Oxidation of 5-hydroxymethylfurfural with a novel aryl alcohol oxidase from *Mycobacterium* sp. MS1601

Mahmoud Sayed,1,2 Yasser Gaber,3,4 Fredrik Junghus,1 Eric Valdés Martín,1,+ Sang-Hyun Pyo1 and Rajni Hatti-Kaul1,∗

1Division of Biotechnology, Department of Chemistry, Center for Chemistry and Chemical Engineering, Lund University, Lund, SE-22100, Sweden.
2Department of Botany and Microbiology, Faculty of Science, South Valley University, Qena, 83523, Egypt.
3Department of Microbiology and Immunology, Faculty of Pharmacy, Beni-Suef University, Beni-Suef, 62511, Egypt.
4Department of Pharmaceutics and Pharmaceutical Technology, Faculty of Pharmacy, Mutah University, Al-Karak, 61710, Jordan.

Summary

Bio-based 5-hydroxymethylfurfural (HMF) serves as an important platform for several chemicals, among which 2,5-furan dicarboxylic acid (FDCA) has attracted considerable interest as a monomer for the production of polyethylene furanoate (PEF), a potential alternative for fossil-based polyethylene terephthalate (PET). This study is based on the HMF oxidizing activity shown by *Mycobacterium* sp. MS 1601 cells and investigation of the enzyme catalysing the oxidation. The *Mycobacterium* whole cells oxidized the HMF to FDCA (60% yield) and hydroxymethyl furan carboxylic acid (HMCA). A gene encoding a novel bacterial aryl alcohol oxidase, hereinafter MycsPAAO, was identified in the genome and was cloned and expressed in *Escherichia coli* BL21 (DE3). The purified MycsPAAO displayed activity against several alcohols and aldehydes; 3,5 dimethoxy benzyl alcohol (veratryl alcohol) was the best substrate among those tested followed by HMF. 5-Hydroxymethylfurfural was converted to 5-formyl-2-furoic acid (FFCA) via diformyl furan (DFF) with optimal activity at pH 8 and 30–40°C. FDCA formation was observed during long reaction time with low HMF concentration. Mutagenesis of several amino acids shaping the active site and evaluation of the variants showed Y444F to have around 3-fold higher kcat/Km and ~1.7-fold lower Km with HMF.

Introduction

Oxidation of 5-hydroxymethylfurfural (HMF), a dehydration product of C6 sugars, has been a reaction of interest to provide oxidized derivatives, especially 2,5-furan dicarboxylic acid (FDCA) for use as building block for the polyester polyethylene furanoate (PEF), a biobased alternative to the widely used polyethylene terephthalate (PET). Replacement of one tonne of PET with PEF is expected to reduce the greenhouse gas emission by 30–50% (Davidson et al., 2021). Polyethylene furanoate has also the advantage of possessing better barrier, thermal and mechanical features (Sousa et al., 2015, 2016). Both HMF and FDCA are identified by the US Department of Energy to be among the important biobased platform chemicals with no negative effect on environment and human health (Werpy et al., 2004; Bozell and Petersen, 2010; Sousa et al., 2015; Chen et al., 2016; Troiano et al., 2020). Besides replacing terephthalic acid, FDCA can also replace adipic acid (Sousa et al., 2015, 2016; Zhang and Deng, 2015; Yuan et al., 2018, 2020) and can serve as a platform to produce other important chemicals such as succinic acid, 2,5-diformyl furan (DFF), 2,5-bis(hydroxymethyl)furan (BHMF) and 2,5-(bishydroxymethyl) tetrahydrofuran (Werpy et al., 2004; Zhang et al., 2015).

2,5-Furan dicarboxylic acid production from HMF involves three oxidation steps and can proceed via two different pathways (Scheme 1). In the first pathway (A), the alcohol group of HMF (1) is oxidized to an aldehyde giving 2,5-diformylfuran (DFF) (2), followed by further oxidation of the two aldehyde groups in DFF to FDCA (5) through 5-formyl-2-furoic acid (FFCA) (4), while the second pathway (B) involves oxidation of the aldehyde group of HMF to 5-hydroxymethyl-2-furoic acid (HMCA) (3), followed by further oxidation of the hydroxyl group of...
HMFCA to FDCA via FFCA (Dijkman, 2015; Dijkman et al., 2015; Troiano et al., 2020; Yuan et al., 2020). Several reports on the oxidation of HMF via chemocatalysis, (photo-)electrochemical catalysis or biocatalysis are available (Ribeiro and Schuchardt, 2003; Chadderdon et al., 2014; Zhang and Deng, 2015; Carro et al., 2018a; Sajid et al., 2018; Sayed et al., 2020; Troiano et al., 2020; Yuan et al., 2020; Kawde et al., 2021). Biocatalytic oxidation of HMF and other aliphatic and aromatic alcohols, diols and polyols using enzymes or microbial cells offers a facile and selective route of reaction under mild conditions. Many microorganisms including bacteria and fungi have shown the ability to metabolize HMF to different products including FDCA (Yuan et al., 2020). In some microorganisms, the HMF oxidizing activity has been attributed to HMF oxidase (HMFO), which is expected to form FDCA as the final product (Dijkman, 2015; Troiano et al., 2020; Vinambres et al., 2020). Besides HMFO, other reports describe the use of a cascade of enzymes for achieving oxidation of HMF to FDCA. For example, a cascade of galactose oxidase M3-5 variant, that oxidizes HMF to DFF, and Escherichia coli periplasmic aldehyde oxidase ABC, oxidizing DFF to FFCA, and consequently to FDCA has been reported (McKenna et al., 2017). This pathway was further developed by including horseradish peroxidase and lipase B from Candida antarctica (McKenna et al., 2015, 2017). The potential drawbacks of an enzyme cascade system could be the differences in the optimum reaction conditions for the enzymes involved and the cost of enzymes production (Troiano et al., 2020).

Glucose–methanol–choline oxidoreductase (GMC) family includes flavoprotein oxidoreductases that are classified into several subfamilies acting on a wide range of substrates from short alcohols such as methanol and choline to more complex substrates such as glucose, aromatic alcohols, fatty alcohols among others (Sütl et al., 2019; Aleksenko et al., 2020; Savino and Fraaije, 2021). According to the Pfam database, GMC oxidoreductases include two domains; the first is Pfam00732 with a typical GxGxxG/A motif, an indicator for the Rossman fold, and responsible for binding the adenosine diphosphate (ADP) moiety of the flavin adenine dinucleotide (FAD) prosthetic group, while the second is the less conserved C-terminal domain, which usually contributes to the substrate specificity of the enzyme. A characteristic of these oxidoreductases is a conserved histidine in the active site that initiates the oxidation via proton subtraction from the substrate and assists in FAD re-oxidation by the molecular oxygen. Among the subfamilies of GMC oxidoreductases are aryl alcohol oxidases (AAOs) that have been reported to oxidize HMF (Seranno et al., 2020; Lappe et al., 2021). Aryl-alcohol oxidases (AAOs; EC 1.1.3.7) catalyse the oxidative dehydrogenation of primary alcohols from aliphatic unsaturated or aromatic alcohols to aldehydes using oxygen as a final electron acceptor resulting in production of hydrogen peroxide (H2O2) (Urlacher and Koschorreck, 2021). Aryl alcohol oxidases have also been classified in the CAZy (the carbohydrate-active enzymes) database into subfamily AA3_2, where AA3 refers to Auxiliary Activity (AA) superfamily that comprises four subfamilies with different GMC oxidoreductases (Sütl et al., 2018). These enzymes are known to play an important role in biomass degradation by way of supplying hydrogen peroxide to the ligninolytic peroxidases.

Mycobacterium sp. MS1601 is a promising source of enzymes for oxidation reactions, which is evident from the abundance of gene sequences encoding oxidases and dehydrogenases in its genome (Sayed et al., 2017a). The organism has shown unique activity for selective oxidation of branched polyols such as...
trimethylolpropane (TMP) to 2,2-bis(hydroxymethyl) butyric acid (BHMB) with high selectivity and yield (Sayed et al., 2016, 2017a) and was also recently shown to oxidize 1,6-hexanediol to 6-hydroxyhexanoic acid and adipic acid (Pyo et al., 2020).

