HAL
open science

Structure–function characterization reveals new catalytic diversity in the galactose oxidase and glyoxal oxidase family

Delu (tyler) Yin, Saioa Urresti, Mickael Lafond, Esther M. Johnston, Fatemeh Derikvand, Luisa Ciano, Jean-Guy Berrin, Bernard Henrissat, Paul H. Walton, Gideon J. Davies, et al.

To cite this version:

Delu (tyler) Yin, Saioa Urresti, Mickael Lafond, Esther M. Johnston, Fatemeh Derikvand, et al.. Structure–function characterization reveals new catalytic diversity in the galactose oxidase and glyoxal oxidase family. Nature Communications, 2015, 6, pp.10197. 10.1038/ncomms10197 . hal-01475825

HAL Id: hal-01475825

https://hal.science/hal-01475825v1

Submitted on 24 Feb 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Structure-function characterization reveals new catalytic diversity in the galactose oxidase/glyoxal oxidase family

DeLu (Tyler) Yin[1,=], Saioa Urresti[2,=], Mickael Lafond[1,3], Esther M. Johnston[2], Fatemeh Derikvand[1], Luisa Ciano[2], Jean-Guy Berrin[4], Bernard Henrissat[5,6,7], Paul H. Walton[2], Gideon J. Davies[2], Harry Brumer[1,*].

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

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

[3] Aix-Marseille University, Institut des Sciences Moléculaires de Marseille – Team BiosCiences UMR 7313-CNRS, Avenue Escadrille Normandie Niemen, 13397 Marseille Cedex 20, France

[4] INRA-BCF - UMR 1163, Polytech'Marseille – Aix-Marseille Université, 163 Avenue de Luminy, 13009 Marseille, France

[5] Architecture et Fonction des Macromolécules Biologiques, CNRS - Aix-Marseille University, 163 Avenue de Luminy, 13288 Marseille, France

[6] INRA, USC 1408 AFMB, 13288 Marseille, France

[7] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[=] Equal contribution

* Author to whom correspondence should be addressed (brumer@msl.ubc.ca)

Abstract

Alcohol oxidases, including carbohydrate oxidases, have a long history of research generating fundamental biological understanding and biotechnological applications. Despite a long history of study, the galactose 6-oxidase/glyoxal oxidase family of mononuclear copper radical oxidases, Auxiliary Activity Family 5 (AA5), is currently represented by only very few characterized members. To explore the wider

biocatalytic potential in AA5, two homologues from the phytopathogenic fungi *Colletotrichum graminicola* and *Colletotrichum gloeosporioides*, *Cgr*AlcOx and *Cgl*AlcOx respectively, were recombinantly produced and subjected to detailed structure-function analyses. EPR spectroscopy and crystallographic analysis confirm a common active-site structure vis-à-vis the archetypal galactose 6-oxidase from *Fusarium graminearum*. Strikingly, however, *Cgr*AlcOx and *Cgl*AlcOx are essentially incapable of oxidizing galactose and galactosides, but instead efficiently catalyze the oxidation of diverse aliphatic alcohols of natural and industrial importance. The results highlight the significant potential of prospecting the evolutionary diversity of AA5 to reveal novel enzyme specificities, thereby informing both biology and applications.

The enzyme-catalyzed oxidation of diverse aliphatic and aromatic alcohols is a ubiquitous reaction in nature that is central to many biological processes and which forms the basis for a range of biotechnological applications [1-6]. Within the Enzyme Commission classification, alcohol oxidoreductases are primarily distinguished by the electron acceptor substrate, with the NAD^+/$NADP^+$-dependent "alcohol dehydrogenases (ADH)" (EC 1.1.1) constituting the historically richest, largest, and most diverse class [1,7] (see also http://www.chem.qmul.ac.uk/iubmb/enzyme/EC1/1/1/). Enzymes from the second most diverse class of alcohol oxidoreductases, traditionally known as "alcohol oxidases (AOX)", employ molecular oxygen as a terminal electron acceptor to generate hydrogen peroxide (EC 1.1.3). Many of these enzymes are flavoproteins that are dependent on the FAD^+ cofactor as an initial electron acceptor [1,6]. EC 1.1.3 also comprises non-FAD^+-dependent enzymes, including copper-radical oxidases (CRO) such as glyoxal oxidases (EC 1.1.3-.), hexose 1-oxidases (EC 1.1.3.5), and galactose 6-oxidases (EC 1.1.3.9) (ref. [6,8], see also http://www.chem.qmul.ac.uk/iubmb/enzyme/EC1/1/3/). Due to their lack of dependence on an organic co-factor, CROs have inspired a generation of "green" small molecule oxidation catalysts, the utility of which is the chemo-selective oxidation of alcohols using only aerobic oxygen as a co-substrate. Additionally, the capacity of these CROs to generate hydrogen peroxide is of increasing biological interest in the context of

recalcitrant plant cell wall degradation by fungal saprotrophs and phytopathogens [8,9]. Specifically, peroxide production is believed to be coupled to lignolytic peroxidases as well as non-enzymatic Fenton chemistry [8].

To catalyze the oxidation of their eponymous substrates with concomitant hydrogen peroxide production, glyoxal oxidases and galactose 6-oxidases have a conserved mononuclear copper-radical active site comprising a cross-linked cysteine-tyrosine residue together with one unmodified axial tyrosine sidechain and two histidine sidechains as additional coordinating residues [8,10]. In galactose 6-oxidase, experimental evidence suggests that the first half-reaction involves proton transfer (PT) from O-6 of galactose to the axial tyrosine anion, hydrogen atom transfer (HAT) from C-6 of galactose to the tyrosine-cysteine radical cofactor, and electron transfer (ET) from the carbohydrate to generate the aldehyde and Cu(I). To complete the catalytic cycle, the second half-reaction is proposed to involve inner-sphere ET from Cu(I) to oxygen to yield superoxide, HAT from the phenolic hydroxyl group of the Tyr-Cys cofactor by the superoxide to produce metal-bound hydroperoxide, and PT from the axial tyrosine to hydroperoxide to produce hydrogen peroxide and the re-oxidized, active Cu(II)-radical state of the enzyme (Fig. 1) [11]. These observations are reflected in the known detailed mechanisms of a range of bio-inspired small molecule copper complexes which act as catalysts for the aerobic oxidation of alcohols [12]. In contrast to galactose 6-oxidase, the mechanism of aldehyde oxidation by glyoxal oxidase has not been studied in as great detail, but it has been suggested that this enzyme may in fact react with the hydrated, *gem*-diol form of the substrate via an analogous mechanism [13,14].

As a complement to the activity-based Enzyme Commission classification, the sequence-based Carbohydrate-Active Enzymes (CAZy; www.cazy.org) classification has, since its inception in 1991, significantly augmented the mapping of structure-function relationships among enzymes involved in glycosidic bond formation and cleavage [15]. More recently, the CAZy classification has been extended to encompass a range of Auxiliary Activity (AA) families (currently 13), which comprise a diversity of redox enzymes acting on polyphenolics and carbohydrates [16]. Among these families, Auxiliary Activity Family 5 (AA5) encompasses known (methyl)glyoxal oxidases in subfamily 1 (AA5_1) and galactose 6-oxidases in subfamily 2 (AA5_2)[16] (Fig. 2). The occurrence of

AA5 family enzymes is wide: as of October 2015, AA5 members in the public CAZy database comprised approximately 90 fungal sequences in subfamilies 1 and 2, as well as a large number of bacterial AA5_2 members (*ca.* 270) primarily from *Burkholderia* sp. genome sequencing efforts (see http://www.cazy.org/AA5_subfamilies.html, ref. [16]).

Strikingly, the recent CAZy classification of AA5 reveals a paucity of sequence-correlated functional and structural information in this family. Indeed, since the initial discovery nearly 30 years ago of (methyl)glyoxal oxidase activity in the white-rot basidiomycete *Phanerochaete chrysosporium* [8], only the archetype *Pch*Glx1 [13,14,17], a strain-specific point variant [18], and a homologue from the phytopathogenic fungus *Ustilago maydis* [9] have been recombinantly produced and enzymatically characterized, while no three-dimensional structures have been solved for any AA5_1 members [8]. Studies on AA5_2 members have essentially exclusively focused on the archetypal galactose 6-oxidase from the cereal head blight ascomycete *Fusarium graminearum* (anamorph of *Gibberella zeae*), encompassing recombinant production, biochemical and kinetic analysis, spectroscopy, crystallography, enzyme engineering, and applications [5,10]. Recently, four additional gene products from other *Fusarium* species have been biochemically defined as galactose 6-oxidases [19-21], yet the number and phylogenetic diversity of characterized AA5 members remain low (Fig. 2). Intriguingly, the three-dimensional structure of a *Streptomyces lividans* AA5 member (not associated with either subfamily) involved in hyphal morphogenesis, but of unclear substrate specificity, was very recently described [22]. However, the diversity and evolution of substrate specificity in AA5 is largely uncharted [8], thereby limiting enzyme discovery for biotechnological applications [1,6].

Accordingly, a key question arises over whether the scope of activity has yet been fully defined for AA5 enzymes. This notion is given further weight from a small handful of early studies of AA5 enzymes which hint at a limited number of aliphatic alkyl alcohols being substrates for AA5 enzymes, albeit at extremely low levels of activity [23-25]. There are also several studies which show that activated alcohols, such as benzyl and allylic alcohols, are oxidized by these enzymes to yield conjugated aldehyde products[10]. Other work has demonstrated that AA5 mutants are active catalysts for the aerobic oxidation of both activated and unactivated secondary alcohols [26,27]. These studies are in

addition to early work on the wide range of galactosides which act as substrates [28,29]. The most compelling signal, however, comes from chemical studies of the suite of small molecule catalysts which incorporate the unique active site features of AA5 family enzymes, most notably the presence of a copper-coordinating redox-active ligand [12]. Several of these catalysts are capable of oxidizing unactivated aliphatic alcohols at high rates, thus indicating that wild-type AA5 members may also catalyse the oxidation of a range of alcohol substrates, which—if true—expands the biochemical role of AA5s and their potential biotechnological applications.

Originally motivated, therefore, by an interest to discover new galactose 6-oxidases with improved recombinant production, biochemical, and kinetic properties for polysaccharide modification [30-32], we selected two homologous gene products from the recently sequenced plant pathogens *Colletotrichum graminicola* M1.001 (ref. [33,34]) and *Colletotrichum gloeosporioides* Nara gc5 (ref. [35]) for functional characterization (see also http://www.colletotrichum.org/genomics/). This choice was driven, in part, by the smaller size and novel predicted domain structure of these proteins – both enzymes lack the N-terminal family 32 Carbohydrate-Binding Module (CBM32) found in the *F. graminearum* GalOx (*Fgr*GalOx) – as well as their apparent phylogenetic distance from the *Fgr*GalOx in a separate clade of AA5_2. As described hereafter, our studies demonstrate that both copper-dependent enzymes had extremely feeble activity on galactose and galactosides, but instead oxidize a range of primary alkyl alcohols (activated and unactivated) with concomitant generation of H_2O_2. Crystallography of the *C. graminicola* alcohol oxidase as the second unique structural representative of AA5 revealed a conserved overall fold and active-site vis-à-vis the archetypal *Fgr*GalOx catalytic module, while detailed comparative structural and phylogenetic analysis revealed key features in the evolution of AA5 enzymes.

These findings now extend the biochemical and structural coverage of AA5, highlight a greater catalytic diversity of AA5_2 members than had been previously appreciated, and deliver new biocatalysts for potential applications in alcohol oxidation under mild aqueous conditions without need for nucleotide cofactors and basic activators.

Results

Phylogenetic relationships of *Cgr*AlcOx and *Cgl*AlcOx in AA5

Sequence alignment of 75 AA5 members in isolation from all predicted signal peptides and other ancillary protein modules, constituting the complete cohort of fungal sequences in the public CAZy Database, resulted in phylogenetic tree revealing a clear distinction between AA5 Subfamily 1 (AA5_1) and Subfamily 2 (AA5_2) (Fig. 2A, see also http://www.cazy.org/AA5_subfamilies.html). Mapping enzyme activity data onto this phylogeny reveals that the three (methyl)glyoxal oxidases fall into AA5_1 and the five *Fusarium* spp. galactose 6-oxidases fall into a clade of AA5_2.

Also within AA5_2, the gene products of *Colletotrichum graminicola* M1.001 locus GLRG_05590.1 (UniProt E3QHV8; GenBank EFQ30446; hereafter "*Cgr*AlcOx") and *Colletotrichum gloeosporioides* Nara gc5 locus CGGC5_13025 (GenBank ELA25906; hereafter "*Cgl*AlcOx") were selected for enzyme characterization due to significant amino acid sequence differences with the archetypal galactose 6-oxidase from *F. graminearum*, *Fgr*GalOx (Genbank AAA16228.1), including the absence of the N-terminal CBM32 of *Fgr*GalOx. Amino acid sequence alignment reveals that *Cgr*AlcOx and *Cgl*AlcOx share 82% sequence identity with each other, but only 46% and 48% identity, respectively, to GalOx from *F. graminearum* (Supplementary Fig. S1). All AA5 sequences known to-date contain the key active site residues of *Fgr*GalOx namely, C228 and Y272 that combine to form the unique crosslinked thioether-tyrosyl cofactor, and Y495, H496, and H581 that also coordinate to the copper ion (ref. [36], see also Fig. 1 and S1).

