Journal Pre-proof

Vanillyl alcohol oxidase from *Diplodia corticola*: Residues Ala420 and Glu466 allow for efficient catalysis of syringyl derivatives

Daniel Eggerichs, Nils Weindorf, Maria Laura Mascotti, Natalie Welzel, Marco W. Fraaije, Dirk Tischler

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
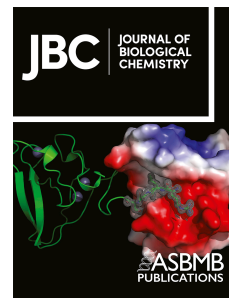
Vanillyl alcohol oxidase from *Diplodia corticola*: Residues Ala420 and Glu466 allow for efficient catalysis of syringyl derivatives

Daniel Eggerichs[1], Nils Weindorf[1], Maria Laura Mascotti[2,3], Natalie Welzel[1], Marco W. Fraaije[2], Dirk Tischler[1]*

[1] Microbial Biotechnology, Ruhr-University Bochum, Universitätsstraße 150, 44801 Bochum, Germany

[2] Molecular Enzymology, University of Groningen, Nijenborgh 4, 9747AG Groningen, The Netherlands

[3] IMIBIO-SL CONICET, Facultad de Química Bioquímica y Farmacia, Universidad Nacional de San Luis, Ejército de los Andes 950, D5700HHW San Luis, Argentina

*Corresponding author: Dirk Tischler, email: dirk.tischler@rub.de

Abstract

Vanillyl alcohol oxidases belong to the 4-phenol oxidases family and are found predominantly in lignin-degrading ascomycetes. Systematical investigation of the enzyme family at the sequence level resulted in discovery and characterization of the second recombinantly produced VAO member, *Dc*VAO, from *Diplodia corticola*. Remarkably high activities for 2,6-substituted substrates like 4-allyl-2,6-dimethoxy-phenol (3.5 ± 0.02 U mg^{-1}) or 4-(hydroxymethyl)-2,6-dimethoxyphenol (6.3 ± 0.5 U mg^{-1}) were observed which could be attributed to a Phe to Ala exchange in the catalytic center. In order to rationalize this rare substrate preference among VAOs, we resurrected and characterized three ancestral enzymes and performed mutagenesis analyses. The results indicate that a Cys/Glu exchange was required to retain activity for ɣ-hydroxylations and shifted the acceptance towards benzyl ethers (up to 4.0 ± 0.1 U mg^{-1}). Our findings contribute to the understanding of the functionality of VAO enzyme group, and with *Dc*VAO, we add a new enzyme to the repertoire of ether cleaving biocatalysts.

Introduction

Vanillyl alcohol oxidases (VAOs) are a group of the 4-phenol oxidases family which occur exclusively in ascomycetous fungi. These enzymes are hypothesized to be involved in degradation of aromatic compounds from natural sources like lignin breakdown or amino acid

transformation as the group of *vao* gene-containing ascomycetes prefers a saprophytic life-style, or are found to be plant-pathogens and endophytes.(1–4) Phylogenetically, fungal VAOs are related to bacterial eugenol oxidases (EUGOs) and 4-ethyl phenol oxidases (4EPOs).(2, 5) Until now, these three groups are characterized solely by a single member each where *Ps*VAO from *Penicillium simplicissimum*, *Rj*EUGO from *Rhodococcus jostii* RHA1 and *Gc*4EPO from *Gulosibacter chungangensis* represent the VAO, EUGO and 4EPO groups, respectively.(6, 7, 5) The putative VAO from *Byssochlamys fulva* V107 has been described as well, however no sequence information is available for this enzyme.(4)

Along with alcohol oxidation, dehydrogenation, hydroxylation and deamination reactions, *Ps*VAO also performs the cleavage of benzyl ethers which are indeed the proposed natural substrates from lignin degradation, as the expression of the corresponding encoding gene was shown to be induced by 4-hydroxy benzyl methyl ether.(8, 1) This hypothesis is corroborated by the subcellular localization of *Ps*VAO in the peroxisomes through its C-terminal peroxisomal targeting signal (PTS1), as these organelles are involved in the oxidative metabolism.(9) *Rj*EUGO and *Gc*4EPO share similar reaction profiles but substantial activity on ether substrates has been reported only for *Ps*VAO so far (Figure 1).(10, 5) This high substrate promiscuity made *Ps*VAO an interesting candidate for biocatalytic applications like the production of vanillin from capsaicin or 2-methoxy-4-methylphenol, and the stereoselective hydroxylation of 4-alkylphenol derivatives.(11–14) Based on detailed mechanistic studies and resolved crystal structures, some crucial active site residues have been identified.(15, 16) This allowed engineering of mutant enzymes displaying for example an inversion of the enantioselectivity for the hydroxylation of ethyl phenol.(17, 18)

The elucidation of the enzymatic mechanism revealed that in a first step, the covalently bound flavin adenine dinucleotide (FAD) cofactor is reduced by a hydride transferred from the phenolic substrate.(19, 11) As a consequence, a quinone methide intermediate is formed, which is a strong electrophile for the addition of a water molecule. Lastly, FAD is regenerated by molecular oxygen, yielding hydrogen peroxide as byproduct.(2)

Despite the high promiscuity of *Ps*VAO, little attention was drawn to other members of this enzyme family which is partially due to difficulties in expression and purification procedures in heterologous hosts yielding dysfunctional enzymes trapped in an intermediate oxidation state of the covalent flavin cofactor.(20) Yet, we decided to continue investigations on this enzyme family since the complexity of lignin degradation in dependency of the type of wood and the respective habitat of each fungus gives valid reason to expect larger functional diversity among other members than a single enzyme can show. We focused especially on ether cleavages as ethers in general, and the β-O-4 bond in particular, are the most abundant linkages in lignin.(21) Besides, ether cleavage itself is challenging from a chemical perspective and often

requires harsh methods, therefore we figured a deeper look into nature's toolbox may reveal potential biocatalysts.

In this work, we systematically evaluated VAO sequences to obtain candidates with altered amino acid compositions in the catalytic center, compared to *Ps*VAO, intending to find novel catalytic features. Based on this sequence analysis, we selected and successfully produced the VAO from *Diplodia corticola* (*Dc*VAO). This ascomycete fungus is called the "bot canker of oak" and represents a plant-pathogen widespread in eastern US and south-western Europe, responsible for infection of members of red and white oaks including the cork oak (*Quercus suber*).(22, 23) As *Dc*VAO contains a markedly different amino acid composition in the catalytic center compared to *Ps*VAO, we figured it might be an interesting candidate to get insight into the overall enzyme family and their potential role in plant biomass degradation. Therefore, we studied the substrate scope and kinetic parameters of *Dc*VAO and generated two point mutations to investigate the role of Glu466 in the catalytic center. Further, we performed an evolutionary analysis and resurrect the common ancestor of the whole clade of fungal VAOs plus some of the internal nodes. By this, we aimed to disclose the catalytic potential of the complete group.

Results

Genome mining and phylogenetic analysis allow identifying enzymes with altered amino acid compositions in the catalytic center

The sequence of *Ps*VAO was used to perform homology searches in fungal genomes resulting in about 200 sequences which cluster into five different clades (I to V)(Figure 2A). For a prediction of altered characteristics, we inspected the key residues in the catalytic center of *Ps*VAO in the multiple sequence alignment of the whole dataset (Figure 2B). Val185, Thr459 and Tyr187 (*Ps*VAO numbering) restrict the catalytic pocket at the side of the of the substrate entry tunnel, while Tyr187 has been hypothesized to work as the gatekeeper for oxygen access.(24) Asp170 is an essential residue for the substrate deprotonation in benzylic position and assists in covalent flavinylation.(25, 26) Thr457 and Cys470 are as well in close proximity to the benzylic position of the substrate molecule (Figure 2C) and probably modulate the overall reactivity. In the crystal structure of *Ps*VAO (PDB: 2VAO), *Rj*EUGO (5FXE) and *Gc*4EPO (7BPI), His422 (*Ps*VAO numbering) is responsible for a covalent linkage of the FAD cofactor to the enzyme and restricts, together with Phe424, the catalytic pocket at the opposing side of the substrate entry tunnel.(16, 27, 5) Notably, this Phe is replaced by an Gly residue in *Rj*EUGO.(27) With the exception of two enzymes, His422 is conserved among all the sequences analyzed. Further, the catalytic triad formed by Tyr108, Tyr503 and Arg504, which is responsible for the binding of the phenol group of the substrate, is also conserved among all VAOs (data not shown).(16, 28) The same is true for Tyr187 and Asp170 which were both

3

described as essential for the reaction mechanism. Differences in the family are rather found for secondary interactions with the substrate. Here, position 459 stands out as fairly variable. Thus, clade III and IV (*Ps*VAO) are distinct from the other clades by the presence of a Thr in position 459 over predominantly Val in clade I and Cys in clade V. In clade II, Thr and Val are common, while this clade also differs from clade III and IV by the occurrence of Glu in position 457 over otherwise Thr. In clade V, Ala is mainly found at this site.