This report presents a study demonstrating the ability of Mycobacterium sp. MS1601 to oxidize HMF to FDCA and other oxidized intermediates, identification of a novel aryl alcohol oxidase (MycspAAO) in the bacteria that efficiently catalyses oxidation of HMF, heterologous production of the enzyme in Escherichia coli and determination of its substrate scope. Furthermore, through sequence and structure analysis based on homology modelling, an important mutation was generated that resulted in improved HMF oxidation and also revealed FDCA production.

**Results**

**Oxidation of HMF using Mycobacterium sp. MS 1601 cells**

Mycobacterium sp. MS1601 cells grown in the medium with different carbon sources including glycerol, glucose and sorbitol, respectively, were used for reaction with 5 mg ml⁻¹ (39.6 mM) HMF for 120 h. Complete consumption of HMF was observed in all cases, but only the cells activated with glycerol showed the ability to produce FDCA with around 60% yield and selectivity and the remaining 40% was represented by HMFCFA (Fig. 1A). Formation of FFCA and BHMF was observed in the initial stage of the reaction and reached a maximum yield of 30.8% and 16%, respectively, before being converted to FDCA and HMFCFA, respectively (Fig. 1A). Meanwhile, the cells grown with glucose or sorbitol as carbon source, led to the accumulation of BHMF and HMFCFA as products (Fig. 1B and C). BHMF reached a yield of 82% and 77.3% after 12 h of the reaction with the Mycobacterium cells activated with glucose and sorbitol, respectively, and was thereafter gradually converted to give 61.6% and 65.8% of HMFCFA in the two cases at 120 h of the reaction. The reaction profiles revealed that the enzyme(s) catalysing the reduction of HMF to BHMF were activated more by glucose.

**MycspAAO identification and sequence analysis**

To identify the putative HMF oxidizing enzymes in the Mycobacterium sp. M1601, BLAST search focused only on the genome using MetspHMFO sequence, the HMFO from Methylovorus sp., as a query sequence, was performed (Dijkman and Fraaije, 2014). The search resulted in three candidate sequences, namely A0A1P8X5Y5, A0A1P8XFW1 and A0A1P8XIE5 that showed putative oxidase features. The three sequences were preliminary annotated as GMC oxidoreductases and showed reasonable sequence identities of 33%, 29% and 25%, respectively, to MetspHMFO. Further analysis of the A0A1P8X5Y5 sequence using HAMMER server and Pfam databases indicated the presence of the two GMC Pfam domains i.e., GMC_oxred_N (PF00732) and GMC_oxred_C (PF05199). BLAST search using A0A1P8X5Y5 sequence to the Uniprot database retrieved 250 hits that have been annotated as choline dehydrogenase, alcohol dehydrogenase and GMC oxidoreductases among others. The highest sequence identity retrieved was 78.4% while the lowest was 43.5% (Fig. S1A–C), however none of these 250 hits have been experimentally characterized up-to-date. Alignment with the characterized CAZY AA3 proteins showed A0A1P8X5Y5 sequence to cluster with AA3_2 subfamily (Fig. S1B). Alignment with the fungal GMC oxidoreductases further revealed certain conserved residues
including the catalytic histidine (Fig. S1C and D). To this end, we decided to codon optimize and synthesize the DNA sequence of A0A1P8X5Y5 (Fig. S2), henceforth called as MycspAAO, for expression in E. coli.

Phylogenetic analysis of the A0A1P8X5Y5 sequence against the HMFOs and AAOs that have been experimentally characterized so far, showed the relatedness of the bacterial HMFOs to each other, while that of the fungal AAOs to each other (Fig. 2). MycspAAO has been located in-between these two subfamilies since it is a bacterial AAO. Since there is no published sequence of a characterized bacterial AAO up-to-date, they were not included in this phylogenetic analysis. The multiple sequence alignment of MycspAAO to the three characterized HMFOs showed specific variations in the sequence motifs that have been recently described (Vinambres et al., 2020). These variations are as follows: the ADP-binding domain – GAG instead of GGG conserved motif, in the active site region – VYYH instead of WVWH, in the PS000623 motif – PRGH instead of EQGR and in the PS000624 motif – GGA instead of AGA (Fig. S3).

**Overexpression of MycspAAO and its mutants in E. coli**

*Escherichia coli* BL21 (DE3) transformed with MycspAAO gene was grown in NYAT and Terrific Broth (TB) medium, and the gene expression after induction was tested at 15, 20 and 30°C, respectively. Determination of HMF oxidizing activity in the clarified cell lysate after cell disruption showed conversion of HMF (4 mM) to be higher with the recombinant cells from the TB medium, corresponding to 97%, 88% and 90% for the cells with gene expression at 15, 20 and 30°C, respectively (Fig. S4A). In the NYAT medium, only the cells induced at 15°C provided an active cell lysate giving HMF conversion of 84%, while the lysates from the cultures at 20 and 30°C gave HMF conversion of only 14 and 0.5%, respectively (Fig. S4B). *Escherichia coli* cells carrying the plasmid without the MycspAAO gene showed no conversion of HMF. Hence, for further experiments, cultivation in TB and gene expression at 15°C were chosen for MycspAAO production in E. coli BL21 (DE3). The enzyme was purified from the cell lysate by metal ion affinity chromatography and was obtained as a protein with an estimated molecular weight of 54 kDa by SDS-PAGE. The purified yellow-coloured MycspAAO, obtained at a concentration of about 1 mg ml⁻¹ was used for further experiments. The enzyme was confirmed as a FAD dependent oxidase through its thermal denaturation and spectrophotometric measurement of the released FAD at 450 nm (not shown).

**MycspAAO substrate spectrum and optimum conditions for activity**

Preliminary screening of MycspAAO activity was performed against a range of compounds using a common assay coupled with horseradish peroxidase catalysed determination of H₂O₂ released (Fig. 3A). MycspAAO exhibited higher specific activity, 4.95, 4.09, 3.79, 3.63 and 3.16 U mg⁻¹ with 3,4-DMBA, HMF, DFF, BHMF and MBA, respectively, as substrates. The specific activity values were 1.2, 0.94, 0.66, 0.23 and 0.12 U mg⁻¹ with 4HBA, furfural, BA, FFCA and HMFCA, respectively. While the activity against the aliphatic alcohol, n-propanol was in the same range (0.13 U mg⁻¹), activity against choline, pyridoxine, glycerol and glucose was extremely low (Fig. 3C). No colour was observed in the control reaction mixtures without MycspAAO (Fig. S5).

Optimum temperature and pH for oxidation of HMF by MycspAAO were then determined by chromatographic analysis of the substrate and products. The MycspAAO activity showed a rather broad pH profile between pH 5–9, with the highest activity at pH 8 giving 93% HMF conversion in 2 h (Fig. 3B). Reactions performed with 4 mM HMF at pH 8 showed the highest substrate conversion (98%) at 40°C followed by 94% at 30°C, while the enzyme activity was drastically decreased at 60°C and above (Fig. 3C).

Subsequently, MycspAAO was used for oxidation of 4 g l⁻¹ HMF (31.7 mM) at pH 8 and 30°C for a longer time period. The reaction profile in Fig. 4 shows DFF to

---

**Fig. 2.** Phylogenetic analysis of the A0A1P8X5Y5 sequence (MycspAAO) in the genome of Mycobacterium sp. MS1601 with bacterial HMFO and fungal AAO sequences described with enzymatic activities in literature. Analysis is done using Mega X software.
be the first oxidation product, which was totally converted to FFCA with 100% yield and selectivity. No FDCA was observed in the reaction.

Homology modelling and mutagenesis of MycspAAO
The three-dimensional structure of MycspAAO was predicted through homology modelling, and the final model obtained was assessed using the Swiss model assessment server (Fig. S6). The Ramachandran plot showed that 95.2% of the residues are located in the allowed region, and zero outliers were recorded. The normalized QMEAN Score for the model was < 1, which indicates acceptable model quality (Fig. S7).