Facile production of *Cgr*AlcOx and *Cgl*AlcOx by *Pichia pastoris*

The methylotrophic yeast *Pichia pastoris* was chosen as recombinant production host for both AA5 members, based on previous work with *Fgr*GalOx [30,31]. During cloning, predicted native secretion signal peptides were replaced with the *Saccharomyces cerevisiae* α-factor secretion signal for extracellular secretion; hereafter, all residue numbering refers to that of the predicted mature sequence of the native protein shown in Supplementary Fig. S1. Maximum production (30 to 40 mg/L with specific activities

toward 1-butanol of 71 and 55 U/mg for *Cgr*AlcOx and *Cgl*AlcOx, respectively) was generally observed after 3 days with induction using 1% methanol in BMMY media at 16°C in well-aerated, beveled shake flasks. As previously observed for *Fgr*GalOx production [31], the addition of copper (II) sulfate was essential to obtain maximal activity, with 0.05 mM, 0.2 mM, and 0.5 mM $CuSO_4$ all providing similar benefit, whereas 2 mM $CuSO_4$ was detrimental to production (data not shown); 0.05 mM $CuSO_4$ was used routinely. Subsequent two-step purification by immobilized metal-ion affinity and size-exclusion chromatographies yielded 35 mg/mL of *Cgr*AlcOx and 14 mg/mL of *Cgl*AlcOx with specific activities of 119 and 86 U/mg, respectively, toward the standard substrate 1-butanol. These enzyme preparations were found to be >90% pure by SDS-PAGE, and enzymatic deglycosylation indicated that both enzymes were N-glycosylated (Supplementary Fig. S2). The observation of recombinant protein glycosylation by the yeast host is consistent with analysis by the NetNGlyc server (ver. 1.0), which predicted two glycosylation sites in *Cgr*AlcOx (N58 and N186) and *Cgl*AlcOx (N58 and N191). After purification and treatment with EDTA, residual tightly bound copper was present in in the as-purified proteins, which could be detected by EPR spectroscopy (*vide infra*) and quantified by EPR spin quantitation to be 0.9±0.1 equivalents of Cu(II) per monomer.

*Cgr*AlcOx and *Cgl*AlcOx are competent primary alcohol oxidases

Based on the common membership of *Cgr*AlcOx and *Cgl*AlcOx in the same subfamily as *Fgr*GalOx, we anticipated that these enzymes would likewise catalyze the oxidation of the primary alcohol in galactose and galactosides to the corresponding aldehyde. Initial activity screening of *Cgr*AlcOx against a panel of carbohydrate substrates reveal extremely weak activity towards galactose, raffinose, and xyloglucan (< 0.3 U/mg), which are good substrates for *Fgr*GalOx [29,31]. Lactose, glucose, xylose and arabinose are similarly poor substrates, yet glycerol was oxidized with a 50-fold greater specific activity than galactose (Fig. 3). This discovery prompted us to test activity against a wider selection of alcohols (Table 1), which reveal a significantly greater (150-fold) specificity, based on k_{cat}/K_m values, of *Cgr*AlcOx for aliphatic primary alcohols such as 1-butanol compared to polyols such as glycerol, xylitol, and sorbitol. A similar trend is observed for the homologous *Cgl*AlcOx (Table 1). Indeed, plots of initial rate

kinetics versus substrate concentration are classically Michaelian for all competent substrates, except for those with apparently high K_m values, which appear as strictly linear plots (representative data for *Cgr*AlcOx is shown in Supplementary Fig. S3; similar data for *Cgl*AlcOx are not shown). 1-Butanol was selected as the benchmark substrate and was used to determine the pH-rate profile and temperature stability of *Cgr*AlcOx and *Cgl*AlcOx. Both enzymes exhibit bell-shaped pH-rate profiles, consistent with two ionizable groups involved in catalysis as proposed for *Fgr*GalOx [11], with basic pH optima (Supplementary Fig. S4). Both enzymes are significantly stable below 30°C, but rapidly lose activity above 50°C, in keeping with the mesophilic nature of their natural host (Supplementary Fig. S5). A pH of 8.0 and temperature of 23°C was used for all subsequent enzymology, including the generation of all data in Table 1.

The extensive panel of kinetic data for *Cgr*AlcOx in Table 1 reveals general substrate preferences with respect to the benchmark 1-butanol, which are mirrored in the homologue *Cgl*AlcOx. Extension of the alkyl chain to pentyl, hexyl, or heptyl has a limited effect on k_{cat} values and only slightly reduces K_m values (less than 4-fold across the series) with correspondingly little effect on k_{cat}/K_m values. On the other hand, systematic reduction in alkyl chain length from butyl to methyl greatly increases K_m values and reduces k_{cat}/K_m values, with methanol being a poor substrate (k_{cat}/K_m is 700-fold lower than for 1-butanol). It is interesting to note here that activity of the AA5 AlcOx enzymes on methanol does not significantly complicate protein production under control of the methanol-dependent AOX promoter in *Pichia pastoris* under shake-flask conditions. However, high-yield production using a continuous methanol feed in a fermenter was unsuccessful for *Cgr*AlcOx (data not shown).

Product analysis clearly indicated that 1-butanol is oxidized to butanal by *Cgr*AlcOx (Supplementary Fig. S6), however we were routinely unable to obtain substrate conversions beyond 30%. This was despite several attempts including the use of high enzyme loadings and the addition of a large excess of catalase to breakdown the peroxide co-product; the presence of horseradish peroxidase (previously shown to aid *Fgr*GalOx substrate conversion [32]) also has no effect (data not shown). The reason for this is unclear in light of the complete conversion of conjugated primary alcohols (*vide infra*).

In contrast to primary alkanols, AA5 AlcOx oxidation of the secondary alkanols 2-butanol and 2-propanol is extremely poor: Substrate saturation could not be obtained and k_{cat}/K_m values extracted from linear plots of v_o vs. [S] are at least 3500-fold lower than for 1-butanol (Table 1). On the other hand, 1,3-propanediol and 1,4-butanediol are competent substrates, the kinetics of oxidation of which are almost identical to the corresponding 1-alkanols. 1,2-propanol and the linear polyols glycerol, xylitol, and sorbitol are comparatively poor substrates, indicating that while 1,3-diol substitution is well tolerated by these AA5 AlcOxs, 1,2-diol substitution is not. Notably in all cases, 1,2-diols substrates have nearly identical k_{cat} values to 1-butanol and the other diols, but K_m values are substantially increased, yielding up to 300-fold lower specificity constants (k_{cat}/K_m values, Table 1). Product analysis using glycerol as a substrate for *Cgr*AlcOx again reveals oxidation of the primary alcohol, as observed for 1-butanol, yielding glyceraldehyde hydrate in this case (Supplementary Fig. S7). As with 1-butanol, we again were unable to effect complete conversion of the substrate despite repeated attempts, including optimization of reaction conditions.

The detrimental effect of 2-substitution is mirrored in the kinetics of (±)-2-methyl-1-butanol and (S)-(-)2-methyl-1-butanol which, despite exhibiting similar k_{cat} values to 1-butanol, has 20-fold and 125-fold higher K_m values and correspondingly poorer k_{cat}/K_m values, respectively (Table 1). Product analysis by NMR indicates that (±)-2-methyl-1-butanol is converted to the corresponding aldehyde (Supplementary Fig. S8), yet as for 1-butanol and glyceraldehyde, conversion levels >30% could not be obtained. The observed 6-fold difference in k_{cat}/K_m values suggests enantioselectivity for (R)-2-methyl-1-butanol (commercially unavailable), which is also indicated by chiral GC/MS analysis of the products, yet the low substrate conversion prevents quantitative determination of an enantiomeric-excess value. The hydroxy-amino acid serine is not detectibly oxidized by either AlcOx, and the buffer TRIS is only weakly oxidized by *Cgr*AlcOx (Table 1).

Benzyl alcohol is an excellent substrate for both AA5 AlcOx enzymes, exhibiting kinetic constants similar to 1-butanol (Table 1). In contrast to the 1-alkanols, oxidation of benzyl alcohol to benzaldehyde by *Cgr*AlcOx is essentially complete (Supplementary Fig. S9). Moreover, 4-hydroxlyation, 4-methoxylation, and 2,3-dimethoxylation of the

aromatic ring abolishes activity below the limit of detection of the assay at high enzyme loadings (200 ng). 2-phenyl-ethan-1-ol is a competent substrate, with kinetic constants comparable to those of 1-propanol. Homologous to benzyl alcohol, cinnamyl alcohol is an excellent substrate for both AA5 AlcOx enzymes, with *Cgr*AlcOx demonstrating an especially low K_m value (Table 1). NMR analysis indicates the facile (>99%) conversion of cinnamyl alcohol by *Cgr*AlcOx to cinnamaldehyde (Supplementary Fig. S10), which was also apparent due to the fragrance of the product. Again in analogy with benzyl alcohol, hydroxylation and methoxylation of the aromatic ring is not tolerated, such that the monolignols *p*-coumaryl alcohol, sinapyl alcohol, and coniferyl alcohol are all not detectibly oxidized by the either AA5 AlcOx at high enzyme loadings.

The oxidation of hexa-2,4-diene-1-ol by both *Cgr*AlcOx and *Cgl*AlcOx exhibits excellent kinetics, including the highest k_{cat} value of all substrates tested and a low K_m value (essentially identical to cinnamyl alcohol, Table 1), as well as near-complete substrate conversion (Supplementary Fig. S11). Notably, the k_{cat}/K_m value of *Cgr*AlcOx for hexa-2,4-diene-1-ol is 4.7-fold higher than for the corresponding 1-hexanol and over an order of magnitude higher than for 1-butanol. The hydroxylated terpene geraniol, which contains a 3-methyl substituent on the double bond, is also good substrate for *Cgr*AlcOx, with a k_{cat}/K_m value similar to those of the 1-alkanols with C_4-chains and higher. Removal of potential conjugation in the product, in the case of cis-3-hexen-1-ol, results in a substrate specificity of *Cgr*AlcOx again similar to the 1-alkanols.

Tertiary structure of *Cgr*AlcOx and metal site geometry

The 3D structure of *Cgr*AlcOx was solved using the homologous two domains of *Fgr*GalOx as the model. Surprisingly, given the difficulty of demetallating with EDTA, crystals grown in 0.1M imidazole were copper-free. A copper-containing structure was eventually obtained from different conditions augmented with a 15 day soak with powdered $CuSO_4$ in the droplets. The *apo* and Cu structures have an r.m.s.d of 0.3 Å (and all subsequent discussion uses the Cu crystal structure as a reference).

As expected, the *Cgr*AlcOx structure displays two domains composed mainly of β-sheets and connecting loops (Fig. 4A,B). The N-terminus (residues 1 to 370) displays seven "Kelch" motifs which together form a classical seven bladed β-propeller (and

which contains the active centre, *vide infra*). The C-terminal domain, comprising residues 371 to 488, forms an immunoglobulin like beta sandwich domain from which residues 418 to 430 pierce the N-terminus propeller axis as a finger, providing the apical His423 which is one of the copper coordination residues (described below). As expected for proteins with around 40-50% sequence identity, the two common domains of *Cgr*AlcOx and *Fgr*GalOx are extremely similar with an r.m.s.d of 0.98 Å over 465 equivalent Cα atoms (48% identical amino acids); the best match was with the *Fgr*GalOx W290F variant (PDB code 2EIC[37], PDBeFOLD Q score 0.64). The most obvious structural difference between *Fgr*GalOx and *Cgr*AlcOx is that the latter lacks the N-terminal family 32 carbohydrate binding module (CBM32) [36] (Fig. 4C). In other enzymes, CBM32 members are generally known to bind galactosyl residues [38], and indeed galactose binding was revealed in the seminal crystal structure of *Fgr*GalOx [36]. In light of the kinetic data, the absence of this module in both *Cgr*AlcOx and *Cgl*AlcOx is entirely consistent with the observed non-carbohydrate substrate specificity of these enzymes.

Consistent with the EPR quantitation of 0.9±0.1 equivalents of Cu(II) per protein, the crystal structures of copper-containing *Cgr*AlcOx and *Cgl*AlcOx show the presence of one mononuclear copper site. The copper coordination site of *Cgr*AlcOx is essentially as described for *Fgr*GalOx: Copper coordination is provided by Tyr120, Tyr334, His335 and His423 at the centre of the β-propeller (Fig. 5A). As found in the homologous galactose 6-oxidase, Cys73 is linked to Tyr120 (equating to Cys228 and Tyr272 in *Fgr*GalOx) via a thioether bond. The observed ligation geometry of the copper site is three coordinate and T-shaped, with short interactions with Tyr120 (Cu-O 2.1 Å), His423 (Cu-N 2.2 Å) and His335 (Cu-N 2.4 Å), while a fourth interaction with Tyr334 (Cu-O 2.8 Å) is too long to be considered a bonding interaction.