Most important interaction points of substrates with groups *ortho* to the phenolic group, are residues in position 424 and 185 (*Ps*VAO numbering). Among the VAO group, the combination of Phe and Val is by far the most common. For position 424, exceptions are found in clade III and V. In clade III, only two VAOs, from *Diplodia corticola* (*Dc*VAO) and *Macrophomina phaseolina* (*Mp*VAO) respectively, contain an Ala in position 424. In contrast, Ala, Gly or Phe are found for enzymes of clade V with decreasing abundance. The introduction of another amino acid in position 424 often coincides with an amino acid change in position 470. These changes are found most of the time in clade V. In fact, *Dc*VAO and *Mp*VAO are the only two VAOs outside of clade V which do not contain a Cys in position 470. In general, the combination of Ala424 and Glu470 is found predominantly, but Leu, Met or in single cases Ser and Asn are also observed in position 470.

The combined occurrence of both amino acid variations raised the question how the catalytic properties of the respective enzymes are affected and which factors caused this unusual variation. Therefore, the habitats of selected fungi were included into the considerations. In this regard, *D. corticola* and *M. phaseolina* are both known plant-pathogens with an overlap in habitat on oak trees.(29, 23) A similar habitat was observed for *Daldinia childiae* which is a member of clade V, and also contains the amino acid combination of Glu470 and Ala424 (according to the homologous positions in *Ps*VAO). Thus, it is possible that the Phe/Ala exchange alters the catalytic pocket to accommodate for 2,6-dimethoxylated substrates as they occur in syringyl-rich S-lignin found in hardwood. As we intended to test this hypothesis, *Dc*VAO turned to our focus as *D. corticola* inhabits a narrow habitat of hardwood oaks. Further, the Cys/Glu exchange in position 470 alone is a remarkable change in the catalytic pocket (Figure 2D) and we were interested if this residue influences the substrate scope of the enzymes. Interested in ether cleavages, we expected an altered performance as Glu470 could act as additional proton shuttle. Based on these assumptions we selected *Dc*VAO for further experimental characterization.

_Dc_VAO can be produced heterologously in _E. coli_ as an FAD-saturated enzyme

The *vao* gene from *D. corticola* was acquired as a DNA fragment *via* gene synthesis and was successfully cloned into the pET16bp vector using the Gibson assembly method. Difficulties in producing fungal VAOs in bacterial hosts were described before and we also experienced issues in that aspect. Therefore, we tested four different *E. coli* expression strains and confirmed

the production of the enzyme by formation of vanillin and coniferyl alcohol in a whole-cell assay. *E. coli* NiCo21 (DE3) pKJE7 emerged as best suited host which encodes for three additional chaperones on the pKJE7 plasmid to promote correct folding of the protein.(30, 31) Overexpression was performed at 4.5 L scale in a Braun Biotech bioreactor and the obtained cells were used to provide a crude extract for subsequent protein purification. After His-tag affinity chromatography, a yellow enzyme was obtained with a yield of 9.1 mg per L broth. As reference, we also produced *Ps*VAO as described earlier.(32)

Spectral analysis of *Dc*VAO showed three peaks at 273, 360, 441 nm whereof the peak at 360 nm is characteristic for a covalent bound flavin cofactor (Figure S1).(33) We obtained a ratio 280 to 441 nm of ~9.4 which indicates full cofactor saturation. This was validated by comparison of the total protein content determined by Bradford analysis and the concentration for the FAD cofactor using the absorption at 441 nm (ε = 14,200 M^{-1} cm^{-1}). Size-exclusion chromatography revealed the protein to occur in a dimeric and an octameric complex (Figure S2). The same oligomerization state was described for *Ps*VAO which was attributed to the presence of a so-called octamerization loop which is present for all VAOs but absent for exclusively dimeric EUGOs and 4EPOs.(34) Our results strengthen this hypothesis. *Dc*VAO was found to be stable up to a temperature of 50°C after an incubation of 2 h at the respective temperature (Figure S3). Stability across a broad pH range from 6.5 to 9.5 (Figure S4) was observed. No activity could be determined at pH values of 4.5 and below while at pH of 10.5 a remaining activity of 13% was found.

*Dc*VAO accepts 2,6-dimethoxylated substrates and shows high activity for ether cleavages

Enzyme activity for *Dc*VAO and *Ps*VAO were measured in an endpoint kinetic assay, determining the hydrogen peroxide concentration in the solution (xylenol orange assay).(10) Product formation was validated by GC-MS. In total, 38 substrates were tested and both enzymes were found to be active on about two thirds of them (*Dc*VAO: 23, *Ps*VAO: 24) (Figure 3, Table S1). As described in literature, *Ps*VAO was found to perform hydroxylations in α- and γ-position, oxidation of benzyl alcohols, dehydrogenation as well as oxidative ether cleavages. The same reaction pattern was also found for *Dc*VAO (Table S2). Similar to *Ps*VAO, we detected a preference for hydroxylations on substrates (Figure S5 to S7). *Dc*VAO performed the hydroxylation of 1,2-bicyclic substrates and ethyl phenol derivatives, while for 4-cyclopentyl phenol (**35**) only dehydrogenation products were observed. Both enzymes performed rather poorly on deamination reactions, although *Ps*VAO has been described to be capable of performing this type of reaction.(14) Notable differences between the two enzymes were observed with regard to the substitution pattern *ortho* to the phenol moiety, as *Dc*VAO was found to accept 2,6-dimethoxylated substrates. Activities of 3.5 ± 0.02 U mg^{-1} and 6.3 ± 0.5 U mg^{-1} on

5

4-allyl-2,6-dimethoxy-phenol (**3**) and 4-(hydroxymethyl)-2,6-dimethoxyphenol (**6**) were observed, respectively, while for PsVAO only residual activities < 0.15 U mg^{-1} were detected. Further, DcVAO achieved high conversions on vanillyl ether derivatives with activities up to 4.0 ± 0.1 U mg^{-1} for ethyl vanillyl ether (**14**). We further investigated this behavior by kinetic studies for both enzymes. The K_M value for eugenol is similar for both enzymes and a cooperative effect according to Hill was observed for both as well (Table 1, Figure S8). The cooperative effect is less pronounced for DcVAO (n = 2.6 ± 0.2) as for PsVAO (n = 4.8 ± 0.6), which cannot be fully explained yet as both enzymes were shown to have the same multimerization state. It is likely that the substrate inhibition observed for PsVAO at high concentrations (K_i = 12.3 ± 4.5 mM) contributes to this effect. No inhibition was observed for DcVAO. Further, DcVAO has an about two-fold increased K_M values for vanillyl alcohol and vanillyl ether derivatives and was found to be more active on vanillyl ethyl (**14**) and vanillyl butyl ether (**15**)(Figure S9 to S11). For the later, no activity was observed for PsVAO while DcVAO had a turnover frequency of 2.1 ± 0.1 s^{-1} (Figure S12). On the other hand, PsVAO was found to be more active on vanillyl methyl ether (**13**).

The Michaelis-Menten kinetics for DcVAO and PsVAO showed that DcVAO has overall higher K_M values for all substrates except for eugenol. Regarding the k_{cat}, PsVAO has a higher turnover frequency on eugenol (**2**), vanillyl alcohol (**5**) and vanillyl methyl ether (**13**) while DcVAO is twice as active on vanillyl ethyl ether (**14**) and the only active enzyme on vanillyl butyl ether (**15**). This highlights a preference of DcVAO for longer chains compared to PsVAO.

As the highest activity of DcVAO was observed for vanillyl ethyl ether, also the highest total turnover number (> 13,500 TTN) was reached for this substrate while for vanillyl methyl and butyl ether a lower initial rate and a lower TTN was observed (> 9,200 and > 7,300 TTN, respectively, Figure S13).