The obtained MycspAAO 3D model (Fig. S6) showed the basic features of GMC oxidoreductases, including the Rossman fold that holds the FAD cofactor and the C-terminal domain. MycspAAO does not have the conserved His/His residues, instead a His/Asn was observed. Moreover, the enzyme binds non-covalently to the FAD cofactor due to the missing conserved His in the position 91, which is replaced by Ile. The model has been compared with the PDB 3FIM (fungal AAO of Pleurotus eryngii) (Fig. 5). The comparison showed the presence of the characteristic loop located in front of the base of the enzyme.

Fig. 3. Specific activity of MycspAAO-WT against different aliphatic, aromatic alcohols, and putative substrates for the GMC oxidoreductase enzymes at pH 8 and 30°C (A). Effect of (B) pH and (C) reaction temperature on the conversion (%) of HMF (4 mM) using purified MycspAAO-WT (40 μg ml⁻¹) after 2 h reaction. The reaction at different temperatures was performed at pH 8, while the effect of pH was studied at 30°C. All experiments in (A) were carried out in 4 replicates, and experiments for (B) and (C) were done in duplicates. The data in all the figures represent the mean and standard deviation from the replicates.

Fig. 4. Profiles of HMF (●) oxidation to DFF (◆) and FFCA (▲) during the reaction of 31.7 mM HMF (4 g l⁻¹) with the purified MycspAAO-WT (200 μg ml⁻¹) at 30°C and pH 8. The experiment was done in duplicate and the data shown represents the mean and standard deviation from the replicates.
active site that has been described to control the access of the substrates to the active site (Fig. 5B) (Carro et al., 2018b).

Upon investigation of the active site of MycspAAO based on the model structure generated in this study, we noticed the narrow active site was due to the presence of two tyrosine residues Y443 and Y444 in the substrate-binding site facing the FAD molecule (Fig. S8). These two residues were thus listed for site-directed mutagenesis. The residues that were probably interacting and to the substrate binding were as follows: Y443, Y444, L298, M337, V360, A358, A93 and V95. Based on these observations, 11 mutations were proposed (Table S1) and were carried out using the primers designed for site-directed mutagenesis of the different amino acid residues.

Out of the eleven mutations planned, expression of only seven enzyme variants was successfully obtained. Screening of the activity of the variants against 2.5 mM HMF using clarified cell lysates (obtained from cell suspensions with similar OD600) and HPLC analysis of the reaction sample after 48 h, showed huge activity reduction for most of the variants i.e., Y444A, L298S, V95A, M377F and A358R, while L298A variant showed slightly lower activity than the wild-type MycspAAO (Fig. 6). Y444F was the only variant with higher activity compared with MycspAAO-WT.

Fig. 5. Comparison of the active site of an aryl alcohol oxidase, PDB 3FIM (A) to the model of MycspAAO (B) in the present study. The loop covering the FAD controlling the substrate entrance and product release is shown in the panel. The comparable residues are as follows 3FIM: I500, F501, H502, H546, I391, F397 and the counterpart residues in MycspAAO are Y443, Y444, H445, N487, M337 and F343. The aromatic residues in MycspAAO W52, H316, F343, Y444 and Y443 are probably contributing to the substrate specificity and oxygen activation by the FAD cofactor.

Fig. 6. Residual HMF concentration after 48 h during activity screening of the clarified cell lysates of E. coli expressing MycspAAO-WT and the mutant enzymes against 2.5 mM HMF at 30°C and pH 8. The higher bars indicate lower activity. All reactions were carried out in duplicates and results in the figure represent the mean and standard deviation from the replicates.

© 2022 The Authors. Microbial Biotechnology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Microbial Biotechnology, 15, 2176–2190
Kinetic parameters and FDCA production

Y444F and L298A variants as well as the wild-type Mycsp AAO were then produced in one-litre medium and purified. The purity of the proteins was checked using SDS-PAGE (Fig. S9), and concentrations were determined using BCA reagent based on BSA standard (Table S3). Figure S10 shows the initial reaction rate (V) for HMF oxidation at different concentrations by the wild-type (WT), Y444F and L289A enzymes during 3 h reaction time. Y444F showed a significant increase in the reaction rate up to 10 mM HMF, above which there was a drastic decrease in activity and the reaction rate remained constant up to 60 mM HMF (Fig. S10). For the WT and L289A enzymes, the reaction rate decreased above 4 mM HMF, and further decrease was observed above 40 mM HMF. At 10 mM HMF, the initial activity of Y444F was around 10 times higher than the other MycspAAO variants.

Determination of the steady-state kinetic parameters of the enzyme variants revealed MycspAAO-Y444F to have about two-fold higher $k_{cat}$ and $V_{max}$, 1.7-fold lower $K_m$ and about 3-fold higher $k_{cat}/K_m$ compared with MycspAAO-WT (Table 1). On the other hand, the L298A variant showed almost same $k_{cat}$ and $V_{max}$ and $K_m$ with slight difference compared with the wild-type enzyme (Table 1).

MycspAAO-WT and -Y444F (200 µg ml$^{-1}$ total protein) were further used for oxidation of 10 mM HMF (lower concentration than that in Fig. 3). Both enzymes oxidized HMF completely within 6 h mainly to DFF; the formation of DFF is seen already at 0 h reaction (immediately after addition of the enzymes) and is higher with the Y444F variant (Fig. 7A). Diformyl furan is subsequently converted to FFCA. Further oxidation of FFCA to FDCA by both WT and the mutant enzyme was observed after 96 h of reaction (Fig 7A). The oxidation rate of HMF to DFF as well as DFF to FFCA was higher with MycspAAO-Y444F compared with the wild-type enzyme, confirming the findings from the colorimetric assay as well as the kinetic parameters above (Fig. 7 and Table 1).

Discussion

The genome sequence of Mycobacterium sp. MS1601 shows the presence of many gene sequences that are annotated as oxidative enzymes (296 dehydrogenases; 168 oxidoreductases and 53 oxidases) (Sayed et al., 2017b). The organism has so far been used as an efficient biocatalyst for the highly selective oxidation of polyhydric alcohols (Sayed et al., 2016, 2017b), and in this study for the oxidation of HMF. Interestingly, oxidation of HMF by Mycobacterium sp. MS1601 seems to follow different routes depending on the carbon source used for growing the cells. While the cells grown on glycerol, oxidized HMF to FDCA via FFCA (pathway A in Scheme 1) (Fig. 1A), a typical pathway used by HMFO from Methylovorus sp. MP688 (Dijkman and Fraaije, 2014), oxidation by the cells grown on glucose or sorbitol followed an alternative route in Scheme 1 yielding BHMF and HMFCa that was not oxidized further (Fig. 1B and C). In the latter case, the aldehyde group of HMF is either oxidized directly to HMFCa or after first being reduced to BHMF. Activation of the enzyme(s) responsible for FDCA production with glycerol was in agreement with the previous report involving selective oxidation of trimethylolpropanol (TMP) to its corresponding carboxylic acid by Mycobacterium sp. MS1601 (Sayed et al., 2017b).