EPR analysis was performed on as-isolated *Cgr*AlcOx and *Cgl*AlcOx to determine the electronic structure of the resting oxidized state. Both proteins gave identical spectra (Supplementary Fig. S12). These show, in accord with previous EPR studies on galactose 6-oxidase enzymes [37,39,40], that the SOMO of the resting state is largely d(x^2-y^2) in character, indicating a single copper site with an axial coordination geometry ($g_z \neq g_x = g_y$) (Fig. 5B). Unusually, there is clear resolution of super-hyperfine

coupling to the coordinating nitrogen atoms, revealing a well-ordered active site with some covalence in the Cu–N bonds. The spectrum could be satisfactorily simulated using a CuN_2 spin system with $g_z = 2.29$ and $|A_z| = 178\times10^{-4}$ cm^{-1} and $g_x = g_y = 2.06$, $|A_x| = 38\times10^{-4}$ cm^{-1}, $|A_y| = 13\times10^{-4}$ cm^{-1}, placing the copper site squarely within the Type 2 definition of Peisach and Blumberg [41]. Second differential plots afforded accurate values of the super-hyperfine coupling constants to the coordinating nitrogen atoms of 14×10^{-4} cm^{-1}, in accord with the expected values for coupling to two sp^2-hybridized nitrogen atoms [42]. Previous studies on an inactive, oxidized state of *Fgr*GalOx, which contains a Cu(II) site and a 1-electron reduced (i.e. neutral) Cys-Tyr cofactor, produced a low temperature EPR spectrum very similar to that observed for as-isolated *Cgr*AlcOx and *Cgl*AlcOx (g_z 2.29 and A_z 191×10^{-4} cm^{-1}) including well-defined Cu-N coupling [37,39,43,44]. Thus, we conclude that as-isolated *Cgr*AlcOx and *Cgl*AlcOx also contain a Cu(II) site and a one-electron reduced form of the Cys-Tyr cofactor, and the active site structure observed by X-ray crystallography likely reflects this inactive state. As has been demonstrated for *Fgr*GalOx, this inactive state can be activated under typical assay conditions. We investigated the redox properties of the inactive resting state of *Cgr*AlcOx by incubating the enzyme with the reductant sodium dithionite and the oxidant potassium ferricyanide (Supplementary Fig. S13). Upon addition of dithionite (E^o +660 mV)[45] to resting *Cgr*AlcOx, a loss of ~70% of the EPR spin intensity is observed, accompanied by a color change from blue to colorless, showing that dithionite reduces the resting inactive state. However, *Cgr*AlcOx is not oxidized by ferricyanide (E^o +436 mV),[46] which contrasts previous observations on *Fgr*GalOx [39]. This difference indicates a lower potential of the Cys-Tyr cofactor in *Cgr*AlcOx compared to the archetypal GalOx, which may account for the ability of the AlcOx enzymes to oxidize unactivated alcohol substrates. The structural origin of the lower redox potential of the Cys-Tyr cofactor in *Cgr*AlcOx is discussed below.

The main difference between the *Cgr*AlcOx and *Fgr*GalOx active sites is that a key tryptophan (Trp290) in *Fgr*GalOx is replaced by a phenylalanine (Phe138) in *Cgr*AlcOx (Fig. 5C, *cf.* Supplementary Fig. S1). Previous studies of the activity of Trp290 mutants of *Fgr*GalOx have indicated a dramatic loss of oxidative capacity upon mutation, which is correlated to significantly higher K_m values for the natural substrate

galactose with the *Fgr*GalOx W290G/F mutants (~200 to 300 mM compared to 82 mM for wild type), presumably due to the loss of a hydrogen-bonding interaction between W290 and a remote hydroxyl group of the substrate. [37] The natural occurrence of Phe in *Cgr*AlcOx and *Cgl*AlcOx may help account for the very low activity of these enzymes on carbohydrate substrates. The observed K_m values of 0.2 to 10 mM for *Cgr*AlcOx with aliphatic alcohols provide an indication that a lack of hydrogen-bonding capacity in residue 290 does not compromise binding of aliphatic substrates (assuming K_m reflects K_s). Additionally, Trp290 in *Fgr*GalOx has been implicated in stabilizing the radical form of the Cys-Tyr cofactor, although substitution with either Phe (the wild-type residue in *Cgr*AlcOx) or Gly also stabilized the tyrosine radical with retention of catalytic activity, while other substitutions were detrimental to the enzyme [37]. A decrease in the stability of the oxidized radical cofactor in *Cgr*AlcOx compared to *Fgr*GalOx, due to poorer stabilization by Phe compared to Trp, would account for the lower redox potential of the inactive state of *Cgr*AlcOx (*vide supra*). This also is consistent with small molecule studies, which suggest that aromatic stacking interactions may tune the redox potential of radicals. Thus, substitution of Trp290 by Phe may enhance the ability of the enzyme to oxidize unactivated alcohols by enhancing the ability of the Cys-Tyr cofactor to accept a hydrogen atom during the HAT step [47,48]. Indeed, the apparent second order rate constants for the oxidation of longer chain primary aliphatic alcohols by *Cgr*AlcOx and *Cgl*AlcOx are of the order of 10^4 to 10^5 $M^{-1}s^{-1}$, some 10–100 times higher than that for oxidation of galactose by *Fgr*GalOx, suggesting that the active site in these enzymes is tuned for better reactivity with aliphatic substrates.

In the other direction, production of the *Cgr*AlcOx F138W variant causes an essentially complete loss of the ability to oxidize alkanols, while the ability to oxidize substrates with extended conjugation (2,4-hexadienol and cinnamyl alcohol) increases ca. 1.2-fold. At the same time, activity toward benzyl alcohol is reduced ca. 2-fold (Supplementary Fig. S15B). The precise reason for this altered specificity is unclear, and indeed the opposite mutation in *Fgr*GalOx (W290F) has been also shown to reduce turnover and increase the Michaelis constant for the primary substrate, galactose, as discussed above [37]. Overall, it is clear that this second shell coordination residue plays a central role in substrate recognition and catalysis through the presence or absence of a

hydrogen bond to the substrate and through a stacking interaction with the Cys-Tyr cofactor (Fig. 4), yet it is not primarily responsible for modulating between galactose 6-oxidase and general alcohol oxidase activity in AA5_2.

More distant from the active center, four loop regions show significant differences in position between *Fgr*GalOx and the AlcOx enzymes (Supplementary Fig. S14). Three of these are close to the copper site and were selected for chimera loop substitutions. To explore the effect of these differences on substrate specificity in AA5_2, these elements were introduced individually onto the the *Cgr*AlcOx scaffold. Unfortunately, none of the four new variants exhibits a gain of GalOx activity, although these mutations strongly affect the substrate specificity of *Cgr*AlcOx, as judged by screening specific activity at near-saturating substrate concentrations on a selection of alcohols (Supplementary Fig. S15). Of the three active-site loop variants, the replacement of Loop 3 in *Cgr*AlcOx with that of *Fgr*GalOx has remarkably little effect on catalysis, with specific activity values for all substrate tested reduced approximately equally to 50-75% of wild-type levels. Loop 1 and Loop 2 variants, on the other hand, are significantly impaired with activities on all substrates tested reduced to ca. 10% and ca. 1% of wild-type levels (Supplementary Fig. S15C-E). Although all loop variant proteins appear to be well-behaved following purification, including with regard to solubility, we are unable to rule out folding problems leading to a fraction of inactive protein as one source of the observed reductions in specific activities. Nonetheless, as for the *Cgr*AlcOx F138W variant, it is clear that none of these loops is individually responsible for the chemoselectivity of AA5_2 members. Resolving protein structure-function relationships in AA5_2 will therefore require far more extensive mutagenesis efforts, including the production of combinations of variants.

Discussion

Many *Colletotrichum* species are hemibiotrophs that cause anthracnose diseases in a wide range of plant hosts, including many important food crops. For example, anthracnose leaf blight and stalk rot of maize caused by *Colletotrichum graminicola* have been estimated to result in annual yield losses of ca. 6% in the United States alone [33].

Consequently, there has been recent, sustained interest in *Colletotrichum* genomics to understand the unique biology of specific host-pathogen interactions [33-35,49,50]. In the present work, we have utilized this wealth of sequence data and precedence from small molecule copper-containing oxidation catalysts to discover two novel alcohol oxidases from CAZy Auxiliary Activity Family 5, which provides essential biochemical data to seed bioinformatics analyses, understand host biology, and expand the range of biocatalysis for applications.

Indeed, the classification of *Cgr*AlcOx and *Cgl*AlcOx in the same subfamily as *Fgr*GalOx may result in the erroneous extrapolation of functional annotation as galactose 6-oxidases to these enzymes. For example, *Cgl*AlcOx is currently listed under the NCBI GenBank ID ELA25906.1 as a "galactose oxidase precursor." Although our enzymological data clearly demonstrate that *Cgr*AlcOx and *Cgl*AlcOx are not galactose 6-oxidases, they do share a common Kelch-repeat β-propeller fold, conserved copper and crosslinked Cys-Tyr cofactor active site, and mechanistic commonality with their closest AA5_2 relative, the archetypal *Fgr*GalOx. Caution is therefore warranted in the annotation of enzyme function on the basis of primary and tertiary structural data alone.

*Cgr*AlcOx and *Cgl*AlcOx exhibit a broad substrate range, resulting in the oxidation of saturated and unsaturated primary alcohols to the corresponding aldehydes with the concomitant production of hydrogen peroxide from molecular oxygen. Comparison with several mechanistically diverse alcohol oxidases characterized on the same substrates reveals that *Cgr*AlcOx and *Cgl*AlcOx are competent enzymes. For example, the second-order rate constants (k_{cat}/K_m) of the oxidation of benzyl alcohol, cinnamyl alcohol, and 2,4-hexadiene-1-ol by both enzymes are equivalent to or significantly greater than the FAD-dependent aryl-alcohol oxidase from *Pleurotus eryngii* [51] (Table S2). Benzyl alcohol is also known to be a substrate for *Fgr*GalOx [52], however *Cgr*AlcOx and *Cgl*AlcOx demonstrate 330- and 195-fold higher k_{cat}/K_m values, respectively for this aryl alcohol. *Cgr*AlcOx has a k_{cat}/K_m value for the oxidation of 1-heptanol that is a notable 3 orders of magnitude higher than for the FAD-dependent alcohol oxidase from *Aspergillus terreus* [53], while both *Cgr*AlcOx and *Cgl*AlcOx are much less efficient at methanol oxidation than an FAD-dependent *Gloeophyllum trabeum* alcohol oxidase [54] (Table S2). Although the natural alcohol substrate is currently

unresolved (e.g. long-chain alcohols representing cuticle waxes were too insoluble to be tested), we speculate that these enzymes may function in the context of the plant cell wall by generating highly diffusible reactive oxygen species for polysaccharide and lignin oxidation (especially in combination with Fenton chemistry), as has been proposed for glyoxal oxidase and other alcohol oxidases [8,54].

In sum, the kinetic data reveal a pattern of reactivity of alcohols with either *Cgr*AlcOx or *Cgl*AlcOx in which unactivated aliphatic alkyl alcohols become better substrates on increasing chain length. Activated alcohols, such as benzyl and hexa-2,4-diene-1-ol are excellent substrates, although they do not exhibit significantly higher k_{cat} values compared to unactivated alcohols. As such, further in-depth mechanistic and kinetic studies are required to elucidate the detailed mechanistic pathway of *Cgr*AlcOx and *Cgl*AlcOx. While *Cgr*AlcOx and *Cgl*AlcOx are unlikely to exhibit a vastly different overall mechanism, the observation that unactivated alkyl alcohols are uniquely competent substrates for the AlcOx enzymes vis-à-vis the archetypal *Fgr*GalOx [25] raises the possibility that the key step, canonically thought to involve hydrogen atom transfer by a coordinated Cys-Tyr radical cofactor, may differ among AA5 members.

Intriguingly, studies of small-molecule copper catalysts for the aerobic oxidation of alcohols indicate that complexes that perform this oxidative step with a coordinated Tyr radical do not oxidize unactivated alcohols. At the same time, other oxidative mechanisms are known, including direct reaction of a copper-bound oxygen species with the alcohol, or the oxidation of unactivated alcohols by copper-TEMPO catalysts in which HAT is effected by a reactive radical species *not* directly coordinated to the copper [12]. These observations regarding the reactivity of small molecule catalysts raise the possibility that the relative stability and positioning of the Cys-Tyr cofactor in the active site may be a key factor in tuning reactivity and modulating the potential range of α-C–H bonds in primary alcohols that can be cleaved by AA5 members, a point previously noted in studies of *Fgr*GalOx [37].

Broadening the substrate range of AA5 enzymes goes hand-in-hand with an extension of the sites where peroxide could be generated by these enzymes. Given the prevalence of galactosyl- and hydroxyl-terminated lipids in cell walls such a local

generation of hydrogen peroxide could then be envisaged as occurring at these sites. This, in addition to the relative efficiency of *Cgr*AlcOx and *Cgl*AlcOx, coupled with the fact that they are independent of complex cofactors such as NAD(P)H and FAD, makes them attractive from a biotechnological perspective not only in enhanced biomass degradation, but also as "green" catalysts for aerobic alcohol oxidation.

Methods

General

Unless specified otherwise, double-distilled water (ddH_2O) was used in the preparation of all buffers, substrates, and enzyme assays. Pre-packed nickel nitrilotriacetic acid (Ni-NTA) was purchased from GE Healthcare (Little Chalfont, UK). All alcohols used in enzyme kinetics and product analysis were purchased from Sigma Aldrich (St. Louis, MO) at 98% or higher purity. Peroxide contamination in commercial alcohols was determined using Quantofix, a semi-quantitative dip-stick test from Sigma Aldrich. Alcohols found to be contaminated with peroxide were passed through neutral alumina (previously activated by strong heating under flame) and re-tested. The glycosidases EndoH, $EndoH_f$ and PNGaseF were purchased from New England Biolabs (Ipswitch, MA), while α-mannosidase was purchased from Sigma Aldrich (St. Louis, MO). Proteomics grade trypsin was purchased from Promega (Madison, WI). Galactose 6-oxidase from *Fusarium graminearum* was expressed in *Pichia pastoris* (Invitrogen, Groningen, Netherlands) and purified using pre-packed Ni-NTA [31]. Bovine serum albumin and pre-cast 4-20% SDS-PAGE TGX were purchased from Bio-Rad (Hercules, CA).