Structural analysis reveals key residues influencing the substrate selectivity

DcVAO was found to accept 2,6-dimethoxylated substrates and to be highly active on vanillyl ethyl derivatives. In order to investigate the molecular origins for these characteristics, we built a homology model for DcVAO based on the PsVAO structure (PDB 2VAO). In addition, a *de novo* model with AlphaFold2 was generated and compared to the previously obtained homology model and the crystal structure of PsVAO. The three structures aligned well, so that we decided to continue with the homology model for further calculations. Subsequently, a structure guided analysis of the catalytic center by docking several substrates to the catalytic cavity of DcVAO was performed and substrate interactions with the amino acids of the active side were evaluated. Although the majority of amino acids is the same for both enzymes (Figure 2C and D), notable differences were observed in the positioning of the substrate molecule: while in PsVAO the methoxy group of 4-ethylguaiacol is directed towards Val185, all docked substrates are turned in DcVAO (Figure S14) and for vanillyl ethyl ether, the single methoxy group is

6

rather pointed in the direction of Ala420 (Figure S15). As already pointed out in the phylogenetic analysis, most VAO sequences contain a Phe in position 424 while *Dc*VAO contains an Ala instead (A420). This amino acid change widens the catalytic pocket and generates space for 2,6-dimethoxylated substrates (Figure S16 and S17), which is in agreement with the higher K_M values observed for *Dc*VAO on 2-methoxylated substrates. It can be speculated, if this amino acid variation is related to a habitat containing high amounts of syringol (4-hydroxy-3,5-dimethoxycinnamic alcohol), as it is the case for *D. corticola*, or if other structural reasons are the determining factor. Nevertheless, the additional space in the catalytic pocket is likely the reason for single methoxylated substates to adopt a different binding pose compared to *Ps*VAO.

In position 466 (470 in *Ps*VAO), *Dc*VAO features a Glu over the more common Cys among other members of the family. This residue is located in close proximity to the *para* residue of the substrate. Docking of 4-allyl-2,6-dimethoxy-phenol (**3**) revealed that Glu466 is also likely to favor ɣ-hydroxylations as the O_{Glu}-$C_ɣ$ distance was determined to be 3.7 Å (Figure 4A). For 4-(hydroxymethyl)-2,6-dimethoxyphenol (**6**), a distance of 4.8 Å was found for the O_{Glu}-$O_β$ and O_{Glu}-$C_α$ interactions (Figure 4B). Similar values were measured for the O_{Glu}-$C_α$ distance (4.7 Å) in the *Dc*VAO homology model docked with vanillyl ethyl ether (**14**) while the O_{Glu}-$O_β$ distance was 0.6 Å shorter (4.1 Å)(Figure 4C). In contrast, in the crystal structure of *Ps*VAO (PDB 2VAO), a S-$C_α$ distance of 5.5 Å was measured (Figure 4D). The positioning of Glu466 leaves the residue in a prime position to possibly act as a proton shuttle for hydroxylation reactions and may also facilitate ether cleavages. To investigate this hypothesis, we replaced Glu466 by a Cys and a Leu in *Dc*VAO. Cys is the most common amino acid for related VAOs, and also present in *Ps*VAO, while Leu occurs naturally in EUGOs and 4EPOs for which little to no activity for ether cleavages was reported (0.004 s^{-1} for *Rj*EUGO).(7)

Variation of the residue in position 466 is probably a result from the repositioning of the substrate molecule

*Dc*VAO variants were generated by means of the QuikChange method and proteins were produced according to the same protocol as for the wildtype enzyme. In comparison, the variants were produced in lower yields (*Dc*VAO E466C: ~3 mg L^{-1}, *Dc*VAO E466L: ~1 mg L^{-1}). The substrate scope of the variants was investigated by using the same endpoint kinetic measurements as described before. In total, 33 substrates were tested as we excluded five from the initial set since no conversion was observed neither for *Dc*VAO or *Ps*VAO. For *Dc*VAO E466C and E466L, a significantly altered substrate profile was found: *Dc*VAO E466C was found to retain 50 to 100% activity for the cleavage of vanillyl ethers where the percentage increased with the chain length (Figure 3B). Unchanged conversion was observed for 5-hydroxyindane (**37**), 5,6,7,8-tetrahydro-2-naphtol (**38**) and 3,4-dihydroxybenzylalcohol (**7**). For other benzyl

alcohol derivatives, like vanillyl alcohol (**5**) the residual activity was 49%. A stronger reduction in activity was observed for hydroxylation reactions. 4-Ethylguaiacol (**20**) represents the only converted ethyl phenol derivative with a residual activity of 12%. Even lower activities were observed for the hydroxylation of allyl phenol derivatives resulting for this group in the overall strongest decrease.

For *Dc*VAO E466L, an overall reduced activity was found as well. For *p*-allyl substrates, a more than ten-fold reduction was observed while for *p*-benzyl alcohol derivatives activities of 17% and lower were found compared to the wildtype enzyme. A residual activity of about 20% was found for vanillyl methyl (**13**) and ethyl ether (**14**) while no conversion could be detected for vanillyl butyl ether (**15**). The highest activity on ether substrates was observed for the ethyl derivative with 0.76 ± 0.05 U mg^{-1}. Notably, a comparably lower reduction was found for catechol-like substrates. For 3,4-dihydroxybenzyl alcohol (**7**) and ethyl catechol (**19**), the residual activity of 90% and 61%, respectively is rather high compared to other substrates. In case of methyl catechol (**24**), an activity of 0.83 ± 0.10 U mg^{-1} was observed which represents an increase of 182% compared to the wildtype. As *Ps*VAO is active on catechol-like substrates as well, it can be hypothesized that a hydrophobic interaction next to His422 (*Ps*VAO numbering) might be beneficial for acceptance of these substrates. In case of *Dc*VAO E466L, the respective Leu would perform this task, while Phe424 interacts with the substrate in *Ps*VAO.

The tendency in decreasing activity from the wildtype over the E466C to the E466L variant highlights the importance of a polar interaction from residue 466 with the substrate in *Dc*VAO. The strongest impact was found for γ-hydroxylations which is in good agreement with the close O_{Glu}-C_{γ} distance. These results contrast with observations for *Ps*VAO, as no impact on the activity on eugenol (**2**) or vanillyl alcohol (**5**) were observed in earlier studies with the C470L variant.(10) But as *Ps*VAO acts on *p*-allyl phenols with up to 17.2 ± 0.2 U mg^{-1} for eugenol (**2**) and contains a Cys in the respective position, clearly other factors determine the substrate acceptance in *Dc*VAO. The changed positioning of the substrate observed in the *in silico* analysis is the most likely explanation for this (Figure S14). This raised the question about when and how the change in substrate orientation occurred. We therefore decided to resurrect three ancestral enzymes to attempt responding this.

Resurrection of ancestral VAOs helps in rationalizing substrate selectivity

To determine the potential substrate scope of the VAO family as well us to understand the sequence determinants of it, ancestral reconstruction was conducted. The common ancestors of the clades including *Ps*VAO (Anc*Ps*VAO, $\overline{PP}$: 0.94) and the one including *Dc*VAO (Anc*Dc*VAO, $\overline{PP}$: 0.94) were reconstructed along with the ancestor of the whole group (AncVAO, $\overline{PP}$: 0.87)(Figure 5A). Genes of ancestral enzymes were ordered as synthesized gene fragments, cloned and expressed in the same way as for the *Dc*VAO gene. Enzyme yields

were generally lower compared to *Dc*VAO. AncVAO and Anc*Ps*VAO were produced with 2.3 and 2.2 mg per L broth, respectively, while 0.3 mg per L broth of Anc*Dc*VAO were obtained. In total, 37 substrates were tested for conversion by the ancestral enzymes (Table S3). To better trace changes in selectivity of each enzyme during evolution, the substrates were grouped into eight categories of chemically similar compounds. The percentage of converted substrates for each category gives the substrate profile of the enzyme (Figure 5B). Ultimately, homology models were calculated for the ancestral proteins to correlate the substrate profile with changes in the amino acid composition.

Over the course of evolution, it seems plausible to speculate that the substrate profile of VAOs first diversified and later specialized as it is the case of modern enzymes. AncVAO shows a narrow selectivity (12 compounds) towards *p*-allyl phenols, benzyl alcohols and amines while ether and alkyl substrates were rarely accepted (Figure 5B). The highest activities of ~1 U mg^{-1} are observed for chavicol (**1**), eugenol (**2**), vanillyl alcohol (**5**), and 4-(1-hydroxyethyl)phenol (**9**). As this enzyme is the closest relative to EUGOs and 4EPOs, it is not surprising that the observed substrate profile is similar to that of *Rj*EUGO which also performs poorly in ether cleavage reactions and hydroxylations of cresol derivatives.(10) This is also reflected in the structural model (Figure S18A): Like *Rj*EUGO, AncVAO harbors a Gly in position 424 and a Val in position 459 (*Ps*VAO numbering). Docking experiments revealed that these residues cause the model substrate eugenol to adopt an alternative binding position which again is comparable to the binding mode observed in the crystal structure of *Rj*EUGO (Figure S18B). Thus, the positioning of the substrate might be the reason for the EUGO-like substrate scope of AncVAO. The fact that AncVAO contains a Cys in position 470 (PsVAO numbering) while in the structure of *Rj*EUGO a Leu is found, seems not to influence the substrate scope.