Based on the similarity of the HMF oxidation profile by the glycerol-activated cells to that by MetspHMFO, the sequence of the Methylovorus sp. enzyme was used for probing the Mycobacterium sp. MS1601 genome for homologous enzymes. The putative gene sequence with > 30% identity with MetspHMFO was chosen for this study. Since the identity to MetspHMFO was low, further sequence analysis was done for the chosen gene from Mycobacterium sp. MS1601 using Pfams and Uniport databases, confirming that the chosen sequence encodes an enzyme belonging to GMC family, to which MetspHMFO also belongs (Dijkman and Fraaije, 2014). However, the first 250 hits from Uniport analysis with high identity were for other GMC members that are not yet experimentally characterized. This observation confirmed the novelty of the chosen protein from Mycobacterium sp. MS1601. The codon-optimized gene of the above-mentioned gene from Mycobacterium sp. MS1601 was expressed in a soluble form in E. coli BL21 (DE3) at 15°C. The purified MycspAAO was shown to be a wide spectrum GMC oxidase enzyme that utilizes oxygen as the final electron receptor and prefers compounds with an aromatic ring as 3,4-DMBA>HMFO>DFF>BHMF>MBA>4-HBA> furfural>BA compared with aliphatic substrates (Fig. 3A). The high specific activity of MycspAAO against 3,4-DMBA, MBA and BA is in agreement with the activity of bacterial aryl alcohol oxidase from Sphingobacterium sp. ATM (Tamboli et al., 2011). Our findings indicate that the new GMC member from Mycobacterium sp. MS1601 belongs to the subfamily of aryl alcohol oxidases, also displaying high activity against aryl alcohol oxidases, including DFF and BHMF (Fig. 3A). Reactions with DFF

Table 1. Steady-state kinetic parameters of the activities of MycspHMFO and its variants against HMF at 30°C and pH 8. The experiments for the kinetic parameters were done in duplicates.

|          | $V_{max}$ (mM min$^{-1}$) | $K_m$ (mM) | $k_{cat}$ (s$^{-1}$) | $k_{cat}/K_m$ (mM$^{-1}$ s$^{-1}$) |
|----------|--------------------------|------------|---------------------|----------------------------------|
| WT       | 0.04 ± 0.004             | 3.3 ± 0.3  | 384 ± 33.5          | 115 ± 1.6                        |
| L289A    | 0.038 ± 0.0002           | 3.1 ± 0.17 | 350 ± 17.8          | 110 ± 0.3                        |
| Y444F    | 0.072 ± 0.0002           | 1.9 ± 0.005| 652 ± 1.7           | 342 ± 0.0                         |

© 2022 The Authors. Microbial Biotechnology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Microbial Biotechnology, 15, 2176–2190
and furfural as substrates suggest that the enzyme is active also on the aldehyde groups; the oxidation of DFF and BHMF resulted in the formation of FFCA. Aryl alcohol oxidases in rare cases have been shown to catalyse the oxidation of aromatic aldehyde substrates to the corresponding acids (Urlacher and Koschorreck, 2021), which may be due to interference by the carboxylic group on the binding of the substrate to the enzyme active site.

More insights were gained from comparison of MycspAAO to the closest fungal aryl alcohol oxidase PDB 3FIM. A characteristic loop covering the active site has been described in the case of PDB 3FIM, and interestingly a similar loop has been identified in the MycspAAO model. Most importantly, the previously reported F397 in 3FIM has an equivalent residue F343 in MycspAAO (Fig. 5). This phenylalanine residue is proposed to attach to the substrates and to be involved in a significant loop displacement necessary to place the substrate inside the active site (Carro et al., 2018b; Serrano et al., 2020). Moreover, the analysis of MycspAAO 3D model showed a narrow active site due to the presence of two tyrosine residues (Y443 and Y444) located very close to the catalytic H445. These two tyrosine residues may be expected to cause interference for the bulkier substrates to be accepted by the enzyme.

MycspAAO-Y444F exhibits favourable kinetic parameters against HMF, 4HBA, and furfural compared with the wild-type enzyme (Table 1 and Fig. S11). With HMF, the improved activity of the enzyme variant was ascribed to the significantly higher $k_{cat}/K_m$ and lower $K_m$ values (Table 1). The higher activity of MycspAAO-Y444F can be further related to the observations recorded for the fungal AAO (PDB 3FIM). F444 in the enzyme variant is exactly equivalent to the F501 in the fungal AAO (3FIM) structures, which is proposed to be a major player in the substrate activity and specificity. This phenylalanine has also been described to control the access of molecular oxygen to the FAD molecule in the active site (Hernandez-Ortega et al., 2011; Ferreira et al., 2015; Carro et al., 2018b). Hence, MycspAAO-Y444F denotes a back-to-ancestral mutation since phenylalanine residue is conserved in many GMC oxidoreductases and specifically in most of the AAOs that have been characterized up to date (Chakraborty et al., 2014; Galperin et al., 2016; Jankowski et al., 2020).

Both MycspAAO-WT and MycspAAO-Y444F variants, when tested against a lower concentration of HMF (10 mM in contrast to 31.7 mM) revealed the formation of FDCA over a long incubation time, the Y444F variant giving the higher activity (Fig. 7). These observations are in agreement with that reported previously for HMFOs and AAOs (Cavener, 1992; Dijkman et al., 2014; Serrano et al., 2019; Vinambres et al., 2020). Very recently, Sánchez-Ruiz et al. (2021) reported that FCA, especially at higher concentration (over 15 mM) has an irreversible inhibition against HMFOs (Sánchez-Ruiz et al., 2021), which could explain FDCA formation from 10 mM HMF (Fig. 7), and not during the oxidation of 31.7 mM HMF with MycspAAO-WT (Fig. 3). This finding was further confirmed from screening Mycsp AAO-WT, -Y444F and L298A variants against concentrations of HMF (Fig. S10), that showed that increasing the concentration of HMF over 10 mM reduced the initial activity of all the variants although the Y444F variant showed a significantly higher activity.
In summary, this report indicates that *Mycobacterium* sp. MS1601 is an interesting source of enzymes for oxidation of HMF. While the whole cells showed the formation of products with different levels of oxidation including FDCA, the analysis of the selected gene sequence from the genome of *Mycobacterium* sp. MS1601 indicates that the HMF oxidizing enzyme represents a novel bacterial AAO that oxidizes HMF preferably to FFCA but even to FDCA at lower substrate concentration and longer reaction time. The site-directed mutagenesis studies on MycspAAO indicated that the Y444 in the enzyme’s active site could be replaced by the conserved phenylalanine (F) for improved enzymatic activity. Further improvement through introducing additional rational mutations in MycspAAO are required to increase the capacity of MycspAAO to oxidize HMF completely to FDCA in a shorter reaction time and higher substrate concentration. On the other hand, being host to several oxidative enzymes, it is more likely that *Mycobacterium* sp. MS1601 cells utilize an aldehyde dehydrogenase/oxidase activity for oxidation of FFCA in vivo.

**Experimental procedures**

**Materials**

*Mycobacterium* sp. MS1601 (previously, *Corynebacterium* sp. ATCC 21245) was obtained from American Type Culture Collection (ATCC, USA). Ampicillin, kanamycin, bovine serum albumin, bicinchoninic acid (BCA), 5-hydroxymethyl furfural (HMF), 2-formyl-5-furan carboxylic acid (FFCA), 2,5-diformyl furan (FFA), 2,5 furan dicarboxylic acid (FDCA), sorbitol and glucose were purchased analytical grade. Protein expression, and the pET-22b[+] and pRSF-Duet vectors were acquired from Novagen (Madison, WI), while the putative oxidoreductases for HMF oxidation, *A0A1P8X5Y5* sequence was analysed in vivo.

**Cultivation of Mycobacterium sp. MS1601 for oxidation of HMF**

The stock culture of *Mycobacterium* sp. MS1601 was prepared by aseptically transferring the lyophilized culture to 25 ml of sterile nutrient broth containing per litre: 5 g peptone and 3 g beef extract (pH 6.8) in a 250 ml Erlenmeyer flask. The flask was incubated in an orbital-shaking incubator (Ecotron, Infors HT, UK) at 31°C and 300 rpm for 48 h. The resulting culture was mixed with sterile 20% (v/v) glycerol and then stored at −20°C.

For activation of the *Mycobacterium* sp. MS1601 cells for HMF oxidation, 200 µl of the glycerol stock culture was added to 50 ml sterile activation medium containing per litre: 5 g glycerol, sorbitol or glucose as carbon source, 2 g sodium acetate and 10 g yeast extract in 250 ml Erlenmeyer flask. The culture was incubated as above for 48 h. Four millilitres of the culture broth were withdrawn, centrifuged and washed with 100 mM phosphate buffer pH 8, which was also used as reaction buffer. The cells were then suspended in 1 ml reaction buffer supplemented with 5 mg ml⁻¹ (39.6 mM) HMF in 4 ml vials. The vials were incubated in a thermomixer (MKR 13, HLC Biotech, Germany) at 30°C and 500 rpm. Samples of 50 µl were collected at regular time intervals for analysis of substrate, product and co-products. All experiments were carried out in duplicates.