Sequence alignment and phylogenetic analysis

The gene models of locus GLRG_05590.1 of *Colletotrichum graminicola* M1.001 (encoding UniProt E3QHV8; GenBank EFQ30446) and locus CGGC5_13025 of *Colletotrichum gloeosporioides* Nara gc5 (encoding GenBank ELA25906) were downloaded from their respective genome databases (refs. [33-35], accessible via http://www.colletotrichum.org/genomics/) and translated to the corresponding protein

sequences, herein referred to as *Cgr*AlcOx and *Cgl*AlcOx, respectively. Seventy-three other amino acid sequences of fungal AA5 members, from both subfamily 1 (AA5_2) and subfamily 2 (AA5_2), were extracted from the public version of the CAZy database (http://www.cazy.org/AA5.html, [16]). Where present, signal sequences and additional modules, such as carbohydrate binding modules, were removed to isolate the catalytic modules for bioinformatics analysis. A multiple sequence alignment was produced using Muscle [55]. A distance matrix based on maximum likelihood was then produced using the BLOSUM62 substitution matrix [56] and transformed into a phylogenetic tree using BioNJ [57]. Subfamilies were inferred with the assistance of Secator [58] and the tree was formatted using Dendroscope [59].

DNA cloning, transformation and selection

cDNA encoding *Cgr*AlcOx without the predicted native signal peptide (`MPTLRSALRNLPAALLALAAACEA`)[60], and including a C-terminal His_6 tag-encoding sequence, was commercially synthesized (BioBasic, Markham, Canada), cloned into pUC57, and subcloned into the *Pichia pastoris* expression vector *pPICZα-C* (Invitrogen) flush with the sequence encoding the *S. cerevisiae* α-factor secretion signal peptide (Table S1 and Fig. S1). The *CgrAlcOx* gene was first amplified by polymerase chain reaction (PCR) using *CgrAlcOx-f* and *CgrAlcOx-r* primers (Table S1), 0.2 mM dNTPs and Q5 high-fidelity DNA polymerase (New England Biolabs) with supplied buffers. The 1.7 kb PCR product was purified using a commercial anion exchange column (OMEGA bio-tek). The purified PCR product and *pPICZα-C* vector were separately digested using *XhoI* and *NotI* restriction enzymes for 1 hour at 37°C and then purified using an ion exchange column. The digested products were combined in a 2:1 insert to vector ratio and ligated using T4 DNA ligase (Lucigen, Middleton, WI) for 2 hours at 20°C. The ligated product was directly transformed into chemically competent *E. cloni 10G* (Lucigen) using the heat shock method. Insertion of the *CgrAlcOx* gene into the *pPICZα–C* vector was verified by DNA sequencing. cDNA encoding *CglAlcOx* without the predicted native signal peptide (`MVTLYSPFGSLPVALLALAALCEA`)[60], and including a C-terminal His_{10} tag-encoding sequence, was commercially synthesized in a codon-optimized form (Genescript, Piscataway, NJ), cloned directly into *pPICZα-A* using the

EcoRI and *XbaI* restriction sites flush with the sequence encoding the *S. cerevisiae* α-factor signal peptide, and transformed into chemically competent *E. coli DH5α* by heat shock.

In both cases, transformants were grown overnight at 37°C on Luria-Bertani-Low-Salt (LBLS) agar plate and selected against 25 μg mL^{-1} Zeocin™ (Invitrogen). Surviving colonies were picked and grown overnight at 37°C in 5 mL LBLS medium with 25 μg mL^{-1} Zeocin™. Plasmids from the overnight cultures were extracted using a commercial miniprep-kit (Geneaid, New Taipei City, Taiwan). The presence of the *pPICZα-C-CgrAlcOx* and *pPICZα-A-CglAlcOx* constructs in positive clones was checked by agarose (1%; w/v) gel electrophoresis and the correct insertion of both *AlcOx* genes into the corresponding vectors was verified by DNA sequencing.

Transformations of *pCgrAlcOx* and *pCglAlcOx* into *P. pastoris* X33 were initiated by digesting 2 and 11 μg of plasmid DNA for 1 hour at 37°C with *SacI* and *PmeI* (New England Biolabs), respectively. The linearized plasmids were purified using an ion exchange column (Geneaid). The digested plasmids were transformed into electrocompetent *P. pastoris* X33 cells prepared on the same day.[61] Transformants were grown on YPDS agar plates for 3 days at 30°C and selected against 100 μg mL^{-1} Zeocin™. Several colonies were streaked onto YPDS agar plates increasing Zeocin™ concentration (250, 500, 750, and 1000 μg mL^{-1}) and grown during 3 days at 30°C to select clones with multi-copy insertions.

Four colonies from each *pCgAlcOx* transformation selected from 500 or 1000 μg mL^{-1} Zeocin™ plates were inoculated into 50 mL sterile conical tubes containing 5 mL of BMGY, grown in a shaking incubator overnight at 30°C and 220 *rpm*. Cells were then pelleted by centrifugation at 1,500 × *g* for 5 min and the BMGY was decanted and replaced with 5 mL of BMMY with 1% (v/v) methanol. The culture was shaken at 200 *rpm* overnight at 18°C. Regular feeding of 1% methanol every 24 h over 3 days was performed to insure continuing protein expression. Proteins exported to the liquid medium, via targeting by the *S. cerevisiae* alpha-factor signal peptide and cleavage by the *P. pastoris Kex*2 enzyme, were separated from cells and solid particulates by centrifugation at 3,000 × *g* for 15 min. Protein production was analyzed by SDS-PAGE

and the clone that yielded the highest amount of protein was retained for large-scale production.

Site-directed Mutagenesis

Loop and point variants of *Cgl*AlcOx were generated using the QuikChange-II Site-directed Mutagenesis kit (Agilent Technologies, city, country) according to the supplier's protocol. The sequences of sense/antisense mutagenic primer pairs and the associated nucleotide replacement are listed in Table S1. The PCR products were treated with 10 units DpnI at 37°C for 1 hour and transformed into the competent *E. coli DH5*α by heat shock method. Positive transformants were selected by plating onto LBLS agar plates containing 25 µg mL^{-1} Zeocin® at 37°C overnight. Plasmids containing the correct mutation were identified by DNA sequencing. Then, the mutant plasmids were linearized with PmeI, transformed into *P. pastoris* X33 by electroporation, and positive transformants were selected as described above.

Large-scale protein production in shake-flask

Single colonies of *P. pastoris* X33 expressing AlcOx variants were individually streaked onto agar plate containing 100 µg mL^{-1} of Zeocin™ and grown for 2 days at 30°C. A single colony of each was then picked and inoculated in 10 mL of BMGY at 30°C shaken at 200 *rpm* for 8 hours. Biomass production was initiated by addition of the 10 mL culture to 1 L of BMGY medium in a 4 L beveled flask with a foam cap shaken at 30°C during 18-20 hours. Cells were harvested by centrifugation at 1,500 x *g* for 5 min under sterile conditions. The cells were quickly resuspended in 1L of BMMY containing 1% methanol and transferred to sterile 4L beveled flask with a foam cap. The culture was shaken at 200 *rpm* overnight at 16°C. Regular feeding of 1% methanol every 24 h was performed to insure continuing protein expression. After 3 days of methanol induction, proteins exported to the liquid medium were separated from cells and solid particulates by centrifugation at 3,000 x *g* during 15 min. The supernatant was quickly decanted and stored at 4°C or used for subsequent purification.

Protein purification

Liquid medium containing *Cgr*AlcOx or *Cgl*AlcOx was first raised to pH 7.5 from 6.0 by drop-wise addition of 5 M KOH while slowly stirring. The medium was allowed to equilibrate at 4°C for 6 to 18 hours, followed by removal of solids by filtration through a sterile 0.45 µm polyethersulfone (PES) membrane (Nalgene, New York, USA) that was previously rinsed with ddH_2O. Purifications were initiated by passing the filtrate through a 5 mL pre-packed Ni-NTA column, pre-equilibrated with 50 mM sodium phosphate buffer at pH 7.5 with 300 mM NaCl, at 2 mL min^{-1}. The column was washed with 70 mL of equilibration buffer (14 column volumes) at 2 mL min^{-1} and 10 mL of equilibration buffer with 25 mM imidazole at 2 mL min^{-1}. Proteins were eluted with a linear gradient from 5% to 95% of 500 mM imidazole in 50 mM sodium phosphate buffer at pH 7.5 with 300 mM NaCl at 2 mL min^{-1}. The total elution volume was 120 mL, which was collected in 1 mL fractions. Both *Cgr*AlcOx and *Cgl*AlcOx eluted at ~90 mM imidazole. Further purification and desalting of *Cgr*AlcOx and *Cgl*AlcOx were performed by size exclusion chromatography. Each time, an aliquot of 1 to 2 mL (4 to 6 $mgmL^{-1}$) of Ni-NTA-purified AlcOx was loaded onto a 16 mm × 600 mm column of Superdex 75 pre-equilibrated with 50 mM sodium phosphate at pH 7.0 at 1 mL min^{-1}. A total volume of 200 mL of equilibration buffer was applied through the column at 1 mL min^{-1}. The AlcOxs eluted at ca. 150 mL. SDS-PAGE was performed using pre-cast 4-20 % (w/v) polyacrylamide gel in the presence of 2% (w/v) SDS under reducing conditions; proteins were visualized by Coomassie blue R-250 or InVision™ His-tag In-gel Stain (Life technologies, Carlsbad, CA). The identities of the purified proteins were confirmed by tandem mass spectrometry following in-gel tryptic digests at the Proteomics Core Facility at the University of British Columbia using standard techniques. Protein concentrations of *Cgr*AlcOx (ε = 101,215 $M^{-1}cm^{-1}$) and *Cgl*AlcOx (ε = 95,715 $M^{-1}cm^{-1}$) were determined by measuring A_{280}; extinction coefficients were calculated using the ProtParam tool on the ExPASy server (http://web.expasy.org/protparam/).

Analytical protein deglycosylation

The presence of protein glycosylation on *Cgr*AlcOx and *Cgl*AlcOx was determined by treatment with endoglycosidase H from *Streptomyces plicatus* (EndoH,

New England Biolabs), endoglycosidase H_f (a recombinant protein fusion of Endo H and maltose binding protein, New England Biolabs), N-glycosidase F from *Flavobacterium meningosepticum* (PNGaseF, New England Biolabs), or α-mannosidase from *Canavalia ensiformis* (Sigma Aldrich). Deglycosylation experiments under denaturing conditions were initiated by adding 20 µg of each protein to denaturing buffer (supplied by manufacturer) and heating to 100°C for 10 min, after which samples were diluted to a final volume of 20 µL with 2 µL of 0.5 M sodium phosphate at pH 7.5 or 0.5 M sodium citrate at pH 5.5 for PNGaseF and EndoH/EndoH$_f$, respectively. Finally, 1 µL of either EndoH/EndoH$_f$ or PNGaseF was added and the samples were incubated for 1 hour at 37°C. For PNGaseF, the reaction required 1% tergitol-type NP-40 for deglycosylation activity [62]. Deglycosylation under non-denaturing conditions with α-mannosidase was performed by adding 0.1 U of this enzyme to 20 µg of AA5 AlcOx in 50 mM sodium citrate at pH 5.5. The reaction was incubated overnight at 23°C. Changes in protein mobility were analyzed by SDS-PAGE followed by staining for protein with Coomassie blue R-250 or carbohydrate with periodic acid-Schiff (PAS) using a SIGMA Glycoprotein Staining Kit®, according to the manufacturer's protocol.

pH profile and temperature stability

To determine enzyme activity across a range of pH values, two sets of stock Britton-Robinson buffers [63] were made; each contained 30 mM citrate, phosphate, and borate at pH 4.5 and 12. The low and high pH buffer sets were mixed in different proportions to obtain the desired pH values. The final concentration of buffer in the enzyme assay was 17 mM of citrate, phosphate, and borate. Enzyme activity was measured using the horseradish peroxidase (HRP) - 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulphonic acid (ABTS) coupled assay described below with 30 mM of 1-butanol.

Temperature stabilities of AA5 AlcOxs were performed by diluting concentrated stock protein into 50 mM sodium phosphate buffer at pH 7.0 to achieve a final protein concentration of 0.35 µg mL^{-1}. The diluted protein was then heated using a thermo cycler at 37, 42, 46, and 51°C. Samples of protein were taken out at different time intervals and the activity was measured using the coupled HRP-ABTS assay described below with 30 mM of 1-butanol as the substrate.

Initial rate kinetics

The kinetics of enzymatic oxidation of alcohols were followed using a colorimetric HRP-ABTS coupled assay. Oxidation of alcohols to the corresponding aldehyde by AA5 enzymes consumes 1 equivalent of O_2 and generates 1 equivalent of H_2O_2; HRP catalyzes the oxidation of ABTS to $ABTS_{ox}$ (λ_{max} 414 nm, ε 36,000 $M^{-1}cm^{-1}$) using H_2O_2. The standard assay routinely used for *Fgr*GalOx from AA5 (ref. [31]) was optimized for sensitivity and linearity (see Supplementary Methods) to contain 50 mM sodium phosphate at pH 8.0, 0.49 mM ABTS, 2.3 µM horseradish peroxidase, 30 mM alcohol and 1.14 pM purified *Cgr*AlcOx or 0.66 pM purified *Cgl*AlcOx (assay temperature 23 °C). Under these conditions, formation of $ABTS_{ox}$ gives a linear response between the 1 to 4 min that was used to derive initial rates at a range of substrate concentrations for the determination of K_m and k_{cat} values. One Unit of AlcOx activity was defined as the amount of enzyme necessary for the oxidation of 2 µmol of ABTS per min, corresponding to the consumption of 1 µmol of O_2 per min.