Anc*Ps*VAO and Anc*Dc*VAO have both a broader substrate scope than AncVAO (19 and 16 compounds, respectively) but do not accept many substrates in the respective categories where *p*-allyl phenols and benzyl alcohols are the exception. Both enzymes contain a Phe in position 424 (*Ps*VAO numbering) and docked eugenol adopts a similar binding mode as observed for *Ps*VAO (Figure S19 and S20A). Further, Anc*Ps*VAO and Anc*Dc*VAO feature a Cys in position 470. With regard to the reaction profile, first tendencies are visible which carry over to the clades later on: Anc*Ps*VAO converts bicyclic compounds and such with longer alkyl chains. Also, the ether cleavage reaction is introduced in the reaction profile which makes Anc*Ps*VAO the most promiscuous ancestral enzyme. *Ps*VAO substrate profile contains 22 converted substrates where a clear selectivity towards longer chains in *para*-positions can be observed (Figure 5B and C). While 40 to 50% of *p*-alkyl substrates (chain length >1) were accepted by Anc*Ps*VAO, the acceptance increases to 60% for *p*-ethyl to 100% for *p*-alkyl substrates with a chain length >2 in *Ps*VAO. Further, all cyclic substrates were accepted, which contain a steric demanding group in *para*-position as well. Anc*Ps*VAO is the only ancestor

9

converting 4-hydroxy benzyl methyl ether indicating that this selectivity was adopted for clade IV during evolution, thus strengthening the assumption that this compound might be the natural substrate for *Ps*VAO.(1) For *p*-allyl phenols, benzyl alcohols and amines, the acceptance of *Ps*VAO remained similar to Anc*Ps*VAO. Structure-wise, the catalytic center of Anc*Ps*VAO and *Ps*VAO contain the same amino acid composition (Figure S19) which is in agreement with the similar substrate profile. As a consequence, the higher activity observed for *Ps*VAO might be due to different structural reasons.

For *Dc*VAO, the same tendency for the chain length of *p*-alkyl substrates was observed as for *Ps*VAO. But apart from this, the enzyme took a different path. *Dc*VAO is rather versatile in ɣ-hydroxylations and the oxidation of benzyl-alcohol derivatives as the enzyme tolerates many residues in the *ortho*-positions of the substrate. This is partially reflected in clade III as Anc*Dc*VAO was found to accept a variety of benzyl alcohols with activities ranging between <0.1 to 0.31 U mg^{-1}. This includes 3-bromo-4-hydroxy benzyl alcohol (**8**, 0.16 ± 0.04 U mg^{-1}) which was not converted by any other ancestral enzyme. Despite reasonable acceptance for *p*-allyl substrates, observed activities for Anc*Dc*VAO were below 0.1 U mg^{-1} for eugenol (**2**) and 4-allyl-2,6-dimethoxy-phenol (**6**). As Anc*Dc*VAO possesses a Cys in position 466 (*Dc*VAO numbering), the enzyme is more comparable to the *Dc*VAO E466C variant (Figure S20). Thus, the low activity on allyl phenols is in agreement with the mutagenesis studies. Further, Anc*Dc*VAO is the most active ancestral enzyme on vanillyl ethyl ether (**14**) being the only one converting vanillyl butyl ether (**15**)(0.23 ± 0.16 and 0.29 ± 0.11 U mg^{-1} respectively). This selectivity is in line with the substrate scope of wt *Dc*VAO and E466C which both accepted 100% of all benzyl ether compounds. A notable difference between Anc*Dc*VAO and *Dc*VAO is the ability of the ancestral enzyme to perform deamination reactions. As this activity could be restored for *Dc*VAO E466C on 4-(1-aminoethyl)phenol (**10**)(0.1 ± 0.03 U mg^{-1}), it is tempting to attribute the loss of this reaction type to the introduction of Glu466 which would strongly interact with the amine group. Moreover, Glu466 seems to be mainly responsible for hydroxylation reactions in α- and ɣ-position as Anc*Dc*VAO and *Dc*VAO E466C were found to possess only residual activities on those substrates. But Glu466 is not solely responsible for the oxidation of alcohol or the cleavage of ether. For the later, the reducing effect in activity can even be offset by longer chain length in *para*-position as observed for butyl vanillyl ether (**15**). This is supported by *in silico* experiments as the distance between the reactive oxygen of Glu466 is closer to the ɣ-carbon atom than to the α-carbon atom of the substrate molecule. Another factor is the orientation of the substrate molecule in the catalytic pocket. Due to the introduction of Ala420, additional space is generated causing the substrates to rotate compared to *Ps*VAO (Figure S14). Thus, it is likely that the Cys/Glu exchange is necessary to retain hydroxylating activity after adoption towards 2,6-methoxylated substrates by the introduction of Ala420. This hypothesis is strengthened by the fact that both mutations occur only in combination.

10

Nevertheless, the fact that AncDcVAO harbors a Phe in the respective position highlights that additional factors in protein folding may play a role as well.

Discussion

In this study, we describe the second member of the fungal VAO family, the enzyme from *D. corticola*, referred to as *Dc*VAO. The enzyme was selected after phylogenetic analysis of the vanillyl alcohol oxidases. Special attention was drawn to its amino acid composition in the catalytic center. *Dc*VAO differs from the majority of VAO group members in the two positions Ala420 and Glu466 (*Dc*VAO numbering). Ala420 widens the space in the catalytic pocket for substrates with moieties positioned *ortho* to the phenolic group. This became obvious comparing the substrate selectivity of *Dc*VAO with the already described VAO from *P. simplicissimum* (*Ps*VAO) containing a Phe in the respective position, as do most VAOs. *Dc*VAO reached up to 6.3 U mg^{-1} on 2,6-substituted substrates like 4-(hydroxymethyl)-2,6-dimethoxyphenol (**6**) while *Ps*VAO was barely active. Since the Ala exchange occurs naturally combined with the exchange of Cys in position 466, we further investigated this change in expectation of an improving effect on ether cleavages. Mutagenesis studies indicated that the introduced Glu466 in *Dc*VAO is important for hydroxylation in α- and ɣ-position, while ether cleavage and oxidation of benzylic alcohols was reduced but still possible in the E466C variant. The fact that *Ps*VAO is active on these compounds, although containing a Cys in the same position, highlighted that other factors might be involved as well. *In silico* experiments revealed a changed orientation of the substrate molecule in the catalytic pocket of *Dc*VAO compared to *Ps*VAO which is required to accommodate for 2,6-substituted compounds. As a consequence, Glu466 restores the original substrate scope of the enzyme. This observation was further supported by the analysis of the resurrected ancestors, on which AncDcVAO has a similar profile to that of *Dc*VAO E466C variant. While AncDcVAO was found to have low activities in ɣ-hydroxylation reactions, the enzyme was able to perform deamination reactions. This reaction type was absent for *Dc*VAO but could be restored in the *Dc*VAO E466C variant.

It is tempting to attribute the wider catalytic pocket and the acceptance of 2,6-dimethoxylated substrates to the habitat of *D. corticola* as the fungus is a known pathogen to oaks. These trees contain hardwood which is rich in 2,6-substituted syringyl lignin. Among the phylogenetic tree, some other examples are found for which the fungal habitat coincides with the amino acid composition in the catalytic center: *M, phaseolina* (clade III) and *D. childiae* (clade V) are both plant-pathogens infecting oak trees and contain the amino acid combination of Glu470 and A424, like *Dc*VAO.(29, 37) Further, plant-pathogenic fungi like *Sphaerulina musiva* or *Dothistroma septosporum* with a habitat on softwood trees like conifers, pine or poplar are found in clade I and harbor the amino acid combination of Cys470 and Phe424 which are the

11

same as found in *Ps*VAO.(38, 39) However, expression of these enzymes and subsequent characterization *in vitro* is required to validate this hypothesis drawn from the sequence data. To summarize the knowledge obtained about the evolution of the VAO family, we can state that AncVAO shows a substrate profile more similar to that of the bacterial EUGO from *R. jostii* RHA1 and 4EPO from *G. chungangensis* as it could be expected from the evolutionary history. From there on, Anc*Ps*VAO evolved to accommodate substrates with large groups in *para*-position of the phenolic moiety like bicyclic compounds and such with longer alkyl chains. In contrast, Anc*Dc*VAO was found to have a high acceptance towards *p*-allyl, ether and benzyl alcohols substrates although the activity for ɣ-hydroxylations is rather poor. These features are in line with the high activity observed for *Dc*VAO on these compounds where ɣ-hydroxylation was restored by the introduction of Glu466. Thus, we observe a varying degree of specialization of VAO enzymes over time, which highlights the catalytic versatility of this protein family.

