R.; Staudigl, P.; Choosri, W.; Haltrich, D.; Leitner, C., Expression, Purification, and Characterization of Galactose Oxidase of Fusarium sambucinum in E. coli. *Protein Expression Purif.* **2015,** *108*, 73-79.

40. Xu, F.; Golightly, E. J.; Schneider, P.; Berka, R. M.; Brown, K. M.; Johnstone, J. A.; Baker, D. H.; Fuglsang, C. C.; Brown, S. H.; Svendsen, A.; Klotz, A. V., Expression and Characterization of a Recombinant Fusarium spp. Galactose Oxidase. *Appl. Biochem. Biotechnol.* **2000,** *88*, 23-32.

41. McPherson, M. J.; Ogel, Z. B.; Stevens, C.; Yadav, K. D.; Keen, J. N.; Knowles, P. F., Galactose Oxidase of Dactylium dendroides. Gene Cloning and Sequence Analysis. *J. Biol. Chem.* **1992,** *267*, 8146-8152.

42. Ito, N.; Phillips, S. E.; Stevens, C.; Ogel, Z. B.; McPherson, M. J.; Keen, J. N.; Yadav, K. D.; Knowles, P. F., Novel Thioether Bond Revealed by a 1.7 Å Crystal Structure of Galactose Oxidase. *Nature* **1991,** *350*, 87-90.

43. Andberg, M.; Mollerup, F.; Parikka, K.; Koutaniemi, S.; Boer, H.; Juvonen, M.; Master, E.; Tenkanen, M.; Kruus, K., Identification and Characterization of a Novel Colletotrichum graminicola Raffinose Oxidase in the AA5 Family. *Appl. Environ. Microbiol.* **2017**, AEM. 01383-17.

44. Mollerup, F.; Aumala, V.; Parikka, K.; Mathieu, Y.; Brumer, H.; Tenkanen, M.; Master, E., A Family AA5_2 Carbohydrate Oxidase from Penicillium rubens Displays Functional Overlap Across the AA5 Family. *PloS One* **2019,** *14*, e0216546.

45. Avigad, G.; Amaral, D.; Asensio, C.; Horecker, B., The D-Galactose Oxidase of Polyporus circinatus. *J. Biol. Chem.* **1962,** *237*, 2736-2743.

46. Paukner, R.; Staudigl, P.; Choosri, W.; Sygmund, C.; Halada, P.; Haltrich, D.; Leitner, C., Galactose Oxidase from Fusarium oxysporum-Expression in E. coli and P. pastoris and Biochemical Characterization. *PLoS one* **2014,** *9*, e100116.

47. Faria, C. B.; de Castro, F. F.; Martim, D. B.; Abe, C. A. L.; Prates, K. V.; de Oliveira, M. A. S.; Barbosa-Tessmann, I. P., Production of Galactose Oxidase Inside the Fusarium fujikuroi Species Complex and Recombinant Expression and Characterization of the Galactose Oxidase GaoA Protein from Fusarium subglutinans. *Mol. Biotechnol.* **2019,** *61*, 633-649.

48. Farmer, V.; Henderson, M. E.; Russell, J., Aromatic-Alcohol-Oxidase Activity in the Growth Medium of Polystictus versicolor. *Biochem. J.* **1960,** *74*, 257-262.

49. Leuthner, B.; Aichinger, C.; Oehmen, E.; Koopmann, E.; Müller, O.; Müller, P.; Kahmann, R.; Bölker, M.; Schreier, P., A H_2O_2-Producing Glyoxal Oxidase is Required for Filamentous Growth and Pathogenicity in Ustilago maydis. *Mol. Genet. Genomics* **2005,** *272*, 639-650.

50. Kersten, P. J.; Cullen, D., Cloning and Characterization of cDNA Encoding Glyoxal Oxidase, a H2O2-Producing Enzyme from the Lignin-Degrading Basidiomycete Phanerochaete chrysosporium. *Proc. Natl. Acad. Sci. U. S. A.* **1993,** *90*, 7411-7413.

51. Shannon, P.; Markiel, A.; Ozier, O.; Baliga, N. S.; Wang, J. T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T., Cytoscape: a Software Environment for Integrated Models of Biomolecular Interaction Networks. *Genome Res.* **2003,** *13*, 2498-2504.

52. O'Connell, R. J.; Thon, M. R.; Hacquard, S.; Amyotte, S. G.; Kleemann, J.; Torres, M. F.; Damm, U.; Buiate, E. A.; Epstein, L.; Alkan, N., Lifestyle Transitions in Plant Pathogenic Colletotrichum Fungi Deciphered by Genome and Transcriptome Analyses. *Nat. Genet.* **2012,** *44*, 1060-1065.