**MycspAAO identification, sequence analysis, homology modelling and mutation design**

BLAST search within the genome of *Mycobacterium* sp. MS1601 was done using *Metsp*HMFO sequence as a search query to find the putative oxidoreductases for HMF oxidation. The *A0A1P8X5Y5* sequence was analysed using HAMMER server (Potter et al., 2018). Multiple sequence alignment was done using PROMALS3D server (Pei et al., 2008; Vinambres et al., 2020). A maximum-likelihood phylogenetic tree was prepared by MEGA X software (Stecher et al., 2020; Tamura et al., 2021) applying 100-iteration bootstrapping (Felsenstein, 1985) using the Whelan and Goldman (2001) model of evolution.

The 3D structure for MycspAAO (*A0A1P8X5Y5*) was modelled using the Swiss-Model server, the nearest 12 templates in terms of coverage and sequence identities were used to build 12 models by the server (Pei et al., 2008; Waterhouse et al., 2018). The best model in terms of QMEAN score was eventually selected and was based on the template 2JBV, a choline oxidase that showed 38.9% sequence identity to MycspAAO (Quaye et al., 2008; Benkert et al., 2009). Another model was also built using YASARA Structure software and based solely on the template 3FIM, a fungal AAO to confirm the positions of the surface loop in front of the active site and the aromatic residues contributing to the active site. The FAD molecule was copied to the model and final energy minimization was performed using YASARA structure software (Krieger et al., 2002). The docking analysis was carried out using AutoDock integrated into YASARA structure software and utilizing the macro dock_run.mer (Morris et al., 2009). Molecular modelling was performed for 2 picoseconds using YASARA structure macro md_run.mer.
Cloning, expression and purification of MycspAAO and its mutants

MycspAAO gene was codon-optimized for expression in E. coli and tagged with a His-tag binding domain at the C-terminal end, and restriction sites for EcoRI and XhoI restriction enzymes were added to N-terminal and C-terminal, respectively (Fig. S1). The gene was synthesized at IDT (Integrated DNA Technologies, USA) and delivered in a pEX-K248 vector. pEX-K248-MycspAAO was co-digested with EcoRI and Xhol, and the desired gene fragment was ligated using T4 DNA Ligase to the expression vector pET-22b(+) that was co-digested with the same restriction enzymes. The plasmid was then transformed into E. coli DH5α, which was propagated in SOC medium for 60 min at 37°C, and grown on Luria-Bertani agar (Applichem, Darmstadt, Germany) in the presence of 100 µg ml⁻¹ ampicillin. Plasmids were extracted from the resulting colonies and the correct clone was chosen after confirmation by colony PCR and sequencing (Eurofins Genomics, Germany), and the plasmid pET-22b(+)MycspAAO was then transformed into chemically competent E. coli BL21 (DE3).

For expression of MycspAAO gene, 250 µl of an overnight culture of E. coli BL21 (DE3) cells bearing the pET-22b (+)-MycspAAO coding plasmid was transferred to 50 ml of Terrific Broth (TB) or NYAT medium containing 100 µg ml⁻¹ ampicillin and grown at 37°C until reaching OD₆₀₀ of 0.5. The cells were then induced with 1 mM IPTG and grown for 48 h at 15, 20 and 30°C, respectively, before harvesting by centrifugation at 4700 g for 15 min at 4°C (Sorvall LYNX 4000, Thermo Scientific, Germany). The cell pellet from 10 ml broth (OD₆₀₀ nm of 5 adjusted for all samples) was resuspended in 1 ml BugBuster™ Protein Extraction Reagent (Novagen, Germany) in 2 ml tubes and then incubated at 30°C for 15 min. Subsequently, all tubes were centrifuged for 30 min at 29 100 g, 4°C to separate the soluble and insoluble fractions. All subsequent steps were performed at 4°C. Escherichia coli BL21 (DE3) cells bearing the pET-22b (+) plasmid without the target gene were used as negative control. The soluble fraction was checked for MycspAAO expression by testing enzyme activity against 5-HMF and SDS-PAGE.

To purify the MycspAAO protein, the cells were cultured in 1L TB medium, induced with 1mM IPTG and incubated at 17°C for 48 h. The cell pellet was collected after centrifugation, lysed as mentioned above, and the protein was purified from the soluble fraction on 5 ml Ni-Sepharose™ High Performance, HisTrap™ HP column (GE-Healthcare) pre-equilibrated with the binding buffer of 50 mM sodium phosphate buffer and 40 mM imidazole. After washing the column, the bound protein was eluted with an elution buffer of 0.1 M phosphate buffer pH 7 supplemented with 500 mM imidazole. The eluted protein was dialysed against 50 mM phosphate buffer pH 8 using a dialysis membrane of 12–14 kDa molecular weight cutoff (Spectra/ Por™, Thermo Fisher Scientific). The purified enzyme was used for further experiments.

The MycspAAO gene was transferred to pRSF-Duet plasmid with kanamycin resistance marker using the same restriction enzymes as indicated above. The target mutations were carried out using the designed primers (Table S2) and QuickChange™ site-directed mutagenesis kit. The PCR reaction of 50 µl contained 10 ng of template, 0.9 µl (125 ng) of primer pairs, 5 µl of 10x reaction buffer, 1 µl of dNTPs and 1 µl of Pfu ultra-DNA polymerase. The reaction was started at 95°C for 30 s for DNA template denaturation, followed by 16 amplification cycles, each of which involved 95°C for 30 s, 55°C for 1 min and 68°C for 5 min for template denaturation, annealing and extension, respectively. The PCR products were mixed gently with 1 µl of DpnI and incubated for 1 h at 37°C and then the products were analysed by gel electrophoresis. Thereafter, 2 µl of the DpnI treated PCR products were transformed into E. coli One Shot Top10 using heat shock and then placed on ice for 30 min.

The transformed cells were propagated in SOC medium for 1 h at 37°C and 225 rpm and then plated on LB agar supplemented with 50 µg ml⁻¹ kanamycin and incubated overnight at 37°C. The colonies obtained were sub-cultured on LB broth supplemented with 50 µg ml⁻¹ kanamycin. The plasmids with wild-type and mutant MycspAAO were purified using Qiagen miniprep kits (Qiagen, Hilden, Germany) and used for PCR analysis using the corresponding primer pairs for further confirmation. Moreover, the mutations were confirmed by DNA sequencing (Eurofins, Hamburg, Germany). The plasmids with successful mutations were transformed into E. coli BL21 (DE3) competent cells, which were cultivated overnight at 37°C in LB medium with 50 µg ml⁻¹ kanamycin. The cells were then suspended in 20% glycerol and stored at –80°C for further use.

Expression of wild-type and mutant MycspAAO genes was carried out in 50 ml TB medium containing 50 µg ml⁻¹ kanamycin. The cells were grown at 37°C until reaching OD₆₀₀ of 0.5 and then induced with 1 mM IPTG, after which the cell growth was continued for 72 h at 17°C. The cells were harvested and lysed as described above. Twenty microlitre supernatant was used for SDS-PAGE, and 200 µl was added to 800 µl of 0.1 M phosphate buffer pH 8 in a total volume of 1 ml containing 2.5 mM HMF, for the activity tests. MycspAAO, MycspAAO-Y444F and MycspAAO-L298A exhibiting relatively higher activity against HMF were chosen for production in 1L of TB medium and
purification of the enzymes was carried out as mentioned above. The purified enzymes after desalting were stored in 40% glycerol at −20°C for further screening and application.

The homogeneity of the purified protein was confirmed by SDS-PAGE using 12% (w/v) separating gel and 3% (w/v) stacking gel (Laemmli, 1970). The protein concentration was determined using a BCA assay kit with bovine serum albumin as standard following the manufacturer’s instructions.