Experimental procedures

Phylogenetic analysis and ancestral sequence reconstruction

Sequence dataset was constructed by homology searches in BLASTp using as query *Ps*VAO (Genbank accession P56216). From the raw data set, redundancies and none-fungal sequences were removed. As outgroups, the sequences from EUGO from *R. jostii* RHA1 (*Rj*EUGO) and 4EPO from *G. chungangensis* (*Gc*4EPO) were included. A multiple sequence alignment was constructed by means of the ClustalW algorithm in the Mega 11 software.(40) A phylogenetic tree was calculated from the alignment using neighbor joining algorithm in Mega 11 and EMBL online services.(41, 42)

Ancestral sequence reconstruction was performed in PAMLX v.4.9 as marginal reconstruction using CODEML.(43) For this, a maximum likelihood phylogeny was inferred in RaxML v8.2.10 with 500 rapid bootstrapping employing a MSA constructed in MAFFT V7 (292 seqs, 652 sites).(44, 45) Sequences were analyzed using an empirical substitution matrix, an empirical equilibrium amino acid frequencies, 4 gamma categories and LG substitution matrix. The posterior probability distribution of the ancestral states at each site was analyzed at nodes corresponding to AncVAO, Anc*Ps*VAO and Anc*Dc*VAO. The length of the ancestors (AncVAO and Anc*Ps*VAO 571 amino acids and Anc*Dc*VAO 570 amino acids) was treated by parsimony analyzing the presence/absence of gaps in the targeted nodes on the basis of the length of the derived sequences in each clade.(46) The overall posterior probabilities ($\overline{PP}$) for each targeted ancestor were determined as the mean value of the PPs per site. Sites were considered ambiguously reconstructed when the alternative states displayed PP > 0.2.

Homology modeling and structural analysis

For structural analysis of *Dc*VAO, a homology model calculated by YASARA and an AlphaFold2 structure were generated using the sequence from the NCBI database (Accession number XP_020127612).(47–49) As template for homology modeling, the crystal structure of *Ps*VAO (PDB 2VAO) was used as both enzymes share 62.5% sequence identity. The structures were inspected using PyMol.(50) Autodocking was performed in YASARA using the VINA autodock software.(51) The structures were refined by performing molecular dynamics

simulation for 5 ns, using the YASARA2 force field, and creating a water-filled periodic simulation box 5 Å around the structures.

Gene synthesis and cloning

All sequences were codon optimized for *E. coli* and synthesized as a DNA fragment by Twist Bioscience. The genes were cloned blunt end into pJET2.1 vectors using the New England Biolabs kit. After sequencing, genes were cloned into the pET16bp vector by means of the Gibson assembly.(52) Primers are listed in Table S4. The success of the PCR was controlled by agarose gel electrophoresis. The Gibson assembly mix was digested with DpnI for 1.5 hours at 37°C before being transformed in competent *E. coli* DH5a cells, which were plated on LB agar containing 100 µg ml^{-1} ampicillin. Clones were picked the next day and grown in a 5 mL LB medium (10 g L^{-1} tryptone, 10 g L^{-1} NaCl, 5 g L^{-1} yeast extract, 100 µg ml^{-1} ampicillin) overnight. Plasmids were isolated by the use of Macherey-Nagel NucleoSpin Plasmid Mini Kit and the success of cloning was validated by sequencing.

Generation of protein variants

For the generation of *Dc*VAO mutants, QuikChange-mutagenesis was applied using the PrimeSTAR Max DNA-polymerase from Takara Bio with 25 µl of total reaction volume, 100 ng of pET16bp_DcVAO plasmid as template and 10 pmol of the respective primer. Primers were designed for the respective mutation targets as fully and partially overlapping pairs (Table S5). In total, 30 PCR-cycles were applied using an annealing temperature recommended for the primer. The success of the PCR was controlled by agarose gel electrophoresis. The PCR-product was digested with DpnI for 1.5 hours at 37°C before being transformed in competent *E. coli* DH5a cells which were plated on LB agar containing 100 µg ml-1 ampicillin. Clones were picked the next day and grown in a 5 mL LB medium (10 g L^{-1} tryptone, 10 g L^{-1} NaCl, 5 g L^{-1} yeast extract, 100 µg ml^{-1} ampicillin) overnight. Plasmids were isolated by the use of Macherey-Nagel NucleoSpin Plasmid Mini Kit and mutations were validated by sequencing.

Screening of the expression host

In an initial experiment, the plasmid encoding for *Dc*VAO was transformed into *E. coli* Rosetta (DE3), *E. coli* SHuffle (DE3), *E. coli* NiCo21 (DE3) pKJE7 and *E. coli* BL21 (DE3). A 50 mL culture of autoinduction medium (12 g L^{-1} tryptone, 24 g L^{-1} yeast extract, 100 mM potassium phosphate buffer pH 7.0, 0.5 g L^{-1} glucose, 2 g L^{-1} lactose, 100 µg ml^{-1} ampicillin) was inoculated by a few colonies and grown at 37°C for 4 h before the temperature was reduced to 20°C for another 20 h of incubation. For cells harboring the pKJE7 plasmid, the lactose was replaced by 0.5 g L^{-1} arabinose and protein expression was induced by a final concentration of 1 mM IPTG after four hours. The cells were harvested, washed once with 100 mM potassium phosphate buffer at pH 7.0 before a final OD$_{600}$ of 30 was established. 500 µL cell suspension was incubated for 2 h in 1.5 mL reaction tubes containing 2 mM substrate. GC-MS samples were prepared as described below to identify product formation.

Protein production and purification

From a 100 mL preculture in LB medium, 4.5 L of autoinduction medium (12 g L^{-1} tryptone, 24 g L^{-1} yeast extract, 100 mM potassium phosphate buffer pH 7.0, 0.5 g L^{-1} glucose, 0.5 g L^{-1} arabinose, 100 µg ml^{-1} ampicillin) in a 5 L fermenter were inoculated to a starting OD$_{600}$ of 0.05. The cultures were stirred at 300 rpm at 37°C for 4 h before the temperature was reduced to 20°C and protein expression was induced by addition of a final concentration of 1 mM IPTG upon reaching 20°C. The cultures were incubated for another 20 h. During the whole time, 2 L min^{-1} air were supplied. The final OD$_{600}$ ranged between 8 and 10. The cultures were harvested at 5,000 x g for 20 minutes at 4°C. Cell pellets were washed in 100 mM potassium

13

phosphate buffer pH 7.0 and centrifuged again. The cell pellets were either directly used for cell lysis or stored at -20°C for further use.

For lysis, cells were resuspended in buffer A (10 mM Tris-HCl pH 7.5, 500 mM NaCl) and lysed by sonication. The lysate was centrifuged at 20,000 x g for 1.5 hours at 4°C. The supernatant was collected, filtered through a 0.2 µm pore and applied to an Äkta Start FPLC system from GE Healthcare. Affinity chromatography was performed using a 5 mL Ni-NTA-HisTrap column from GE Healthcare. The protein was eluted at 100% buffer B (10 mM Tris-HCl pH 7.5, 500 mM NaCl, 500 mM imidazole) while before a washing step of 20% buffer B in buffer A was performed. During the whole process, the absorbance at 280 nm was monitored to collect protein containing fractions. The buffer was exchanged to 50 mM potassium phosphate buffer pH 7.5 by means of dialysis overnight at 8°C. Afterwards, the dialyzed proteins were concentrated by ultrafiltration (Amicon 10 KDa cut-off) and stored with a final concentration of 60% glycerol at -20°C until further use.

Size exclusion chromatography

Analytical size exclusion chromatography was carried out using a Superdex 200 Increase 10/300 GL column (Cytivia) in 50 mM sodium phosphate at pH 7.2 containing 150 mM sodium chloride. A constant flow rate was applied using a Knauer Azura FPLC system. 150 µl of a 25 µM protein solution was injected to the column. As reference, the Cytiva high molecular weight gel filtration calibration kit was used. Detection was performed at 280 nm.