Enzyme product analysis by ^{1}H-NMR

AA5 AlcOx oxidation products of 1-butanol, glycerol, 2-methyl-1-butanol, benzyl alcohol, cinnamyl alcohol, and 2,4-hexa-(*E,E*)-dien-1-ol were determined by ^{1}H-NMR. Oxidation of these alcohols was initiated by addition of 45 U of *Cgr*AlcOx or *Cgl*AlcOx to a 10 mL solution of 5 to 90 mM alcohol (5 to 10 mM for 2,4-hexadien-1-ol), 50 mM sodium phosphate at pH 8.0, and 15,000 U of catalase. Solutions were shaken at 200 *rpm* for 16 hours at 23 ˚C. In parallel, a reaction without enzyme was performed as a negative control. Remaining starting materials and products were extracted by addition of 1 mL of deuterated chloroform ($CDCl_3$) to the reaction mixture in a separatory funnel (glycerol oxidation was directly analyzed in 80% D_2O buffer). Products extracted with $CDCl_3$ were dried with anhydrous sodium sulfate and immediately analyzed using a 300 MHz NMR spectrometer (Bruker).

Electron paramagnetic resonance

Pure AA5 AlcOx enzymes were incubated with 1 mM EDTA at 4°C to remove any weakly bound Cu. After 16 h, EDTA was removed by buffer exchange and the

enzyme was concentrated to 0.2 mM in 20 mM 3-morpholinopropane-1-sulfonic acid (MOPS), pH 7.0. Residual tightly bound copper was still present in the sample and the Cu site occupancy determined by EPR spin quantitation against a standard solution of $CuCl_2 \bullet 2H_2O$ was 90±10%, where the protein concentration was determined by the Bradford assay. Continuous-wave X-band frozen solution EPR spectra of individual samples of *Cgr*AlcOx and *Cgl*AlcOx were collected with and without the addition of 10% *v/v* glycerol (glycerol did not perturb the spectrum) on a Bruker EMX spectrometer operating at 155 K and 9.3 GHz. Spectral simulation was carried out using Easyspin 4.0.0 software. To investigate redox properties, samples of *Cgr*AlcOx were incubated with 20 equivalents of sodium dithionite under N_2 atmosphere or 20 equivalents of potassium ferricyanide in air for one hour, with EPR spectra collected before the addition, after 10 min, and after 1 h of incubation. Raw data files in support of this work can be accessed via the library at the University of York (DOI:10.15124/d1278660-b889-4932-8430-401884176d75).

Crystallography of *Cgr*AlcOx

Prior to crystallization, 5 mg of purified *Cgr*AlcOx were incubated with 10 µL of $EndoH_f$ (New England Biolabs) in a total volume of 5 mL in 50 mM phosphate buffer, pH 7.5 for 2 h at 20 ˚C to reduce sample complexity due to N-glycosylation. The sample was then loaded on to a 1 mL Histrap HP column (GE Healthcare), previously equilibrated in the same incubation buffer. $EndoH_f$ was recovered in the flow-through while *Cgr*AlcOx was eluted with a linear gradient of 0 to 500 mM imidazole in the same buffer. *Cgr*AlcOx was buffer-exchanged to 20 mM MOPS, pH 7.0 and concentrated up to 8 mg mL^{-1} in a 0.5 mL Amicon centrifugal filter (Millipore). Carbohydrate cleavage was confirmed by a ~8 kDa difference (compared with the native protein) in a 12% SDS-PAGE gel (data not shown).

Crystallization trials were set up with the vapor diffusion sitting-drop technique using a Mosquito robot (TTP Labtech, Melbourn, UK). Apo-protein crystals, of hexagonal appearance, were obtained after optimization and micro-seeding in a condition composed of 0.1 M imidazole, pH 8.0, 12% v/v polyethylene glycol 8.000 and 4.5% 1-butanol. Crystals were soaked in mother liquor supplemented with 40% v/v 2-methyl-2,4-

pentanediol (MPD) followed by fast freezing in liquid nitrogen. Diffraction data (Table 2) were collected at the Diamond Light Source (UK), beamline I04, at a wavelength of 0.979 Å.

Crystals of the protein, in complex with copper and of tetragonal appearance, were obtained after optimization and micro-seeding in 50% tacsimate buffer (Hampton Research, Aliso Viejo, CA, USA), at pH 6.5, followed by soaking. The crystal was soaked in a solution of 52% tacsimate buffer, pH 6.0, with the addition of copper sulfate powder to the drop. After fifteen days, the crystals were harvested using 52% tacsimate buffer, pH 6.5, 25% (v/v) ethylene glycol as cryoprotectant and immediately fast frozen in liquid nitrogen. Diffraction data (Table 2) were collected at the Diamond Light Source, beamline I03, at a wavelength of 0.979 Å.

Structure solution and refinement

All diffraction data were indexed, integrated and scaled with XDS [64] and the following processing steps were performed with the CCP4i software suite [65]: Molecular replacement was performed with MOLREP [66] using the protein atoms only, from the two conserved domains, of *Fusarium graminearum* galactose 6-oxidase as the model (2EIE [37]). *Cgr*AlcOx shares 47% identity with the two domains of this model (i.e. excluding the CBM32 of *Fgr*GalOx). The structure was rebuilt and refined with COOT [67] and REFMAC5 [68]. The two Asn-linked *N*-acetylglucosamine residues remaining from the deglycosylation step were validated with Privateer [69]. An anomalous map was built to confirm the heavy atom nature of the positive density localized in the active center (Cu, consistent with the EPR spectra, *vide supra*). For the apo-protein, 0.5% of Ramachandran outliers were detected and 94.6% of the amino acids were in the preferred region. In the case of the copper-containing crystal, 0.42% of Ramachandran outliers were detected and 95.2% of the amino acids were in the preferred regions. Structural figures were made using CCP4mg [70].

Acknowledgements

The authors would like to thank Diamond Light Source (DLS), for granting access to beamlines I03 and I04 (mx-9948) that contributed to the results presented here and

their staff for the onsite assistance. This work is funded, in part, by BBSRC grant BB/I014802/1. Work at the University of British Columbia was generously supported by the Natural Sciences and Engineering Research Council of Canada (Discovery Grant and Strategic Network Grant for the "NSERC Industrial Biocatalysis Network"), the Canada Foundation for Innovation, the British Columbia Knowledge Development Fund, and UBC faculty funding. We thank Ryan Barrett (Brumer lab, UBC) and Anthony Levasseur (INRA, Luminy, Marseille) for assistance with sequence analysis and Nicolas Lenfant (AFMB, Luminy, Marseille) for assistance with phylogeny and tree generation.

Author contributions

DLY performed protein sequence analysis, gene cloning, recombinant protein production, biochemical analysis, enzyme kinetic analysis, and enzyme product analysis on *Cgr*AlcOx. SU performed all crystallography and EPR analysis. ML performed protein sequence analysis, gene cloning, recombinant protein production, biochemical analysis, enzyme kinetic analysis on *Cgl*AlcOx; and mutagenesis and analysis of *Cgr*AlcOx variants. FD assisted enzyme product analysis by NMR and provided experimental guidance. EMJ and LC assisted with EPR analyses. J-GB assisted with protein sequence analysis and provided project guidance. BH guided AA5 phylogenetic analysis. PHW guided EPR analyses. GJD guided crystallography. HB conceived the project and provided overall guidance, including enzymology. All authors contributed to the writing of the article and production of figures.

Conflict of interest statement

The authors declare no conflict of interest.

References

1 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.* **97**, 4259-4275, doi:10.1007/s00253-013-4842-9 (2013).

2 Hernandez-Ortega, A., Ferreira, P. & Martinez, A. T. Fungal aryl-alcohol oxidase: a peroxide-producing flavoenzyme involved in lignin degradation. *Appl. Microbiol. Biotechnol.* **93**, 1395-1410, doi:10.1007/s00253-011-3836-8 (2012).

3 Hollmann, F., Arends, I., Buehler, K., Schallmey, A. & Buhler, B. Enzyme-mediated oxidations for the chemist. *Green Chem.* **13**, 226-265, doi:10.1039/c0gc00595a (2011).
4 Dijkman, W. P., de Gonzalo, G., Mattevi, A. & Fraaije, M. W. Flavoprotein oxidases: classification and applications. *Appl. Microbiol. Biotechnol.* **97**, 5177-5188, doi:10.1007/s00253-013-4925-7 (2013).
5 Parikka, K., Master, E. & Tenkanen, M. Oxidation with galactose oxidase: Multifunctional enzymatic catalysis. *Journal of Molecular Catalysis B: Enzymatic* **120**, 47-59, doi:10.1016/j.molcatb.2015.06.006 (2015).
6 Pickl, M., Fuchs, M., Glueck, S. M. & Faber, K. The substrate tolerance of alcohol oxidases. *Appl. Microbiol. Biotechnol.* **99**, 6617-6642, doi:10.1007/s00253-015-6699-6 (2015).
7 de Smidt, O., du Preez, J. C. & Albertyn, J. The alcohol dehydrogenases of *Saccharomyces cerevisiae*: a comprehensive review. *FEMS Yeast Res.* **8**, 967-978, doi:10.1111/j.1567-1364.2008.00387.x (2008).
8 Kersten, P. & Cullen, D. Copper radical oxidases and related extracellular oxidoreductases of wood-decay Agaricomycetes. *Fungal Genet. Biol.* **72**, 124-130, doi:10.1016/j.fgb.2014.05.011 (2014).
9 Leuthner, B. *et al.* A H_2O_2-producing glyoxal oxidase is required for filamentous growth and pathogenicity in *Ustilago maydis*. *Mol. Genet. Genomics* **272**, 639-650, doi:10.1007/s00438-004-1085-6 (2005).
10 Whittaker, J. W. Free radical catalysis by galactose oxidase. *Chem. Rev.* **103**, 2347-2363, doi:10.1021/cr020425z (2003).
11 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.* **131**, 4657-4663, doi:10.1021/ja807963e (2009).
12 Allen, S. E., Walvoord, R. R., Padilla-Salinas, R. & Kozlowski, M. C. Aerobic Copper-Catalyzed Organic Reactions. *Chem. Rev.* **113**, 6234-6458, doi:10.1021/cr300527g (2013).
13 Whittaker, M. M., Kersten, P. J., Cullen, D. & Whittaker, J. W. Identification of catalytic residues in glyoxal oxidase by targeted mutagenesis. *J Biol Chem* **274**, 36226-36232, doi:10.1074/jbc.274.51.36226 (1999).
14 Whittaker, M. M. *et al.* Glyoxal oxidase from *Phanerochaete chrysosporium* is a new radical-copper oxidase. *J Biol Chem* **271**, 681-687 (1996).
15 Davies, G. J. & Sinnott, M. L. Sorting the diverse: the sequence-based classifications of carbohydrate-active enzymes. *Biochem. J.*, DOI: 10.1042/BJ20080382 (online only) (2008).
16 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* **6**, 14, doi:10.1186/1754-6834-6-41 (2013).
17 Kersten, P. J. & Cullen, D. Cloning and characterization of a cDNA encoding glyoxal oxidase, a H2O2-producing enzyme from the lignin-degrading basidiomycete *Phanerochaete chrysosporium*. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 7411-7413, doi:10.1073/pnas.90.15.7411 (1993).
18 Yamada, Y., Wang, J. Q., Kawagishi, H. & Hirai, H. Improvement of ligninolytic properties by recombinant expression of glyoxal oxidase gene in hyper lignin-

degrading fungus *Phanerochaete sordida* YK-624. *Biosci. Biotechnol. Biochem.* **78**, 2128-2133, doi:10.1080/09168451.2014.946398 (2014).
19 Paukner, R., Staudigl, P., Choosri, W., Haltrich, D. & Leitner, C. Expression, purification, and characterization of galactose oxidase of *Fusarium sambucinum* in *E. coli*. *Protein Expr Purif* **108**, 73-79, doi:10.1016/j.pep.2014.12.010 (2015).
20 Paukner, R. *et al.* Galactose Oxidase from *Fusarium oxysporum* - Expression in *E. coli* and *P. pastoris* and Biochemical Characterization. *PLoS One* **9**, doi:10.1371/journal.pone.0100116 (2014).
21 Cordeiro, F. A., Faria, C. B. & Barbosa-Tessmann, I. P. Identification of new galactose oxidase genes in *Fusarium* spp. *J. Basic Microbiol.* **50**, 527-537, doi:10.1002/jobm.201000078 (2010).
22 Chaplin, A. K. *et al.* GlxA is a new structural member of the radical copper oxidase family and is required for glycan deposition at hyphal tips and morphogenesis of *Streptomyces lividans*. *Biochem. J.* **469**, 433-444, doi:10.1042/bj20150190 (2015).
23 Hamilton, G., DeJersey, J. & Adolf, P. in *Oxidases and Related Redox Systems Vol. 1* (eds T. King, H. Mason, & M. Morrison) 103-124 (University Park Press, 1973).
24 Whittaker, J. W. in *Copper-Containing Molecules* (eds J.S. Valentine & E.B. Gralla) 1-49 (Academic Press, 2002).
25 Siebum, A., van Wijk, A., Schoevaart, R. & Kieboom, T. Galactose oxidase and alcohol oxidase: Scope and limitations for the enzymatic synthesis of aldehydes. *Journal of Molecular Catalysis B-Enzymatic* **41**, 141-145, doi:10.1016/j.molcatb.2006.04.003 (2006).
26 Escalettes, F. & Turner, N. J. Directed evolution of galactose oxidase: Generation of enantioselective secondary alcohol oxidases. *ChemBioChem* **9**, 857-860, doi:10.1002/cbic.200700689 (2008).
27 Arnold, F. H., Sun, L. H. & Petrounia, I. P. Glucose 6-oxidases. Patent US7220563 B2 (2007).
28 Avigad, G., Amaral, D., Asensio, C. & Horecker, B. L. The D-galactose oxidase of *Polyporus circinatus*. *J Biol Chem* **237**, 2736-2743 (1962).
29 Xu, F. *et al.* Expression and characterization of a recombinant *Fusarium* spp. galactose oxidase. *Appl. Biochem. Biotechnol.* **88**, 23-32, doi:10.1385/abab:88:1-3:023 (2000).
30 Anasontzis, G. E., Pena, M. S., 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.* **30**, 728-735, doi:10.1002/btpr.1878 (2014).
31 Spadiut, O., Olsson, L. & Brumer, H. A comparative summary of expression systems for the recombinant production of galactose oxidase. *Microb. Cell. Fact.* **9**, doi:10.1186/1475-2859-9-68 (2010).
32 Xu, C. L., Spadiut, O., Araújo, A. C., Nakhai, A. & Brumer, H. Chemo-enzymatic assembly of clickable cellulose surfaces via multivalent polysaccharides. *ChemSusChem* **5**, 661-665, doi:10.1002/cssc.201100522 (2012).
33 Schliebner, I., Becher, R., Hempel, M., Deising, H. B. & Horbach, R. New gene models and alternative splicing in the maize pathogen *Colletotrichum graminicola*