Substrate spectrum and optimum conditions for activity and of MycspAAO

Activity of pure MycspAAO against HMF (4 mM; 0.5 g l−1) was determined using 40 μg ml−1 enzyme in 200 μl of 100 mM sodium phosphate buffer pH 8 in 1.5 ml Eppendorf tubes and incubation at 30°C and 500 rpm in a thermoshaker (MKR 13, HLC Biotech, Germany) for 2 h. The concentrations of substrate and product(s) were determined by HPLC. For determining optimum pH, the reaction was performed in 100 mM phosphate buffer pH 5, 6, 7, 8 and 9, respectively, at 30°C while maintaining the other conditions as above. The optimum temperature of the HMF oxidizing activity was determined by varying the reaction temperature between 20–70°C. The optimum conditions of pH and incubation temperature were then applied for oxidation of 4 mg ml−1 HMF using 40 μg pure MycspAAO in 1 ml of 100 mM phosphate buffer in 4 ml vials, incubated at 30°C and 700 rpm in a thermomixer for 48 h. All experiments were carried out in duplicates.

The activity and specific activity of the pure MycspAAO was also tested against 5 mM of different substrates besides HMF, including furfural, FFCA, DFF, BHMF, HMFC, 4-hydroxy benzyl alcohol (4-HBA), benzyl alcohol (BA), n-propanol (n-P-OL), veratryl alcohol (3,4 di-methoxy benzyl alcohol, 3,4DMBA), methoxy benzyl alcohol (MBA), choline, pyridoxine, glycerol, methanol and glucose, using 20 μl of the enzyme solution (protein concentration of 8 μg ml−1) in 200 μl total reaction volume in a microtitre plate. The assay was performed at pH 8 and 30°C as a coupled assay in which H2O2 generated in the oxidation reaction is used by horseradish peroxidase (HRP, 15 U ml−1) to oxidize 4-aminoantipyrine (0.1 mM) and 3,5-dichloro-2-hydroxybenzensulfonic acid (1 mM) to form a pink product that is detected at 515 nm (Dijkman and Fraaije, 2014) in a spectrophotometer (Multiskan Go, Thermo Fisher Scientific, Sweden). Blank reactions corresponding to each substrate without enzyme were included as negative controls. The assay was also used to determine the specific activity of MycspAAO and MycspAAO-Y444F against HMF, furfural, 4-HBA, glucose, glycerol, choline, pyridoxine and methanol.

Calculations of activity, yield and kinetic analysis

The activity and specific activity of the wild-type MycspAAO against different substrates was calculated using the following equations:

\[
\text{Activity} = \frac{\text{Sample} \times \text{t} - \text{Blank} \times \text{t}}{\text{total reaction volume} \times \text{dilution factor} \times \text{enzyme volume} \times \epsilon} = U\left(\frac{\mu\text{mole}}{\text{min}}\right) / \text{ml} \quad (1)
\]

Specific activity = \(\frac{\text{U/ml}}{\text{total protein conc. (mg/ml)}} = U\left(\frac{\mu\text{mole}}{\text{min}}\right) / \text{mg} \quad (2)
\]

The product yield \((Y_p/s)\) and selectivity \((S)\) were calculated using the following equations:

\[
Y_{p/s} (%) = \left[\frac{\text{Product}_{\text{final}} - \text{Product}_{\text{initial}}}{\text{Substrate}_{\text{initial}} - \text{Substrate}_{\text{final}}}\right] \times 100
\]

\[
S(%) = \left[\frac{\text{Product (moles)} - \text{byproducts (moles)}}{\text{converted HMF (moles)}}\right] \times 100
\]

The kinetic parameters, \(V_{max}\), \(K_m\), \(k_{cat}\) and \(k_{cat}/K_m\) for MycspAAO, MycspAAO-Y444F and MycspAAO-L298A were determined by Lineweaver–Burk plot with data from the reactions with 0.1–80 mM HMF as substrate using 10 μg ml−1 of the purified enzyme variants in 200 μl total reaction volume of 100 mM sodium phosphate buffer pH 8 at 30°C for 3 h. Fifty microlitre samples were collected every 30 min and the slope of the linear relation between HMF concentration (mM) and time was used to calculate the reaction rate (V). Both \(V_{max}\) and \(K_m\) were obtained directly from Lineweaver–Burk plot and \(k_{cat}\) was calculated according to the following equation \(k_{cat} = V_{max}/[E]_0\) where \([E]_0\) is the initial enzyme concentration (mM).

The activities of pure MycspAAO-WT and MycspAAO-Y444F were compared for oxidation of HMF by mixing 200 μl enzyme solutions (1 mg ml−1) with 0.8 ml of 10 mM HMF in 0.1 mM sodium phosphate buffer pH 8 in 4 ml vials, incubating in a thermoshaker at 30°C and
500 rpm and collecting 20 µl samples at regular time intervals for HPLC analysis.

**Analytical procedures**

The concentrations of 5-HMF, FFA, FFCA and FDCA were determined using HPLC (JASCO, Tokyo, Japan) equipped with a fast acid analysis chromatographic column connected to a guard column (Biorad, Richmond, CA, USA), refractive index detector (ERC, Kawaguchi, Japan), a JASCO UV detector operating at 215 nm and a JASCO intelligent autosampler. The column temperature was maintained at 65°C in a chromatographic oven (Shimadzu, Tokyo, Japan). Samples were diluted with Milli-Q quality water and mixed with 20% (v/v) H2SO4 (20 µl ml⁻¹ sample) and then filtered. A 40 µl aliquot was injected in 0.5 mM H2SO4 mobile phase flowing at a rate of 0.4 ml min⁻¹. The peaks for the different compounds were confirmed and quantified using the corresponding external standards.

To determine the FAD bound to the enzyme, 500 µl of 1 mg ml⁻¹ of pure MycspAAO-WT was filtered using 30 kDa cut-off and washed twice using 500 µl of 0.1 M phosphate buffer pH 8, and then denatured at 100°C for 15 min and separated by centrifugation at 15 000 g for 5 min. The released FAD was detected at 450 nm in a spectrophotometer (Boateng et al., 2015, Dishisha et al., 2019), and the concentration was calculated based on a standard curve.

**Acknowledgements**

The authors gratefully acknowledge the financial support from the Swedish Foundation for Strategic Environmental Research (Mistra) for the research program, Sustainable Plastics and Transition Pathways (STEPS) (grant no. 2016/1489) and the Swedish Research Council FORMAS (grant no. 942-2016-33) for the project Farm2-Furan at Lund University.

**Conflict of interest**

The authors declare no conflict of interest.

**Author contributions**

RHK and MS conceptualized the work. MS and RHK planned the experiments. MS designed and carried out the experiments and the analysis. YG designed and performed the in silico analyses. FJ helped with the mutagenesis experiments and EVM with the experiments on characterization of the wild-type enzyme. MS, RHK, SHP and YG were involved in writing and revising the manuscript.

**References**

Aleksenko, V.A., Anand, D., Remeeva, A., Nazarenko, V.V., Gordeliy, V., Jaeger, K.-E., et al. (2020) Phylogeny and structure of fatty acid photodecarboxylases and glucose-methanol-choline oxidoeductases. *Catalysts* **10**: 1072.

Benkert, P., Kunzli, M., and Schwede, T. (2009) QMEAN server for protein model quality estimation. *Nucl Acids Res* **37(suppl 2)**: W510–W514.

Boateng, M.O., Corrigall, A.V., Sturrock, E., and Meissner, P.N. (2015) Characterisation of the flavin adenine dinucleotide binding region of *Myxococcus xanthus* protoporphyrinogen oxidase. *Biochem Biophys Rep* **4**: 306–311.

Bozell, J.J., and Petersen, G.R. (2010) Technology development for the production of biobased products from bio-refinery carbohydrates—the US Department of Energy’s “Top 10” revisited. *Green Chem* **12**: 539–554.

Carro, J., Amengual-Rigo, P., Sancho, F., Medina, M., Guallar, V., Ferreira, P., and Martinez, A.T. (2018a) Multiple implications of an active site phenylalanine in the catalysis of aryl-alcohol oxidase. *Sci Rep* **8**: 8121.