Determination of protein concentration

The absorption of the covalently bound FAD cofactor was used to measure the amount of active protein in the solution. Measurement was carried out at 441 nm ($\varepsilon = 14,200$ M^{-1} cm^{-1}). (6, 53) The total amount of protein was determined by Bradford assay using the RotiQuant reagent from Roth. Protein dilutions ranging from 10 to 1000-times were prepared in a volume of 200 µl containing a final concentration of 1x RotiQuant. The solution was incubated for 30 min in the dark before the absorption was measured at 595 nm. A standard curve with bovine serum albumin was used for protein quantification.

Xylenol orange assay for enzyme activity measurements

To obtain kinetic data of the enzymes, a 96-well plate-based assay was used to photometrically determine the production of hydrogen peroxide, which is a stochiometric byproduct of VAOs. A variation of this assay was applied before,(10) however, the sample points were increased from two to three to improve data quality. During the assay, triplicates of 100 µL, containing a final concentration of 50 mM potassium phosphate buffer pH 7.5, a varying amount of substrate, and 50 nM of the respective enzymes were incubated and shaken at 25 °C on a 96-well-plate. The reactions were simultaneously started by the addition of enzymes. After 3, 6, and 9 minutes, 20 µl of each reaction were transferred onto a new 96-well-plate, already containing 180 µl of detection solution (250 µM $FeSO_4$, 25 mM H_2SO_4, 100 µM Xylenol-Orange). In presence of hydrogen peroxide, the absorption of the dye shifts from 440 nm to 560 nm which was measured after incubation for 30 min. The amount of hydrogen peroxide was determined by a calibration curve and the slope though all three measurement points was used to calculate the enzyme activity.

pH stability assay

The activity assay was performed similar to the xylenol orange assay but using Briton-Robinson buffer (40 mM acetic acid, 40 mM phosphoric acid, 40 mM boric acid) in 0.5 pH steps while vanillyl methyl ether at a final concentration of 2 mM was used for all reactions. For pH stability,

14

the enzyme stock solution was incubated in Briton-Robinson buffer for two hours before the reaction described in the xylenol orange assay was performed.

Temperature stability assay

The activity assay was performed similar to the xylenol orange assay but the enzyme aliquots were incubated at the respective temperature for two hours before they were cooled on ice. Vanillyl methyl ether was used at a final concentration of 2 mM for all reactions.

Total turnover assay

The total turnover was determined by measurement of the reaction product vanillin at 350 nm. The reaction was performed in triplicates of 200 µl in a 96 well plate containing a final concentration of 50 mM potassium phosphate buffer pH 7.5, 10 mM of substrate, and 5 nM of the respective enzyme. The plate was incubated and shaken at 30°C and the absorption was measured in 15 min. intervals. The turnover was calculated as the produced product molecules per active site.

GC-MS measurements

To identify the formed products of the converted substrates, triplicates of 400 µl in 1.5 ml tubes of each substrate were used, containing 2 mM of substrate, 50 mM of potassium phosphate buffer pH 7.5 and 500 nM of the respective enzyme. The reactions were shaken at 25 °C and 750 rpm for 16 h. For sample preparation, 300 µL of the solution were extracted with the same volume of ethyl acetate by vortexing for 1 min. Phase separation was initiated by centrifugation for 5 minutes at 17,000 x g and room temperature. Afterwards, the organic phase was dried over anhydrous $MgSO_4$ and 100 µl were used for analysis. On a Shimadzu Nexis GC-MS 2030, a 5°C min^{-1} gradient from 150 to 200°C was applied followed by a 5 min hold at 200°C using a FS-Supreme 5ms column (length: 30 m, inner diameter 0.25 mm, outer diameter 0.36 mm). Compounds were identified by NIST 2017 library search.

Data availability
All data will be available upon request.

Supporting Information
This article contains supporting information.

Acknowledgement

Author contributions

DE analyzed sequence, phylogenetic and structural data, performed molecular dynamic simulations, evaluated the substrate scope of ancestral enzymes, and wrote the original draft. NW cloned the genes in expression vectors and introduced point mutations, generated the homology models, supported the structural analysis, measured activity data, and performed GC-MS measurements. MLM performed the ancestral sequence reconstruction and supported writing. NW measured additional enzyme activity and kinetic properties, introduced point mutations, and performed total turnover experiments. MWF supported conceptualization the project and supported writing. DT supervised all lab work, conceptualized the project, and supported writing. All authors approved the final manuscript prior submission.

Funding

Conflict of interest

The authors declare that they have no conflicts of interest with the contents of this article.

References

1. Fraaije, M. W., Pikkemaat, M., and van Berkel, W. J. H. (1997) Enigmatic Gratuitous Induction of the Covalent Flavoprotein Vanillyl Alcohol Oxidase in *Penicillium simplicissimum. Applied and Environmental Microbiology* **63**, 435–439 10.1128/aem.63.2.435-439.1997 PMID 16535508

2. Gygli, G., de Vries, R. P., and van Berkel, W. J. H. (2018) On the Origin of Vanillyl Alcohol Oxidases. *Fungal Genetics and Biology* **116**, 24–32 10.1016/j.fgb.2018.04.003 PMID 29626635

3. Ewing, T. A., Gygli, G., Fraaije, M. W., and van Berkel, W. J. H. (2020) Vanillyl Alcohol Oxidase. *The Enzymes* **47**, 87–116 10.1016/bs.enz.2020.05.003 PMID 32951836

4. Furukawa, H., Wieser, M., Morita, H., Sugio, T., and Nagasawa, T. (1999) Purification and Characterization of Vanillyl Alcohol Oxidase from *Byssochlamys fulva* V107. *Journal of Bioscience and Bioengineering* **87**, 285–290 10.1016/S1389-1723(99)80033-4 PMID 16232469

5. Alvigini, L., Gran-Scheuch, A., Guo, Y., Trajkovic, M., Saifuddin, M., Fraaije, M. W., and Mattevi, A. (2021) Discovery, Biocatalytic Exploration and Structural Analysis of a 4-Ethylphenol Oxidase from *Gulosibacter chungangensis. Chembiochem: A European Journal of Chemical Biology* **22**, 3225–3233 10.1002/cbic.202100457 PMID 34523783

6. de Jong, E., van Berkel, W. J. H., van der Zwan, R. P., and de Bont, J. A. (1992) Purification and Characterization of Vanillyl Alcohol Oxidase from *Penicillium simplicissimum*. A Novel

Aromatic Alcohol Oxidase Containing Covalently Bound FAD. *European Journal of Biochemistry* **208**, 651–657 10.1111/j.1432-1033.1992.tb17231.x PMID 1396672

7. Jin, J., Mazon, H., van den Heuvel, R. H. H., Janssen, D. B., and Fraaije, M. W. (2007) Discovery of a Eugenol Oxidase from *Rhodococcus* sp. strain RHA1. *The FEBS Journal* **274**, 2311–2321 10.1111/j.1742-4658.2007.05767.x PMID 17419730

8. Drijfhout, F. P., Fraaije, M. W., Jongejan, H., van Berkel, W. J. H., and Franssen, M. C. R. (1998) Enantioselective Hydroxylation of 4-Alkylphenols by Vanillyl Alcohol Oxidase. *Biotechnological Bioengineering* **59**, 171–177 10.1002/(SICI)1097-0290(19980720)59:2<171::AID-BIT5>3.0.CO;2-E

9. Fraaije, M. W., Sjollema, K. A., Veenhuis, M., and van Berkel, W. J. H. (1998) Subcellular Localization of Vanillyl Alcohol Oxidase in *Penicillium simplicissimum. FEBS Letters* **422**, 65–68 10.1016/s0014-5793(97)01605-0

10. Ewing, T. A., van Noord, A., Paul, C. E., and van Berkel, W. J. H. (2018) A Xylenol Orange-Based Screening Assay for the Substrate Specificity of Flavin-Dependent *para*-Phenol Oxidases. *Molecules (Basel, Switzerland)* **23** 10.3390/molecules23010164 PMID 29342886

11. Fraaije, M. W., van den Heuvel, R. H. H., Roelofs, J. C., and van Berkel, W. J. H. (1998) Kinetic Mechanism of Vanillyl Alcohol Oxidase with Short-Chain 4-Alkylphenols. *European Journal of Biochemistry* **253**, 712–719 10.1046/j.1432-1327.1998.2530712.x PMID 9654070