53. Abbott, D. W.; Eirín-López, J. M.; Boraston, A. B., Insight into Ligand Diversity and Novel Biological Roles for Family 32 Carbohydrate-Binding Modules. *Mol. Biol. Evol.* **2007,** *25*, 155-167.

54. Tong, S. M.; Chen, Y.; Zhu, J.; Ying, S. H.; Feng, M. G., Subcellular Localization of Five Singular WSC Domain-Containing Proteins and their Roles in Beauveria bassiana Responses to Stress Cues and Metal Ions. *Environ. Microbiol. Rep.* **2016,** *8*, 295-304.

55. Tordai, H.; Bányai, L.; Patthy, L., 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*. **1999,** *461*, 63-67.

56. Anasontzis, G. E.; Salazar Penã, M.; Spadiut, O.; Brumer, H.; Olsson, L., Effects of Temperature and Glycerol and Methanol-Feeding Profiles on the Production of Recombinant Galactose Oxidase in Pichia pastoris. *Biotechnol. Prog*. **2014,** *30*, 728-735.

57. Spadiut, O.; Olsson, L.; Brumer, H., A Comparative Summary of Expression Systems for the Recombinant Production of Galactose Oxidase. *Microb. Cell Fact*. **2010,** *9*, 68.

58. Humphreys, K. J.; Mirica, L. M.; Wang, Y.; Klinman, J. P., Galactose Oxidase as a Model for Reactivity at a Copper Superoxide Center. *J. Am. Chem. Soc*. **2009,** *131*, 4657-4663.

59. Daou, M.; Yassine, B.; Wikee, S.; Record, E.; Duprat, F.; Bertrand, E.; Faulds, C. B., Pycnoporus cinnabarinus Glyoxal Oxidases Display Differential Catalytic Efficiencies on 5-Hydroxymethylfurfural and its Oxidized Derivatives. *Fungal Biol. Biotechnol*. **2019,** *6*, 4.

60. Wang, Y.; DuBois, J. L.; Hedman, B.; Hodgson, K. O.; Stack, T. D. P., Catalytic Galactose Oxidase Models: Biomimetic Cu(II)-Phenoxyl-Radical Reactivity. *Science* **1998,** *279*, 537-540.

61. McNicholas, S.; Potterton, E.; Wilson, K. S.; Noble, M. E. M., Presenting Your Structures: the CCP4mg Molecular-Graphics Software. *Acta Crystallogr., Sect. D: Biol. Crystallogr* **2011,** *67*, 386-394.

62. Whittaker, M. M.; Whittaker, J. W., The Active Site of Galactose Oxidase. *J. Biol. Chem*. **1988,** *263*, 6074-6080.

63. Whittaker, J. W., Free Radical Catalysis by Galactose Oxidase. *Chem. Rev.* **2003,** *103*, 2347-2364.

64. Rogers, M. S.; Tyler, E. M.; Akyumani, N.; Kurtis, C. R.; Spooner, R. K.; Deacon, S. E.; Tamber, S.; Firbank, S. J.; Mahmoud, K.; Knowles, P. F., The Stacking Tryptophan of Galactose Oxidase: a Second-Coordination Sphere Residue that has Profound Effects on Tyrosyl Radical Behavior and Enzyme Catalysis. *Biochemistry* **2007,** *46*, 4606-4618.

65. Knowles, P. F.; Brown III, R. D.; Koenig, S. H.; Wang, S.; Scott, R. A.; McGuirl, M. A.; Brown, D. E.; Dooley, D. M., Spectroscopic Studies of the Active Site of Galactose Oxidase. *Inorg. Chem.* **1995,** *34*, 3895-3902.

66. Peisach, J.; Blumberg, W., Structural Implications Derived from the Analysis of Electron Paramagnetic Resonance Spectra of Natural and Artificial Copper Proteins. *Arch. Biochem. Biophys.* **1974,** *165*, 691-708.

67. Liu, J.; Wu, S.; Li, Z., Recent Advances in Enzymatic Oxidation of Alcohols. *Curr. Opin. Chem. Biol.* **2018,** *43*, 77-86.

68. Sheldon, R. A.; Brady, D., Broadening the Scope of Biocatalysis in Sustainable Organic Synthesis. *ChemSusChem* **2019,** *12*, 2859-2881.

69. Goswami, P.; Chinnadayyala, S. S. R.; Chakraborty, M.; Kumar, A. K.; Kakoti, A., An Overview on Alcohol Oxidases and their Potential Applications. *Appl. Microbiol. Biotechnol.* **2013,** *97*, 4259-4275.

70. Hu, L.; Lin, L.; Wu, Z.; Zhou, S.; Liu, S., Recent Advances in Catalytic Transformation of Biomass-Derived 5-Hydroxymethylfurfural into the Innovative Fuels and Chemicals. *Renewable Sustainable Energy Rev*. **2017,** *74*, 230-257.