Carro, J., Fernández-Fueyo, E., Fernández-Alonso, C., Cañada, J., Ullrich, R., Hofrichter, M., et al. (2018b) Self-sustained enzymatic cascade for the production of 2,5-furandicarboxylic acid from 5-methoxymethylfurfural. *Bio-technol Biofuels* **11**: 1–10.

Carro, J., Ferreira, P., Rodriguez, L., Prieto, A., Serrano, A., Balcells, B., et al. (2015) 5-hydroxymethylfurfural conversion by fungal aryl-alcohol oxidase and unspecific peroxidase. *FEBS J* **282**: 3218–3229.

Cavener, D.R. (1992) GMC oxidoreductases: a newly defined family of homologous proteins with diverse catalytic activities. *J Mol Biol* **223**: 811–814.

Chadderdon, D.J., Xin, L., Qi, J., Qiu, Y., Krishna, P., More, K.L., and Li, W. (2014) Electrocatalytic oxidation of 5-hydroxymethylfurfural to 2, 5-furandicarboxylic acid on supported Au and Pd bimetallic nanoparticles. *Green Chem* **16**: 3778–3785.

Chakraborty, M., Goel, M., Chinnadayala, S.R., Daiya, U.R., Ghosh, S.S., and Goswami, P. (2014) Molecular characterization and expression of a novel alcohol oxidase from *Aspergillus terreus* MTCC6324. *PLOS One* **9**: e95368.

Chen, G., van Straalen, N.M., and Roelofs, D. (2016) The ecotoxicocenogenic assessment of soil toxicity associated with the production chain of 2, 5-furandicarboxylic acid (FDCA), a candidate bio-based green chemical building block. *Green Chem* **18**: 4420–4431.

Davidson, M.G., Elgie, S., Parsons, S., and Young, T.J. (2021) Production of HMF, FDCA and their derived products: a review of life cycle assessment (LCA) and technoeconomic analysis (TEA) studies. *Green Chem* **23**: 3154–3171.

Dijkstra, W.P. (2015) HMF oxidase: characterization, application and engineering of 5-(hydroxymethyl)furfural oxidase. PhD Thesis. Groningen, the Netherlands: University of Groningen.

Dijkstra, W.P., Bindia, C., Fraaije, M.W., and Mattevi, A. (2015) Structure-based enzyme tailoring of 5-hydroxymethylfurfural oxidase. *ACS Catal* **5**: 1833–1839.

Dijkstra, W.P., and Fraaije, M.W. (2014) Discovery and characterization of a 5-hydroxymethylfurfural oxidase from...
McKenna, S.M., Leimkühler, S., Herter, S., Turner, N.J., and Carnell, A.J. (2015) Enzyme cascade reactions: synthesis of furanarboxylic acid (FDA) and carboxylic acids using oxidases in tandem. Green Chem 17: 3271–3275.

McKenna, S.M., Mines, P., Law, P., Kovacs-Schreiner, K., Birmingham, W.R., Turner, N.J., et al. (2017) The continuous oxidation of HMF to FDCA and the immobilisation and stabilisation of periplasmic aldehyde oxidase (PaoABC). Green Chem 19: 4660–4665.

Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F., Belew, R.K., Goodsell, D.S., et al. (2009) AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30: 2785–2791.

Pei, J., Kim, B.-H., and Grishin, N.V. (2008) PROMALS3D: a tool for multiple protein sequence and structure alignments. Nucleic Acids Res 36: 2295–2300.

Potter, S.C., Luciani, A., Eddy, S.R., Park, Y., Lopez, R., and Finn, R.D. (2018) HHMER web server: 2018 update. Nucleic Acids Res 46: W200–W204.

Pyo, S.-H., Park, J.H., Srebny, V., and Hatti-Kaul, R. (2020) A sustainable synthetic route for biobased 6-hydroxyhexanoic acid, adipic acid and ε-caprolactone by integrating bio-and chemical catalysis. Green Chem 22: 4450–4455.

Quaye, O., Lountos, G.T., Fan, F., Orville, A.M., and Gadda, G. (2008) Role of Glu312 in binding and positioning of the substrate for the hydride transfer reaction in choline oxidase. Biochemistry 47: 243–256.

Ribeiro, M.L., and Schuchardt, U. (2003) Cooperative effect of cobalt acetylacetonate and silica in the catalytic cyclization and oxidation of fructose to 2,5-furanarboxylic acid. Catal Commun 4: 83–86.

Sajid, M., Zhao, X., and Liu, D. (2018) Production of 2,5-furanarboxylic acid (FDCA) from 5-hydroxymethylfururfural (HMF): recent progress focusing on the chemical-catalytic routes. Green Chem 20: 5427–5453.

Sánchez-Ruiz, M.I., Martínez, A.T., and Serrano, A. (2021) Optimizing operational parameters for the enzymatic production of furanarboxylic acid building block. Microb Cell Fact 20: 1–13.

Savino, S., and Fraaije, M.W. (2021) The vast repertoire of carbohydrate oxidases: an overview. Biotechnol Adv 51: 107634.

Sayed, M., Dishisha, T., Sayed, W.F., Salem, W.M., Temerk, H.A., and Pyo, S.-H. (2016) Selective oxidation of trimethylolpropane to 2, 2-bis (hydroxymethyl) butyric acid using growing cells of Corynebacterium sp. ATCC 21245. J Biotechnol 221: 62–69.

Sayed, M., Dishisha, T., Sayed, W.F., Salem, W.M., Temerk, H.M., and Pyo, S.-H. (2017a) Enhanced selective oxidation of trimethylolpropane to 2,2-bis (hydroxymethyl) butyric acid using Corynebacterium sp. ATCC 21245. Process Biochem 63: 1–7.

Sayed, M., Sayed, W.F., Hatti-Kaul, R., and Pyo, S.-H. (2017b) Complete genome sequence of Mycobacterium sp. MS1601, a bacterium performing selective oxidation of polyls. Genome Announce 5: e00156-00117.

Sayed, M., Warlin, N., Hulteberg, C., Munslow, I., Lundmark, S., Pajalic, O., et al. (2020) 5-Hydroxymethylfururfural from fructose: An efficient continuous process in a water-dimethyl carbonate biphasic system with high yield product recovery. Green Chem 22: 5402–5413.

Serrano, A., Calviño, E., Carro, J., Sánchez-Ruiz, M.I., Cañada, F.J., and Martinez, A.T. (2019) Complete oxidation of hydroxymethylfururfural to furanarboxylic acid by aryl-alcohol oxidase. Biotechnol Biofuels 12: 1–12.