12. van den Heuvel, R. H. H., Fraaije, M. W., Laane, C., and van Berkel, W. J. H. (1998) Regio- and Stereospecific Conversion of 4-Alkylphenols by the Covalent Flavoprotein Vanillyl Alcohol Oxidase. *Journal of Bacteriology* **180**, 5646–5651 10.1128/JB.180.21.5646-5651.1998 PMID 9791114

13. van den Heuvel, R. H. H., van den Berg, W. A. M., Rovida, S., and van Berkel, W. J. H. (2004) Laboratory-Evolved Vanillyl Alcohol Oxidase Produces Natural Vanillin. *The Journal of Biological Chemistry* **279**, 33492–33500 10.1074/jbc.M312968200 PMID 15169773

14. van den Heuvel, R. H. H., Fraaije, M. W., Laane, C., and van Berkel, W. J. H. (2001) Enzymatic Synthesis of Vanillin. *Journal of Agricultural and Food Chemistry* **49**, 2954–2958 10.1021/jf010093j PMID 11409992

15. Mattevi, A., Fraaije, M. W., Coda, A., and van Berkel, W. J. H. (1997) Crystallization and Preliminary X-Ray Analysis of the Flavoenzyme Vanillyl Alcohol Oxidase from *Penicillium simplicissimum. Proteins* **27**, 601–603 10.1002/(SICI)1097-0134(199704)27:4<601::AID-PROT12>3.0.CO;2-O

16. Mattevi, A., Fraaije, M. W., Mozzarelli, A., Olivi, L., Coda, A., and van Berkel, W. J. H. (1997) Crystal Structures and Inhibitor Binding in the Octameric Flavoenzyme Vanillyl Alcohol Oxidase: The Shape of the Active-Site Cavity Controls Substrate Specificity. *Structure (London, England : 1993)* **5**, 907–920 10.1016/S0969-2126(97)00245-1 PMID 9261083

17. van den Heuvel, R. H. H., Fraaije, M. W., Ferrer, M., Mattevi, A., and van Berkel, W. J. H. (2000) Inversion of Stereospecificity of Vanillyl Alcohol Oxidase. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 9455–9460 10.1073/pnas.160175897 PMID 10920192

18. van den Heuvel, R. H. H., Laane, C., and van Berkel, W. J. H. (2001) Exploring the Biocatalytic Potential of Vanillyl Alcohol Oxidase by Site-Directed Mutagenesis. *Advanced*

Synthesis & Catalysis **343**, 515–520 10.1002/1615-4169(200108)343:6/7<515::AID-ADSC515>3.0.CO;2-U

19. Fraaije, M. W., and van Berkel, W. J. H. (1997) Catalytic Mechanism of the Oxidative De-methylation of 4-(Methoxymethyl)phenol by Vanillyl Alcohol Oxidase. Evidence for For-mation of a *p*-Quinone Methide Intermediate. *The Journal of Biological Chemistry* **272**, 18111–18116 10.1074/jbc.272.29.18111 PMID 9218444

20. Gygli, G., and van Berkel, W. J. H. (2017) Communication. Vanillyl Alcohol Oxidases Pro-duced in *Komagataella phaffii* Contain a Highly Stable Noncovalently Bound Anionic FAD Semiquinone. *Biocatalysis* **3** 10.1515/boca-2017-0002

21. Belgacem, M. N. (2011) *Monomers, Polymers and Composites from Renewable Re-sources,* Elsevier Science, Burlington

22. Ferreira, S. L., Stauder, C. M., Martin, D. K. H., and Kasson, M. T. (2021) Morphological and Phylogenetic Resolution of *Diplodia corticola* and *D. quercivora*, Emerging Canker Pathogens of Oak (*Quercus* spp.), in the United States. *Plant Disease* **105**, 1298–1307 10.1094/PDIS-05-20-0977-RE PMID 32852252

23. Martin, D. K. H., Turcotte, R. M., Miller, T. M., Munck, I. A., Aćimović, S. G., Macias, A. M., Stauder, C. M., and Kasson, M. T. (2017) First Report of *Diplodia corticola* Causing Stem Cankers and Associated Vascular Occlusion of Northern Red Oak (*Quercus rubra*) in West Virginia. *Plant Disease* **101**, 380 10.1094/PDIS-08-16-1181-PDN

24. Gygli, G., Lucas, M. F., Guallar, V., and van Berkel, W. J. H. (2017) The Ins and Outs of Vanillyl Alcohol Oxidase: Identification of Ligand Migration Paths. *PLoS Computational Bi-ology* **13**, e1005787 10.1371/journal.pcbi.1005787 PMID 28985219

25. van den Heuvel, R. H. H., Fraaije, M. W., Mattevi, A., and van Berkel, W. J. H. (2000) Asp-170 is Crucial for the Redox Properties of Vanillyl Alcohol Oxidase. *The Journal of Biolog-ical Chemistry* **275**, 14799–14808 10.1074/jbc.275.20.14799 PMID 10809721

26. Fraaije, M. W., van den Heuvel, R. H. H., van Berkel, W. J. H., and Mattevi, A. (2000) Structural Analysis of Flavinylation in Vanillyl Alcohol Oxidase. *The Journal of Biological Chemistry* **275**, 38654–38658 10.1074/jbc.M004753200 PMID 10984479

27. Nguyen, Q.-T., de Gonzalo, G., Binda, C., Rioz-Martínez, A., Mattevi, A., and Fraaije, M. W. (2016) Biocatalytic Properties and Structural Analysis of Eugenol Oxidase from *Rhodo-coccus jostii* RHA1: A Versatile Oxidative Biocatalyst. *Chembiochem: A European Journal of Chemical Biology* **17**, 1359–1366 10.1002/cbic.201600148 PMID 27123962

28. Ewing, T. A., Nguyen, Q.-T., Allan, R. C., Gygli, G., Romero, E., Binda, C., Fraaije, M. W., Mattevi, A., and van Berkel, W. J. H. (2017) Two Tyrosine Residues, Tyr-108 and Tyr-503, are Responsible for the Deprotonation of Phenolic Substrates in Vanillyl Alcohol Oxidase. *The Journal of Biological Chemistry* **292**, 14668–14679 10.1074/jbc.M117.778449 PMID 28717004

29. Kishore Babu, B., Saxena, A. K., Srivastava, A. K., and Arora, D. K. (2007) Identification and Detection of *Macrophomina phaseolina* by Using Species-Specific Oligonucleotide Pri-mers and Probe. *Mycologia* **99**, 797–803 10.3852/mycologia.99.6.797 PMID 18333503

30. Nishihara, K., Kanemori, M., Kitagawa, M., Yanagi, H., and Yura, T. (1998) Chaperone Coexpression Plasmids: Differential and Synergistic Roles of DnaK-DnaJ-GrpE and GroEL-GroES in Assisting Folding of an Allergen of Japanese cedar pollen, Cryj2, in *Esch-erichia coli*. *Applied and Environmental Microbiology* **64**, 1694–1699 10.1128/AEM.64.5.1694-1699.1998 PMID 9572938

31. Nishihara, K., Kanemori, M., Yanagi, H., and Yura, T. (2000) Overexpression of Trigger Factor Prevents Aggregation of Recombinant Proteins in *Escherichia coli*. *Applied and Environmental Microbiology* **66**, 884–889 10.1128/AEM.66.3.884-889.2000 PMID 10698746

32. Benen, J. A., Sánchez-Torres, P., Wagemaker, M. J., Fraaije, M. W., van Berkel, W. J. H., and Visser, J. (1998) Molecular Cloning, Sequencing, and Heterologous Expression of the *vaoA* Gene from *Penicillium simplicissimum* CBS 170.90 Encoding Vanillyl Alcohol Oxidase. *The Journal of Biological Chemistry* **273**, 7865–7872 10.1074/jbc.273.14.7865 PMID 9525880

33. Fraaije, M. W., van den Heuvel, R. H. H., van Berkel, W. J. H., and Mattevi, A. (1999) Covalent Flavinylation Is Essential for Efficient Redox Catalysis in Vanillyl Alcohol Oxidase. *Journal of Biological Chemistry* **274**, 35514–35520 10.1074/jbc.274.50.35514

34. Ewing, T. A., Gygli, G., and van Berkel, W. J. H. (2016) A Single Loop is Essential for the Octamerization of Vanillyl Alcohol Oxidase. *The FEBS Journal* **283**, 2546–2559 10.1111/febs.13762 PMID 27214042