revealed by RNA-Seq analysis. *BMC Genomics* **15**, 13, doi:10.1186/1471-2164-15-842 (2014).
34 O'Connell, R. J. *et al.* Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genet.* **44**, 1060-1065, doi:10.1038/ng.2372 (2012).
35 Gan, P. *et al.* Comparative genomic and transcriptomic analyses reveal the hemibiotrophic stage shift of *Colletotrichum* fungi. *New Phytol.* **197**, 1236-1249, doi:10.1111/nph.12085 (2013).
36 Ito, N. *et al.* Novel thioether bond revealed by a 1.7 A crystal structure of galactose oxidase. *Nature* **350**, 87-90, doi:10.1038/350087a0 (1991).
37 Rogers, M. S. *et al.* The stacking tryptophan of galactose oxidase: a second-coordination sphere residue that has profound effects on tyrosyl radical behavior and enzyme catalysis. *Biochemistry* **46**, 4606-4618, doi:10.1021/bi062139d (2007).
38 Ficko-Blean, E. & Boraston, A. Carbohydrate Binding Module Family 32. *CAZypedia*, Available at URL: http://www.cazypedia.org/index.php/Carbohydrate_Binding_Module_Family_32 (Accessed Aug. 2015).
39 Whittaker, M. M. & Whittaker, J. W. The active site of galactose oxidase. *J Biol Chem* **263**, 6074-6080 (1988).
40 Whittaker, M. M. & Whittaker, J. W. Cu(I)-dependent biogenesis of the galactose oxidase redox cofactor. *J Biol Chem* **278**, 22090-22101, doi:10.1074/jbc.M300112200 (2003).
41 Peisach, J. & Blumberg, W. E. Structural implications derived from the analysis of electron paramagnetic resonance spectra of natural and artificial copper proteins. *Arch Biochem Biophys* **165**, 691-708 (1974).
42 Iwaizumi, M., Kudo, T. & Kita, S. Correlation between the hyperfine coupling constants of donor nitrogens and the structures of the first coordination sphere in copper complexes as studied by nitrogen-14 ENDOR spectroscopy. *Inorg. Chem.* **25**, 1546-1550, doi:10.1021/ic00230a008 (1986).
43 Knowles, P. F. *et al.* SPECTROSCOPIC STUDIES OF THE ACTIVE-SITE OF GALACTOSE-OXIDASE. *Inorg. Chem.* **34**, 3895-3902, doi:10.1021/ic00119a010 (1995).
44 Whittaker, M. M. *et al.* Spectroscopic and magnetochemical studies on the active site copper complex in galactose oxidase. *Journal of Molecular Catalysis B-Enzymatic* **8**, 3-15, doi:10.1016/s1381-1177(99)00072-7 (2000).
45 Mayhew, S. G. The redox potential of dithionite and SO_2^- from equilibrium reactions with flavodoxins, methyl viologen and hydrogen plus hydrogenase. *European Journal of Biochemistry* **85**, 535-547 (1978).
46 Kwong, H. L. in *Encyclopedia of Reagents for Organic Synthesis* (ed L. Paquette) (J. Wiley & Sons, 2004).
47 Liu, X. M., Barrett, S. A., Kilner, C. A., Thornton-Pett, M. & Halcrow, M. A. Syntheses of new hydroxy- 3.3 orthocyclophanes as models for the galactose oxidase Tyr-Cys cofactor. *Tetrahedron* **58**, 603-611, doi:10.1016/s0040-4020(01)01175-9 (2002).

48 Liu, X. M. *et al.* An intramolecular pi-pi interaction has no effect on the lifetime of an aryl radical cation. *Chemical Communications*, 1947-1948, doi:10.1039/b006005o (2000).
49 Baroncelli, R., Sanz-Martin, J. M., Rech, G. E., Sukno, S. A. & Thon, M. R. Draft genome sequence of *Colletotrichum sublineola*, a destructive pathogen of cultivated sorghum. *Genome announcements* **2**, doi:10.1128/genomeA.00540-14 (2014).
50 Baroncelli, R., Sreenivasaprasad, S., Sukno, S. A., Thon, M. R. & Holub, E. Draft Genome Sequence of *Colletotrichum acutatum Sensu Lato* (*Colletotrichum fioriniae*). *Genome announcements* **2**, doi:10.1128/genomeA.00112-14 (2014).
51 Ferreira, P. *et al.* Spectral and catalytic properties of aryl-alcohol oxidase, a fungal flavoenzyme acting on polyunsaturated alcohols. *Biochem. J.* **389**, 731-738 (2005).
52 Whittaker, M. M. & Whittaker, J. W. Catalytic reaction profile for alcohol oxidation by galactose oxidase. *Biochemistry* **40**, 7140-7148, doi:10.1021/bi0103031 (2001).
53 Kumar, A. K. & Goswami, P. Purification and properties of a novel broad substrate specific alcohol oxidase from *Aspergillus terreus* MTCC 6324. *BBA-Proteins Proteomics* **1784**, 1552-1559, doi:10.1016/j.bbapap.2008.06.009 (2008).
54 Daniel, G. *et al.* Characteristics of *Gloeophyllum trabeum* alcohol oxidase, an extracellular source of H2O2 in brown rot decay of wood. *Appl. Environ. Microbiol.* **73**, 6241-6253, doi:10.1128/aem.00977-07 (2007).
55 Edgar, R. C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-1797, doi:10.1093/nar/gkh340 (2004).
56 Henikoff, S. & Henikoff, J. G. Amino-acid substitution matrices from protein blocks. *Proc. Natl. Acad. Sci. U. S. A.* **89**, 10915-10919, doi:10.1073/pnas.89.22.10915 (1992).
57 Gascuel, O. BIONJ: An improved version of the NJ algorithm based on a simple model of sequence data. *Molecular Biology and Evolution* **14**, 685-695 (1997).
58 Wicker, N., Perrin, G. R., Thierry, J. C. & Poch, O. Secator: A program for inferring protein subfamilies from phylogenetic trees. *Molecular Biology and Evolution* **18**, 1435-1441 (2001).
59 Huson, D. H. & Scornavacca, C. Dendroscope 3: An Interactive Tool for Rooted Phylogenetic Trees and Networks. *Syst. Biol.* **61**, 1061-1067, doi:10.1093/sysbio/sys062 (2012).
60 Petersen, T. N., Brunak, S., von Heijne, G. & Nielsen, H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**, 785-786, doi:10.1038/nmeth.1701 (2011).
61 Cregg, J. M. & Russell, K. A. in *Methods in Molecular Biology* Vol. 103 27-39 (1998).
62 Maley, F., Trimble, R. B., Tarentino, A. L. & Plummer, T. H. Characterization of glycoproteins and their associated oligosaccharides through the use of endoglycosidases. *Analytical Biochemistry* **180**, 195-204 (1989).

63 Britton, H. T. S. & Robinson, R. A. CXCVIII. Universal buffer solutions and the dissociation constant of veronal. *Journal of the Chemical Society (Resumed)*, 1456, doi:10.1039/jr9310001456 (1931).
64 Kabsch, W. *XDS*. *Acta Crystallographica Section D Biological Crystallography* **66**, 125-132, doi:10.1107/S0907444909047337 (2010).
65 Potterton, E., Briggs, P., Turkenburg, M. & Dodson, E. A graphical user interface to the CCP4 program suite. *Acta Crystallographica. Section D, Biological Crystallography* **59**, 1131-1137 (2003).
66 Vagin, A. & Teplyakov, A. Molecular replacement with *MOLREP*. *Acta Crystallographica Section D Biological Crystallography* **66**, 22-25, doi:10.1107/S0907444909042589 (2010).
67 Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of *Coot*. *Acta Crystallographica Section D Biological Crystallography* **66**, 486-501, doi:10.1107/S0907444910007493 (2010).
68 Murshudov, G. N., Vagin, A. A. & Dodson, E. J. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallographica Section D: Biological Crystallography* **53**, 240–255 (1997).
69 Agirre, J. & Cowtan, K. Validation of carbohydrate structures in CCP4 6.5. *Computational Crystallography Newsletter* **6**, 10-12 (2015).
70 McNicholas, S., Potterton, E., Wilson, K. S. & Noble, M. E. M. Presenting your structures: the *CCP* 4 *mg* molecular-graphics software. *Acta Crystallographica Section D Biological Crystallography* **67**, 386-394, doi:10.1107/S0907444911007281 (2011).
71 McPherson, M. J. *et al.* Galactose oxidase of *Dactylium dendroides* - gene cloning and sequence analysis. *J Biol Chem* **267**, 8146-8152 (1992).
72 Haynes, W. H. e. *CRC Handbook of Chemistry and Physics*. 94th edn, (CRC Press, 2013).

Figure legends

Figure 1. Abbreviated mechanism of galactose oxidase. a) The first half reaction in which the enzyme is reduced while oxidizing the alcohol to aldehyde. b) The second half reaction reduces O_2 to H_2O_2 while oxidizing copper^{1+} to the resting +2 oxidative state (reviewed in ref. [10]; see ref. [11] for a detailed mechanism). PT = proton transfer, ET = electron transfer.

Figure 2. Phylogeny of AA5. Subfamilies 1 (AA5_1) and 2 (AA5_2) are indicated. GenBank identifiers (UniProt identifier Q01745 in the case of the *Fusarium graminarium* GalOx) are given for all sequences available in the public release of the CAZy Database [16] as of June 2015; full genus and species names of source organisms can be obtained from the respective GenBank entries. AA5 members for which enzymological data has been conclusively linked to protein sequence are indicated as (methyl)glyoxyl oxidases (GlyoxOx) [9,13,14,17,18], galactose oxidases (GalOx) [19-21,29,36,37,71], or general alcohol oxidases (AlcOx) [this study].

Figure 3. Initial activity screen of *Cgr*AlcOx against carbohydrates. Saccharides and glycerol were assayed at 50 mM and xyloglucan was assayed at 1 g L^{-1}. Results shown are averages of duplicate measurements from a single experimental replicate.

Figure 4. 3-D structure of *Cgr*AlcOx. A. In the N terminus (color-ramped blue to yellow), 7 Kelch motifs enclose the copper binding site in a β-propeller arrangement whilst the C terminus displays a 9-stranded β-barrel (orange to red) B. Orthogonal view. The Cu ion is showed as a grey shaded sphere. C. Overlap of *Cgr*AlcOx (dark) and GalOx (pale, 2EIE), in divergent ("wall-eyed") stereo highlighting the additional N-terminal CBM32 domain of the latter (pale pink).

Figure 5. Active site of *Cgr*AlcOx. A. Observed electron density map for the active site, maximum likelihood (REFMAC) weighted $2F_{obs}$-F_{calc} map contoured at 0.59 $Å^3$ Da^{-1}. with the anomalous difference density map (red, $\lambda = 0.979$ Å) contoured at 2σ. Copper is coordinated by residues Y120, Y334, H335 and H423 in the center of the N-terminal β-propeller. B. X-band EPR spectrum (9.3 GHz, 155 K) of Cu-*Cgr*AlcOx (black) with

simulation (red), 10% *v/v* glycerol (the spectrum of *Cgl*AlcOx was identical, Supplementary Figure S12). C. Divergent stereoscopic view of the overlay of *Cgr*AlcOx (dark) and *Fgr*GalOx (pale) copper binding centers. The most obvious difference is that F138 in *Cgr*AlcOx is present as a Trp, W290, in GalOx.