71. Lv, G.; Wang, H.; Yang, Y.; Deng, T.; Chen, C.; Zhu, Y.; Hou, X., Graphene Oxide: a Convenient Metal-Free Carbocatalyst for Facilitating Aerobic Oxidation of 5-Hydroxymethylfurfural into 2, 5-Diformylfuran. *ACS Catal*. **2015,** *5*, 5636-5646.

72. Ning, L.; Liao, S.; Sun, Y.; Yu, L.; Tong, X., The Efficient Oxidation of Biomass-Derived 5-Hydroxymethyl Furfural to Produce 2, 5-Diformylfuran Over Supported Cobalt Catalysts. *Waste Biomass Valorization* **2018,** *9*, 95-101.

73. Mishra, D. K.; Cho, J. K.; Kim, Y. J., Facile Production of 2, 5-Diformylfuran from Base-Free Oxidation of 5-Hydroxymethyl Furfural Over Manganese–Cobalt Spinels Supported Ruthenium Nanoparticles. *J. Ind. Eng. Chem*. **2018,** *60*, 513-519.

74. Grasset, F. L.; Katryniok, B.; Paul, S.; Nardello-Rataj, V.; Pera-Titus, M.; Clacens, J.-M.; De Campo, F.; Dumeignil, F., Selective Oxidation of 5-Hydroxymethylfurfural to 2, 5-Diformylfuran Over Intercalated Vanadium Phosphate Oxides. *RSC Adv*. **2013,** *3*, 9942-9948.

75. Van Deurzen, M.; Van Rantwijk, F.; Sheldon, R., Chloroperoxidase-Catalyzed Oxidation of 5-Hydroxymethylfurfural. *J. Carbohydr. Chem*. **1997,** *16*, 299-309.

76. Hanke, P. D., Enzymatic Oxidation of Hydroxymethylfurfural. *Patent US8183020 B2* **2012**.

77. Qin, Y.-Z.; Li, Y.-M.; Zong, M.-H.; Wu, H.; Li, N., Enzyme-Catalyzed Selective Oxidation of 5-Hydroxymethylfurfural (HMF) and Separation of HMF and 2, 5-Diformylfuran Using Deep Eutectic Solvents. *Green Chem*. **2015,** *17*, 3718-3722.

78. Kadowaki, M.; Godoy, M.; Kumagai, P.; Costa-Filho, A.; Mort, A.; Prade, R.; Polikarpov, I., Characterization of a New Glyoxal Oxidase from the Thermophilic Fungus Myceliophthora thermophila M77: Hydrogen Peroxide Production Retained in 5-Hydroxymethylfurfural Oxidation. *Catalysts* **2018,** *8*, 476.

79. McKenna, S.; Mines, P.; Law, P.; Kovacs-Schreiner, K.; Birmingham, W.; Turner, N.; Leimkühler, S.; Carnell, A., The Continuous Oxidation of HMF to FDCA and the Immobilisation and Stabilisation of Periplasmic Aldehyde Oxidase (PaoABC). *Green Chem*. **2017,** *19*, 4660-4665.

80. Sützl, L.; Laurent, C. V.; Abrera, A. T.; Schütz, G.; Ludwig, R.; Haltrich, D., Multiplicity of Enzymatic Functions in the CAZy AA3 Family. *Appl. Microbiol. Biotechnol*. **2018,** *102*, 2477-2492.

81. Hernández-Ortega, A.; Ferreira, P.; Martínez, A. T., Fungal Aryl-Alcohol Oxidase: a Peroxide-Producing Flavoenzyme Involved in Lignin Degradation. *Appl. Microbiol. Biotechnol*. **2012,** *93*, 1395-1410.

82. Birmingham, W. R.; Turner, N. J., A Single Enzyme Oxidative "Cascade" via a Dual-Functional Galactose Oxidase. *ACS Catal*. **2018,** *8*, 4025-4032.

83. Siebum, A.; van Wijk, A.; Schoevaart, R.; Kieboom, T., Galactose Oxidase and Alcohol Oxidase: Scope and Limitations for the Enzymatic Synthesis of Aldehydes. *J. Mol. Catal. B: Enzym*. **2006,** *41*, 141-145.

84. Lippow, S. M.; Moon, T. S.; Basu, S.; Yoon, S.-H.; Li, X.; Chapman, B. A.; Robison, K.; Lipovšek, D.; Prather, K. L., Engineering Enzyme Specificity Using Computational Design of a Defined-Sequence Library. *Chem. Biol.* **2010,** *17*, 1306-1315.