Serrano, A., Carro, J., and Martinez, A.T. (2020) Reaction mechanisms and applications of aryl-alcohol oxidase. In
Enzymes. Chaiyen, P., and Tamanoi, F. (eds). Amsterdam, the Netherlands: Elsevier, pp. 167–192.
Sousa, A.F., Coelho, J.F.J., and Silvestre, A.J.D. (2016) Renewable-based poly(ether)esters from 2,5-furandicarboxylic acid. Polymer 98: 129–135.
Sousa, A.F., Vilela, C., Fonseca, A.C., Matos, M., Freire, C.S.R., Gruter, G-J., et al. (2015) Biobased polyesters and other polymers from 2,5-furandicarboxylic acid: a tribute to furan excellency. Polym Chem 6: 5961–5983.
Stecher, G., Tamura, K., and Kumar, S. (2020) Molecular evolutionary genetics analysis (MEGA) for macOS. Mol Biol Evol 37: 1237–1239.
Sützl, L., Foley, G., Gillam, E.M., Bodén, M., and Haltrich, D. (2019) The GMC superfamily of oxidoreductases revisited: analysis and evolution of fungal GMC oxidoreductases. Biotechnol Biofuels 12: 1–18.
Sützl, L., Laurent, C.V., Abrera, A.T., Schütz, G., Ludwig, R., and Haltrich, D. (2018) Multiplicity of enzymatic functions in the CAZY AA3 family. Appl Microbiol Biotechnol 102: 2477–2492.
Tamboli, D.P., Telke, A.A., Dawkar, V.V., Jadhav, S.B., and Govindwar, S.P. (2011) Purification and characterization of bacterial aryl alcohol oxidase from Sphingobacterium sp. ATM and its uses in textile dye decolorization. Biotechnol Bioproc Eng 16: 661–668.
Tamura, K., Stecher, G., and Kumar, S. (2021) MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol 38: 3022–3027.
Troiano, D., Orsat, V., and Dumont, M.-J. (2020) Status of biocatalysis in the production of 2, 5-furandicarboxylic acid (FDCA). ACS Catal 10: 9145–9169.
Urlacher, V.B., and Koschorreck, K. (2021) Peculiarities and applications of aryl-alcohol oxidases from fungi. Appl Microbiol Biotechnol 105: 4111–4126.
Viña-Gonzalez, J., Martinez, A.T., Guallar, V., and Alcalde, M. (2020) Sequential oxidation of 5-hydroxymethylfurfural to furan-2,5-dicarboxylic acid by an evolved aryl-alcohol oxidase. Biochim Biophys Acta 1868: 140293. https://doi.org/10.1016/j.bbabap.2019.140293
Vinambres, M., Espada, M., Martinez, A.T., and Serrano, A. (2020) Screening and evaluation of new hydroxymethylfurfural oxidases for furandicarboxylic acid production. Appl Environ Microbiol 86: e00842-20. https://doi.org/10.1128/AEM.00842-20.
Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumieny, R., et al. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. Nucleic Acids Res 46: W296–W303.
Werpy, T., Petersen, G., Aden, A., Bozell, J., Holladay, J., White, J., et al. (2004) Top Value Added Chemicals from Biomass. Volume 1-Results of Screening for Potential Candidates from Sugars and Synthesis Gas. Washington, DC, USA: Department of Energy.
Whelan, S., and Goldman, N. (2001) A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. Mol Biol Evol 18: 691–699.
Yuan, H., Li, J., Shin, H.-D., Du, G., Chen, J., Shi, Z., and Liu, L. (2018) Improved production of 2,5-furandicarboxylic acid by overexpression of 5-hydroxymethylfurfural oxidase and 5-hydroxymethylfurfural/furfural oxidoreductase in Raoultella ornithinolytica BF60. Bioresour Technol 247: 1184–1188.
Yuan, H., Liu, H., Du, J., Liu, K., Wang, T., and Liu, L. (2020) Biocatalytic production of 2,5-furandicarboxylic acid: recent advances and future perspectives. Appl Microbiol Biotechnol 104: 527–543.
Zhang, J., Li, J., Tang, Y., Lin, L., and Long, M. (2015) Advances in catalytic production of bio-based polyether monomer 2, 5-furandicarboxylic acid derived from lignocellulosic biomass. Carbohydr Polym 130: 420–428.
Zhang, Z., and Deng, K. (2015) Recent advances in the catalytic synthesis of 2,5-furandicarboxylic acid and its derivatives. ACS Catal 5: 6529–6544.

Supporting information

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

Fig. S1. Sequence analysis of A0A1P8X5YS5 (MycspAAO). A) BLAST search of the sequence in the UniProt database, the retrieved 250 hits did not show any biochemically characterized protein. B) A0A1P8X5YS5 in red color was aligned with the characterized protein found in the CAZY AA3 proteins and it was clustered within the AA3_2 subfamily. AA3_1 subfamily sequences are colored green, AA3_2 subfamily are colored blue, AA3_3 are colored black, and AA3_4 are colored pink. C) A screenshot shows part of the whole alignment of A0A1P8X5YS5, highlighted in coral color, against 4220 fungal GMC oxidoreductases that have been recently classified into 5 clusters by Sützl et al. The A0A1P8X5YS5 was aligned under the A0A0H2S5F3, a putative aryl-alcohol oxidase from Schizopora paradoxa and within the AAO–PDH cluster according to Sützl et al (2019). Catalytic histidine is highlighted yellow, with the highest conservancy score (9). D) Comparison to characterized AAOs from fungal sources from Pleurotus eryngii, Pleurotus sapidus and Thermothelomycyes thermophilus. The multiple sequence alignment was done using PROTAL3SSD server.

Fig. S2. Nucleotide sequence of MycspAAO optimized gene for expression in E. coli.

Fig. S3. Comparison of MycspAAO described in the current study to the HMFOs characterized in literature, MethylHMO, PsenHMO, and PaeospHMFO. The numbered boxes show the positions where MycspAAO differs compared to the three HMFOs, the yellow highlighted sequence parts show the ADP-binding motif, PS000623 and PS000624 motifs. The table shows the specific differences in MycspAAO compared to the three HMFOs. Code for the alignment: secondary structure predictions (red: alpha-helix, blue: beta-strand). Consensus predicted secondary structure symbols: alpha-helix: h; beta-strand: e. Consensus amino acid symbols: conserved amino acids are in bold and uppercase letters; aliphatic (I, V, L); I: aromatic (Y, H, W, F); @: hydrophobic (W, F, Y, M, L, I, V, A, C, T, H); h: alcohol (S, T); o: polar residues (D, E, H, K, N, Q, R, S, T); p: tiny (A, G, C, S); t: small (A, G, C, S, V, N, D, T, P); s: bulky residues (E, F, I, K, L, M, Q, R, W, Y); b: positively charged (K,

© 2022 The Authors. Microbial Biotechnology published by Society for Applied Microbiology and John Wiley & Sons Ltd., Microbial Biotechnology, 15, 2176–2190
Conversion of HMF (4 mM) using the soluble fraction containing *MycspAAO-WT* from *Mycobacterium* sp. MS1601 expressed in *E. coli* BL21 (DE3) at 15, 20, and 30 °C, respectively, in (A) NYAT and (B) TB media. The experimental details are described in the text.

**Fig. S5.** The colorimetric assay results of *MycspAAO-WT* and -Y444F for the oxidation of 5 mM of different substrates after 30 min reaction time.

**Fig. S6.** Overall view of the homology model constructed for *MycspAAO*.

**Fig. S7.** A) Ramachandran plot for *MycspAAO*, B) Z-score of the homology model *MycspAAO* (UniProt accession code: A0A1P8X5Y5). The Z-score is <1, indicating a good model.

**Fig. S8.** The possible catalytic binding mode of the substrate HMF in the active site of *MycspAAO* 3D model (A) based on the docking experiments. The hydroxyl moiety of HMF is at hydrogen bond distance to the catalytic H445 and is probably stabilized by hydrogen bonding to Q335. The amino acids V95, L97, M337, N487, Y444, Y443, A317 are shaping the active site and are probably critical for substrate selectivity. T356 is a candidate for mutation because of its proximity to the substrate-binding site. (B) Another pose of the HMF in the active site that was detected in the docking experiments. The substrates (either 5-HMF or hydrated 5-HMF) are possibly involved in hydrogen bonding with Y444 that misorient the oxidizable Cα of the substrate against the N5-FAD atom. Mutating Y444 into F showed improved enzymatic activity.

**Fig. S9.** SDS-PAGE representing the purification of *MycspAAO*: wild type (WT), and variants Y444F, and L298A expressed in *E. coli* BL21 (DE3).

**Fig. S10.** Initial reaction rate of HMF oxidation at different concentrations using *Mycsp AAO* variants. The reaction was carried out at pH 8 and 30 °C for 3 hours. HMF and FFCA concentrations were measured by HPLC.

**Fig. S11.** Screening of the activity of purified wild type *MycspAAO*-WT (red) and *MycspAAO*-Y444F (blue) (8 µg/mL of total protein) for the oxidation of different substrates used at 5 mM concentration, pH 8, and 30 °C using the colorimetric coupled assay with horseradish peroxidase.

**Table S1.** Proposed mutations for *MycspAAO*, and their expected outcomes. The mutations that were not successfully obtained are marked with an asterix.

**Table S2.** List of primers used for site-directed mutagenesis.

**Table S3.** Protein concentrations of wild type *MycspAAO* and variants based on BCA assay, used for determination of activity with HMF.