Tables

Table 1. Substrate specificity[a] of *Cgr*AlcOx and *Cgl*AlcOx

			*Cgr*AlcOx			*Cgl*AlcOx		
	Substrate	**Solubility (mM)[b]**	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}M^{-1}$)	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($s^{-1}M^{-1}$)
primary alcohols	Methanol	miscible	n.d.[c]	n.d.[c]	$2 \times 10^2 \pm 7$	n.d.[c]	n.d.[c]	83.7 ± 2.3
	ethanol	miscible	92 ± 1	10.7 ± 0.4	8.6×10^3	23.7 ± 0.8	18.4 ± 1.2	1.29×10^3
	1-propanol	miscible	95 ± 1	2.2 ± 0.1	4.4×10^4	67.1 ± 1.5	4.27 ± 0.25	1.57×10^4
	1-butanol	1000	96 ± 1	0.68 ± 0.04	1.4×10^5	119 ± 6	3.46 ± 0.48	3.46×10^4
	1-pentanol	250	97 ± 1	0.47 ± 0.01	2.1×10^5	n.a.[d]		
	1-hexanol	59	89 ± 1	0.25 ± 0.01	3.6×10^5	n.a.[d]		
	1-heptanol	20	100 ± 1	0.19 ± 0.01	5.3×10^5	n.a.[d]		
	(+/-)-2-methyl-1-butanol	320	86 ± 2	13 ± 1	6.5×10^3	n.a.[d]		
	(S)-(-)2-methyl-1-butanol	320	91 ± 2	85 ± 1	1.1×10^3	n.a.[d]		
unsaturated primary alcohols	cis-3-hexen-1-ol	~ 20	105 ± 1	0.5 ± 0.02	2.1×10^5	n.a.[d]		
	2,4-hexadiene-1-ol	~20	117 ± 2	0.07 ± 0.004	1.7×10^6	94.6 ± 1.6	0.13 ± 0.01	7.51×10^5
	geraniol	~ 2.6	86 ± 2	0.25 ± 0.02	3.5×10^5	n.a.[d]		
secondary alcohols	2-propanol	miscible	n.d.[c]	n.d.[c]	25 ± 0.7	n.a.[d]		
	2-butanol	390	n.d.[c]	n.d.[c]	41 ± 0.7	n.d.[c]	n.d.[c]	8.45 ± 0.91
diols	1,2-propanediol	miscible	90 ± 1	67 ± 4	1.4×10^3	27.5 ± 1.1	38.5 ± 4.1	7.13×10^2
	1,3-propanediol	miscible	98 ± 1	7.3 ± 0.2	1.3×10^4	n.a.[d]		
	1,4-butanediol	miscible	97 ± 1	1.2 ± 0.04	8.3×10^4	n.a.[d]		
polyols	glycerol	miscible	96 ± 2	104 ± 4	9.2×10^2	43.8 ± 3.9	270 ± 37	1.63×10^2
	sorbitol	miscible	83 ± 2	130 ± 10	6.2×10^2	n.a.[d]		
	xylitol	miscible	86 ± 2	210 ± 20	4.2×10^2	n.a.[d]		
aryl alcohols	benzyl alcohol	400	94 ± 1	0.69 ± 0.04	1.4×10^5	106 ± 2	1.27 ± 0.08	8.31×10^4
	4-methoxy benzyl alcohol		no activity[e]			n.a.[d]		
	2,3-dimethoxy benzyl alcohol		no activity[e]			n.a.[d]		
	4-hydroxy benzyl alcohol		no activity[e]			n.a.[d]		
	2-phenyl ethanol		92 ± 2	2.0 ± 0.1	4.6×10^4	n.a.[d]		
	cinnamyl alcohol	13	93 ± 1	0.06 ± 0.003	1.6×10^6	111.5 ± 2.6	0.33 ± 0.02	3.39×10^5
	coniferyl alcohol			no activity[e]		n.a.[d]		
	sinapyl alcohol			no activity[e]		n.a.[d]		
	p-coumaryl alcohol			no activity[e]		n.a.[d]		
other alcohols	tris	miscible	n.d.[c]	n.d.[c]	1.8 ± 0.1	n.a.[d]		
	serine	480	n.d.[c]	n.d.[c]	n.d.[c]	n.a.[d]		

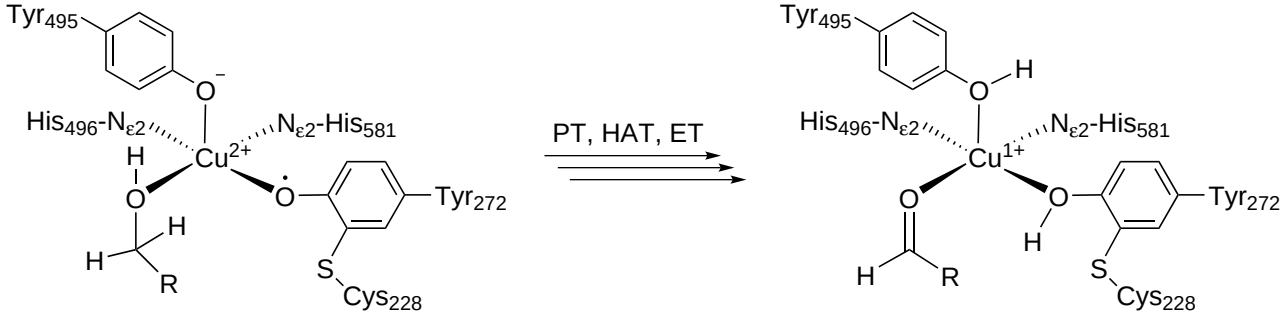
[a] The modified ABTS assay was used to measure initial rates of these alcohols. A standard 1 mL assay solution contains 50 mM sodium phosphate at pH 8.0, 0.25 mg mL^{-1} of ABTS, 0.1 mg horseradish peroxidase, 0.01 to 800 mM of substrate, and 7 ng to 1.83 µg of purified *Cgr*AlcOx or *Cgl*AlcOx. Standard errors of means from curve-fitting to triplicate measurements at each substrate concentration in v_o vs. [S] plots are given; values shown are from a single experimental replicate in all cases.

[b] Solubility data is from ref. [72]
[c] n.d. – individual k_{cat} and K_m values not determinable; k_{cat}/K_m values obtained from slope of linear v_o vs. [S] plots.
[d] n.a. – not assayed
[e] No activity detected with a specific activity limit of detection of 6 x 10^{-2} s^{-1} using 200 ng of protein, which is 28-fold higher than used in a typical kinetics assay.

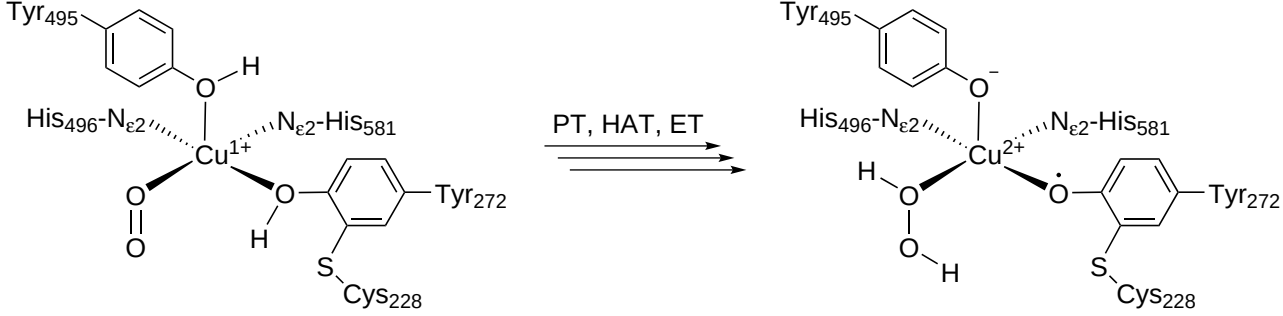
Table 2. Data collection and refinement statistics (molecular replacement)

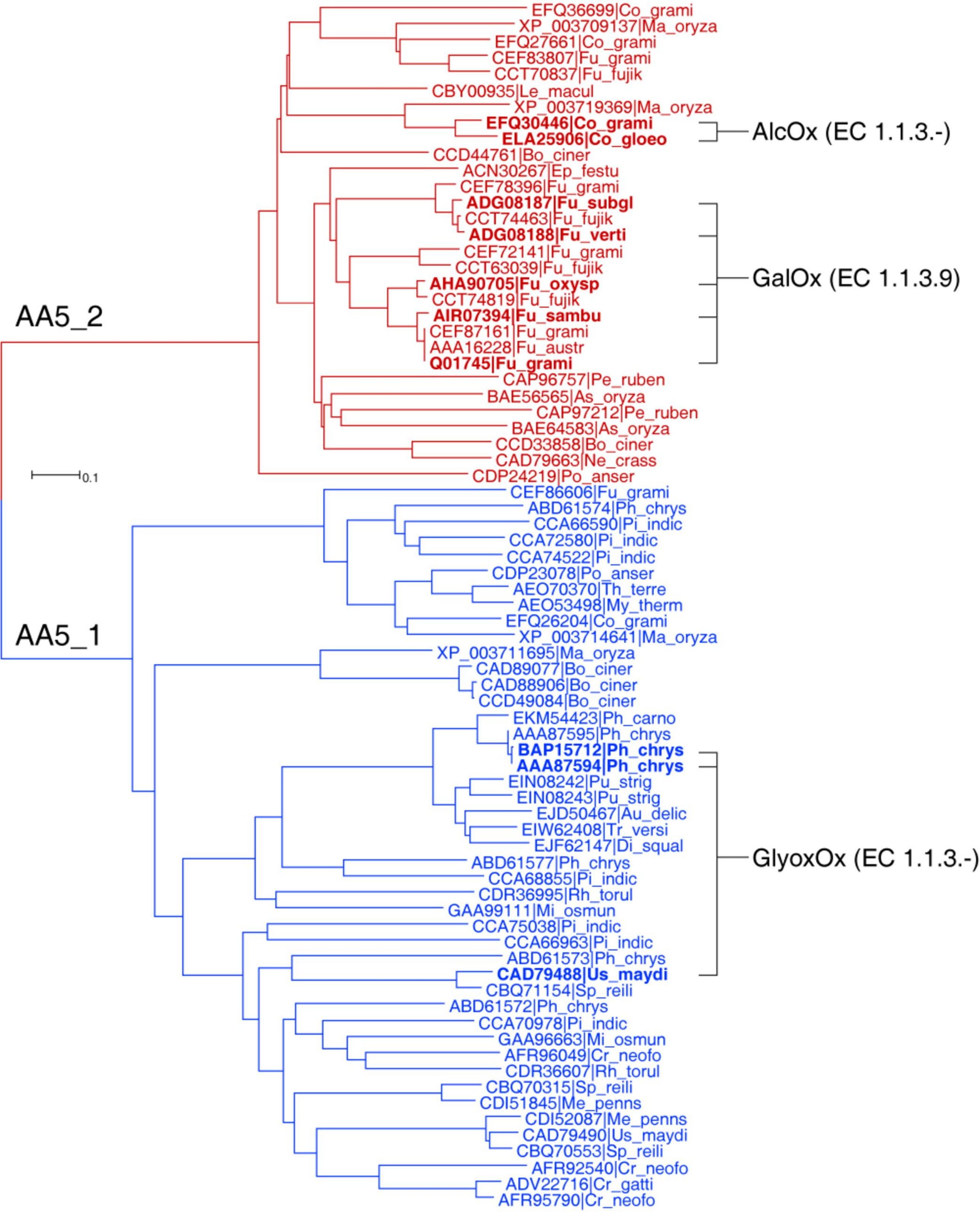
	***Apo-Cgr*AlcOx**	**Cu-*Cgr*AlcOx**
Beamline	I04	I03
PDB code	5C86	5C92
Data collection		
Space group	$P6_1$	$P4_32_12$
Cell dimensions		
a, *b*, *c* (Å)	123.7, 123.7, 52.6	81.6, 81.6, 154.3
α,β,γ (º)	90, 90, 120	90, 90, 90
Resolution (Å)	52.64 to 1.51 (1.55 to 1.51)	72.12 to 2.10 (2.15 to 2.10)
R_{merge}	0.08 (0.84)	0.1 (0.72)
R_{sym}	0.03 (0.28)	0.03 (0.21)
$I / \sigma I$	18.1 (3.5)	20.6 (4.3)
Completeness (%)	99.9 (99.8)	100 (99.9)
Redundancy	9.9 (10.0)	12.9 (13.2)
Refinement		
Resolution (Å)	1.51	2.10
No. reflections	68603	29704
R_{work} / R_{free}	0.15/0.18	0.14/0.20
No. atoms		
Protein	3767	3614
Ligand/ion		33
Water	403	224
B-factors (Å^2)		
Protein	17.96	32.45
Ligand/ion		46.93
Water	28.28	35.19
R.m.s. deviations		
Bond lengths (Å)	0.02	3.46
Bond angles (º)	2.08	4.29