85. Sun, L.; Bulter, T.; Alcalde, M.; Petrounia, I. P.; Arnold, F. H., Modification of Galactose Oxidase to Introduce Glucose 6-Oxidase Activity. *ChemBioChem* **2002,** *3*, 781-783.

86. Rannes, J. B.; Ioannou, A.; Willies, S. C.; Grogan, G.; Behrens, C.; Flitsch, S. L.; Turner, N. J., Glycoprotein Labeling Using Engineered Variants of Galactose Oxidase Obtained by Directed Evolution. *J. Am. Chem. Soc.* **2011,** *133*, 8436-8439.

87. Deacon, S. E.; Mahmoud, K.; Spooner, R. K.; Firbank, S. J.; Knowles, P. F.; Phillips, S. E.; McPherson, M. J., Enhanced Fructose Oxidase Activity in a Galactose Oxidase Variant. *ChemBioChem* **2004,** *5*, 972-979.

88. Wachter, R. M.; Branchaud, B. P., Molecular Modeling Studies on Oxidation of Hexopyranoses by Galactose Oxidase. An Active Site Topology Apparently Designed to Catalyze Radical Reactions, Either Concerted or Stepwise. *J. Am. Chem. Soc.* **1996,** *118*, 2782-2789.

89. Grigoriev, I. V.; Nikitin, R.; Haridas, S.; Kuo, A.; Ohm, R.; Otillar, R.; Riley, R.; Salamov, A.; Zhao, X.; Korzeniewski, F.; Smirnova, T.; Nordberg, H.; Dubchak, I.; Shabalov, I., MycoCosm Portal: Gearing up for 1000 Fungal Genomes. **2014,** *42*, D699-D704.

90. Grigoriev, I. V.; Cullen, D.; Goodwin, S. B.; Hibbett, D.; Jeffries, T. W.; Kubicek, C. P.; Kuske, C.; Magnuson, J. K.; Martin, F.; Spatafora, J. W.; Tsang, A.; Baker, S. E., Fueling the Future with Fungal Genomics. *Mycology* **2011,** *2*, 192-209.

91. Standley, D. M.; Katoh, K., MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013,** *30*, 772-780.

92. Miller, M. A.; Pfeiffer, W.; Schwartz, T. "Creating the CIPRES Science Gateway for Inference of Large Phylogenetic Trees" in Proceedings of the Gateway Computing Environments Workshop (GCE) 14 Nov. **2010**, 1-8.

93. Stamatakis, A., RAxML Version 8: a Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics* **2014,** *30*, 1312-1313.

94. Camacho, C.; Coulouris, G.; Avagyan, V.; Ma, N.; Papadopoulos, J.; Bealer, K.; Madden, T. L., BLAST+: Architecture and Applications. *BMC Bioinf.* **2009,** *10*, 421.

95. Petersen, T. N.; Brunak, S.; von Heijne, G.; Nielsen, H., SignalP 4.0: Discriminating Signal Peptides from Transmembrane Regions. *Nat. Methods* **2011,** *8*, 785-786.

96. Cregg, J. M.; Russell, K. A., Transformation. Humana Press: Totowa, NJ, 1998; pp 27-39.

97. Quesenberry, M.; Lee, Y., A Rapid Formaldehyde Assay Using Purpald Reagent: Application Under Periodation Conditions. *Anal. Biochem.* **1996,** *234*, 50-55.

98. Stoll, S.; Schweiger, A., EasySpin, a Comprehensive Software Package for Spectral Simulation and Analysis in EPR. *J. Magn. Reson.* **2006,** *178*, 42-55.

99. Waterman, D. G.; Winter, G.; Gildea, R. J.; Parkhurst, J. M.; Brewster, A. S.; Sauter, N. K.; Evans, G., Diffraction-Geometry Refinement in the DIALS Framework. *Acta Crystallogr., Sect. D: Struct. Biol.* **2016,** *72*, 558-575.

100. Kabsch, W., Xds. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66*, 125-132.

101. Evans, P. R.; Murshudov, G. N., How Good are my Data and what is the Resolution? *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013,** *69*, 1204-1214.

102. Vagin, A.; Teplyakov, A., MOLREP: an Automated Program for Molecular Replacement. *J. Appl. Crystallogr.* **1997,** *30*, 1022-1025.

103. Cowtan, K., The Buccaneer Software for Automated Model Building. 1. Tracing Protein Chains. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2006,** *62*, 1002-1011.

104. Murshudov, G. N.; Skubák, P.; Lebedev, A. A.; Pannu, N. S.; Steiner, R. A.; Nicholls, R. A.; Winn, M. D.; Long, F.; Vagin, A. A., REFMAC5 for the Refinement of Macromolecular Crystal Structures. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011,** *67*, 355-367.

105. Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K., Features and Development of Coot. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *66*, 486-501.