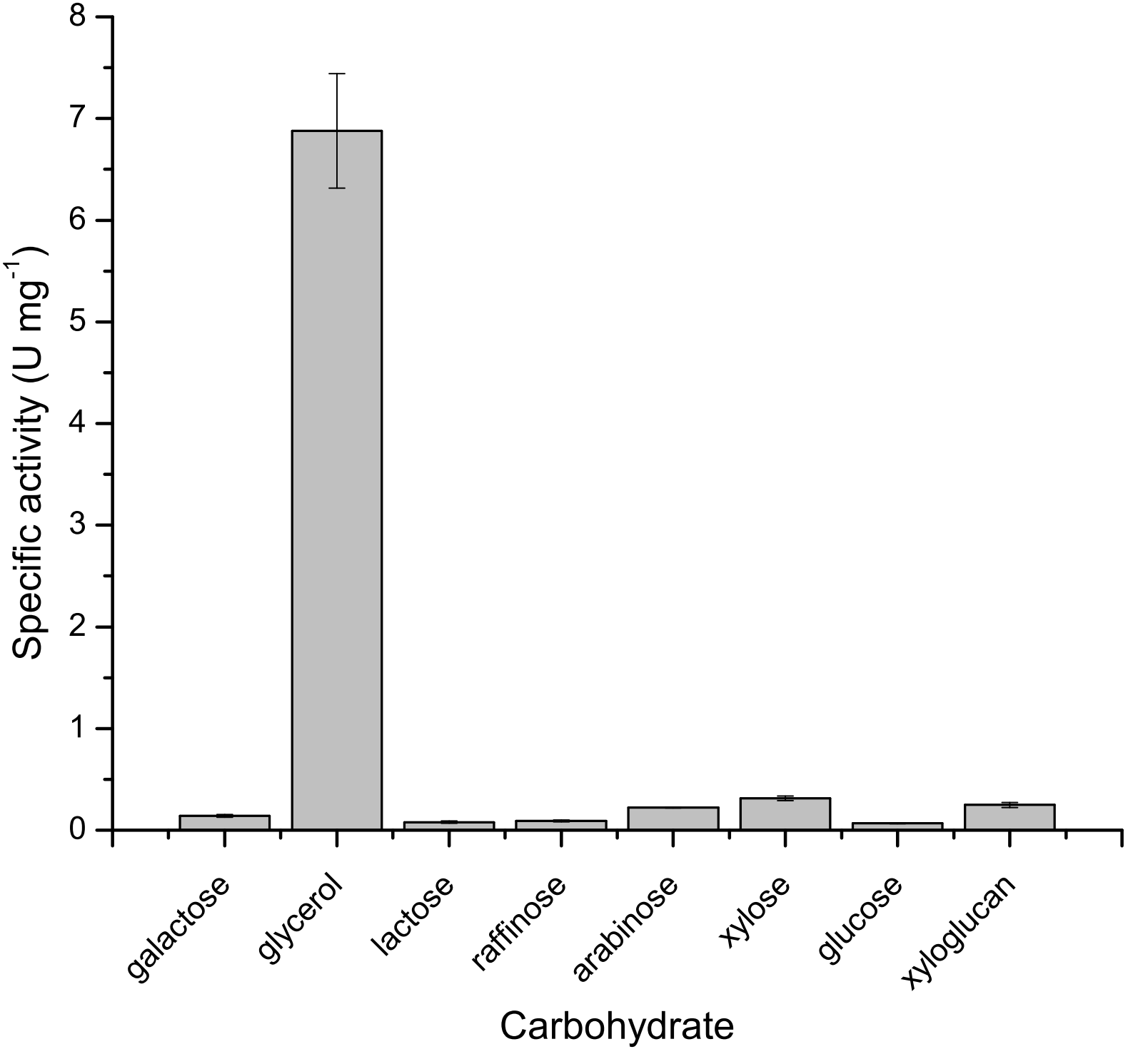
A.

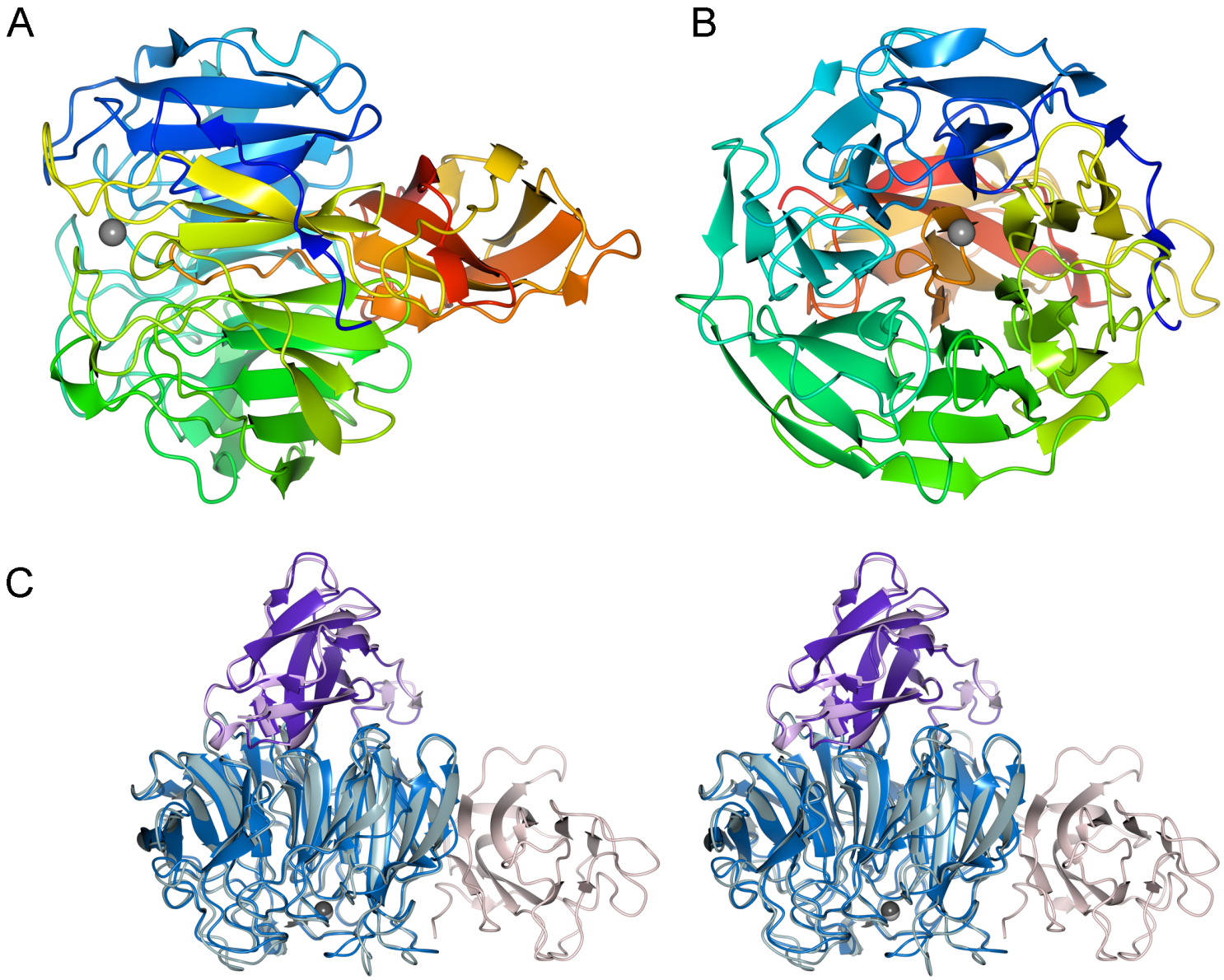

Tyr495
His496-Nε2
Nε2-His581
Cu2+
Tyr272
Cys228
PT, HAT, ET
Cu1+

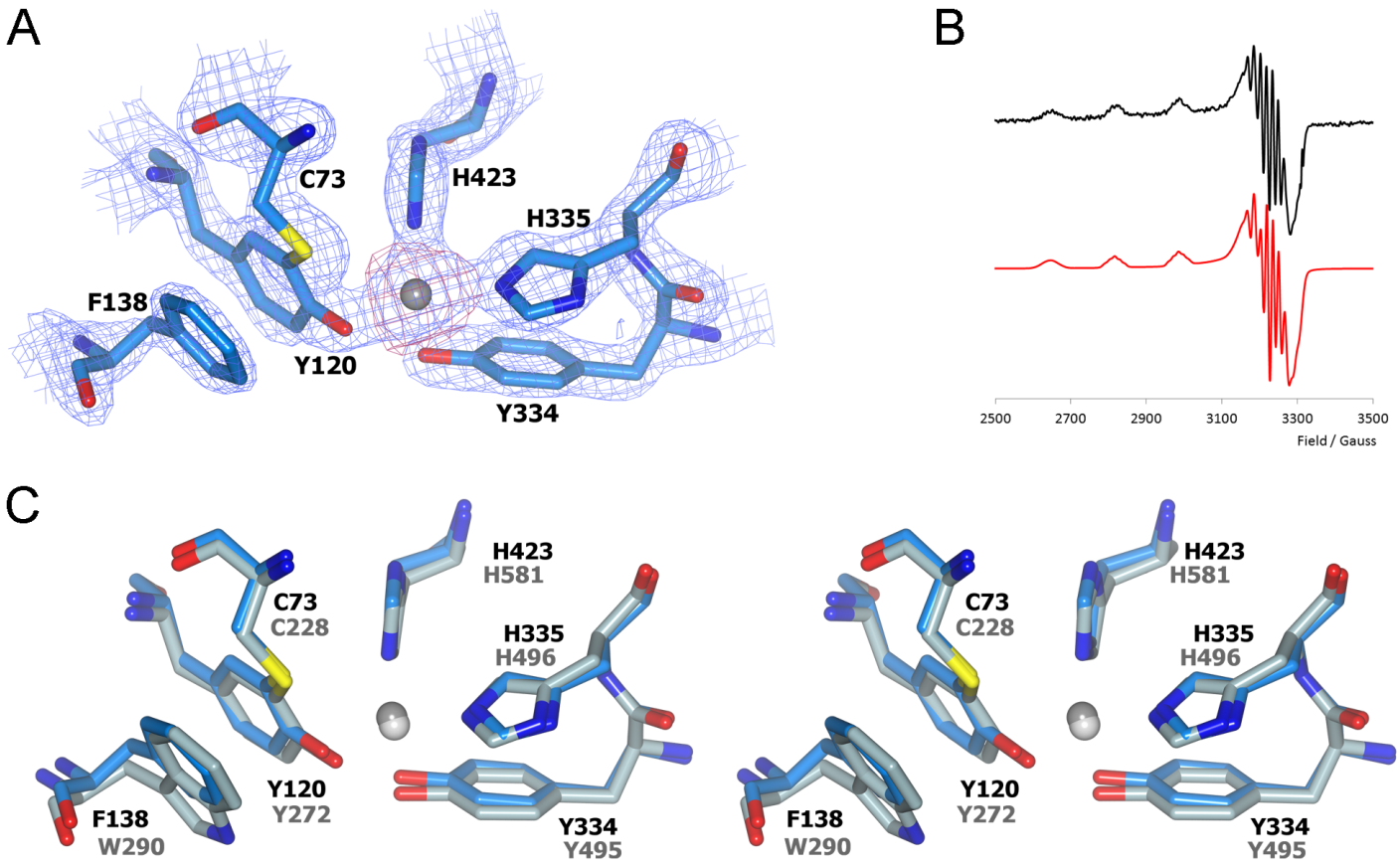
B.


Tyr495
His496-Nε2
Nε2-His581
Cu1+
Tyr272
Cys228
PT, HAT, ET
Cu2+


EFQ36699|Co_grami
XP_003709137|Ma_oryza
EFQ27661|Co_grami
CEF83807|Fu_grami
CCT70837|Fu_fujik
CBY00935|Le_macul
XP_003719369|Ma_oryza
EFQ30446|Co_grami
ELA25906|Co_gloeo
AlcOx (EC 1.1.3.-)
CCD44761|Bo_ciner
ACN30267|Ep_festu
CEF78396|Fu_grami
ADG08187|Fu_subgl
CCT74463|Fu_fujik
ADG08188|Fu_verti
CEF72141|Fu_grami
CCT63039|Fu_fujik
AHA90705|Fu_oxysp
CCT74819|Fu_fujik
AIR07394|Fu_sambu
CEF87161|Fu_grami
AAA16228|Fu_austr
Q01745|Fu_grami
GalOx (EC 1.1.3.9)
AA5_2
CAP96757|Pe_ruben
BAE56565|As_oryza
CAP97212|Pe_ruben
BAE64583|As_oryza
CCD33858|Bo_ciner
CAD79663|Ne_crass
CDP24219|Po_anser
0.1
CEF86606|Fu_grami
ABD61574|Ph_chrys
CCA66590|Pi_indic
CCA72580|Pi_indic
CCA74522|Pi_indic
CDP23078|Po_anser
AEO70370|Th_terre
AEO53498|My_therm
EFQ26204|Co_grami
XP_003714641|Ma_oryza
AA5_1
XP_003711695|Ma_oryza
CAD89077|Bo_ciner
CAD88906|Bo_ciner
CCD49084|Bo_ciner
EKM54423|Ph_carno
AAA87595|Ph_chrys
BAP15712|Ph_chrys
AAA87594|Ph_chrys
EIN08242|Pu_strig
EIN08243|Pu_strig
EJD50467|Au_delic
EIW62408|Tr_versi
EJF62147|Di_squal
ABD61577|Ph_chrys
CCA68855|Pi_indic
CDR36995|Rh_torul
GAA99111|Mi_osmun
CCA75038|Pi_indic
CCA66963|Pi_indic
ABD61573|Ph_chrys
CAD79488|Us_maydi
CBQ71154|Sp_reili
GlyoxOx (EC 1.1.3.-)
ABD61572|Ph_chrys
CCA70978|Pi_indic
GAA96663|Mi_osmun
AFR96049|Cr_neofo
CDR36607|Rh_torul
CBQ70315|Sp_reili
CDI51845|Me_penns
CDI52087|Me_penns
CAD79490|Us_maydi
CBQ70553|Sp_reili
AFR92540|Cr_neofo
ADV22716|Cr_gatti
AFR95790|Cr_neofo


Specific activity (U mg-1)
0
1
2
3
4
5
6
7
8
galactose
glycerol
lactose
raffinose
arabinose
xylose
glucose
xyloglucan
Carbohydrate


A
B
C


A
C73
H423
H335
F138
Y120
Y334
B
2500
2700
2900
3100
3300
3500
Field / Gauss
C
C73
C228
H423
H581
H335
H496
Y120
Y272
F138
W290
Y334
Y495
C73
C228
H423
H581
H335
H496
Y120
Y272
F138
W290
Y334
Y495