35. Hill, A. V. (1910) A New Mathematical Treatment of Changes of Ionic Concentration in Muscle and Nerve Under the Action of Electric Currents, with a Theory as to Their Mode of Excitation. *The Journal of Physiology* **40**, 190–224 10.1113/jphysiol.1910.sp001366 PMID 16993004

36. LiCata, V. J., and Allewell, N. M. (1997) Is Substrate Inhibition a Consequence of Allostery in Aspartate Transcarbamylase? *Biophysical Chemistry* **64**, 225–234 10.1016/S0301-4622(96)02204-1 PMID 9127947

37. Binion, D. E. (2008) *Macrofungi Associated with Oaks of Eastern North America,* West Virginia Univ. Press, Morgantown, WVa

38. Feau, N., and Bernier, L. (2004) First Report of Shining Willow as a Host Plant for *Septoria musiva*. *Plant Disease* **88**, 770 10.1094/PDIS.2004.88.7.770B PMID 30812495

39. Tsopelas, P., Barnes, I., Soulioti, N., and Wingfield, M. J. (2013) *Dothistroma septosporum* Identified in Greece on *Pinus brutia* and *Pinus nigra* Plantations. *Plant Disease* **97**, 1247 10.1094/PDIS-11-12-1107-PDN PMID 30722443

40. Tamura, K., Stecher, G., and Kumar, S. (2021) MEGA11: Molecular Evolutionary Genetics Analysis Version 11. *Molecular Biology and Evolution* **38**, 3022–3027 10.1093/molbev/msab120 PMID 33892491

41. Saitou, N., and Nei, M. (1987) The Neighbor-Joining Method: A New Method for Reconstructing Phylogenetic Trees. *Molecular Biology and Evolution* **4**, 406–425 10.1093/oxfordjournals.molbev.a040454 PMID 3447015

42. Madeira, F., Pearce, M., Tivey, A. R. N., Basutkar, P., Lee, J., Edbali, O., Madhusoodanan, N., Kolesnikov, A., and Lopez, R. (2022) Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Research* **50**, W276-9 10.1093/nar/gkac240 PMID 35412617

43. Yang, Z. (2007) PAML 4: Phylogenetic Analysis by Maximum Likelihood. *Molecular Biology and Evolution* **24**, 1586–1591 10.1093/molbev/msm088 PMID 17483113

44. Stamatakis, A. (2014) RAxML Version 8: A Tool for Phylogenetic Analysis and Post-Analysis of Large Phylogenies. *Bioinformatics (Oxford, England)* **30**, 1312–1313 10.1093/bioinformatics/btu033 PMID 24451623

45. Katoh, K., Rozewicki, J., and Yamada, K. D. (2019) MAFFT Online Service: Multiple Sequence Alignment, Interactive Sequence Choice and Visualization. *Briefings in Bioinformatics* **20**, 1160–1166 10.1093/bib/bbx108 PMID 28968734

46. Mascotti, M. L. (2022) Resurrecting Enzymes by Ancestral Sequence Reconstruction. *Methods in Molecular Biology (Clifton, N.J.)* **2397**, 111–136 10.1007/978-1-0716-1826-4_7 PMID 34813062

47. Krieger, E., and Vriend, G. (2014) YASARA View - Molecular Graphics for All Devices - from Smartphones to Workstations. *Bioinformatics (Oxford, England)* **30**, 2981–2982 10.1093/bioinformatics/btu426 PMID 24996895

48. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., Back, T., Petersen, S., Reiman, D., Clancy, E., Zielinski, M., Steinegger, M., Pacholska, M., Berghammer, T., Bodenstein, S., Silver, D., Vinyals, O., Senior, A. W., Kavukcuoglu, K., Kohli, P., and Hassabis, D. (2021) Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **596**, 583–589 10.1038/s41586-021-03819-2 PMID 34265844

49. Mirdita, M., Schütze, K., Moriwaki, Y., Heo, L., Ovchinnikov, S., and Steinegger, M. (2022) ColabFold: Making Protein Folding Accessible to All. *Nature Methods* **19**, 679–682 10.1038/s41592-022-01488-1 PMID 35637307

50. Schrödinger, LLC and Warren DeLano (2020-05-20) *PyMol*

51. Trott, O., and Olson, A. J. (2010) AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. *Journal of Computational Chemistry* **31**, 455–461 10.1002/jcc.21334 PMID 19499576

52. Gibson, D. G., Young, L., Chuang, R.-Y., Venter, J. C., Hutchison, C. A., and Smith, H. O. (2009) Enzymatic Assembly of DNA Molecules up to Several Hundred Kilobases. *Nature Methods* **6**, 343–345 10.1038/nmeth.1318 PMID 19363495

53. Kiuchi, K., Nishikimi, M., and Yagi, K. (1982) Purification and Characterization of *L*-Gulonolactone Oxidase from Chicken Kidney Microsomes. *Biochemistry* **21**, 5076–5082 10.1021/bi00263a035 PMID 7138847

Figure legends

Figure 1. *Ps*VAO catalyzes a broad spectrum of reactions. **A:** oxidative ether cleavage **B:** alcohol oxidation **C:** oxidative deamination **D:** γ-hydroxylation **E:** α-hydroxylation **F:** dehydrogenation. Molecular oxygen is always employed as electron acceptor. Reactions E and F are performed in parallel resulting in a product mixture where the ratio of alcohol and alkene products depends on the amino acid composition in the catalytic center.

Figure 2. A: Phylogenetic tree of fungal VAOs with EUGO from *R jostii* RHA1 and 4EPO from *G. chungangensis* as outgroup. Enzymes investigated in this work are shown with names in bold: VAOs from *D. corticola* (*Dc*VAO) and from *P. simplicissimum* (*Ps*VAO) respectively. The tree is divided in five clades (I to V). **B:** Color-coded amino acids of the catalytic center from *Ps*VAO extracted from the multiple sequence alignment. Numbering corresponds to *Ps*VAO. **C:** Key amino acids of the catalytic center of *Ps*VAO are depicted as sticks. The figure is based on a crystal structure (PDB: 2VAO). The FAD cofactor is shown in orange and the bound substrate 4-ethylguaiacol in pink. **D:** Key amino acids of the catalytic center of *Dc*VAO are depicted as sticks. The figure is based on a homology model of the enzyme. The FAD cofactor is shown in orange.

Figure 3. A: Overview of the chemical structures of selected substrates. For the complete list see table S1. **B:** Substrate selectivity of *Dc*VAO wildtype and variants compared to *Ps*VAO wildtype. The reactions were performed with 50 nM of enzyme in potassium phosphate buffer at pH 7.5 and 25°C. Product formation was determined by the hydrogen peroxide production in a colorimetric assay.

Figure 4. Distances of Glu466 to indicated residues of respective substrate molecules in the homology model of *Dc*VAO after docking and subsequent energy minimization by molecular dynamics simulations. **A:** For 4-allyl-2,6-dimethoxy-phenol (**3**), the distance from the oxygen atom to the C_γ atom is half an angstrom shorter than to the C_α atom (O_{Glu}-C_α: 4.2 Å, O_{Glu}-C_γ: 3.7 Å). The O_{Glu}-C_γ distance is the closest distance measured for any of the substrates. **B:** For 4-(hydroxymethyl)-2,6-dimethoxyphenol (**6**), the distances towards the C_α atom and the benzylic oxygen is the same (O_{Glu}-O_β: 4.8 Å, O_{Glu}-C_α: 4.8 Å). **C:** For vanillyl ethyl ether (**14**), a closer distance towards the ether oxygen is observed than to the C_α atom (O_{Glu}-C_α: 4.7 Å, O_{Glu}-O_β: 4.1 Å). **D:** Distance in the *Ps*VAO crystal structure (2VAO) between the sulfur atom of Cys470 to the C_α atom of the bound substrate molecule 4-ethylguiacol (S_{Cys}-C_α: 5.5 Å). In general, the distance in *Ps*VAO is much larger and beyond hydrogen bond distance compared to *Dc*VAO.

Figure 5. A: Compressed phylogeny employed for ASR. The VAO group is shown with black branches. The Clade IV including *Ps*VAO in shown in red, while Clades I, II and III (this including *Dc*VAO) in light blue and uncharacterized Clade V is shown in orange. Bacterial EUGO & 4EPO homologs were used as outgroup (shown in grey). The resurrected ancestors are marked with larger dark colored circles and the extant sequences with smaller and lighter ones. All circles are colored according to their clades. **B:** Substrate profile of AncVAO. The percentage of converted compounds of the respective categories A to H (see below) is shown. **C:** Substrate profile overlap of Anc*Ps*VAO (red) and Anc*Dc*VAO (blue). The percentage of converted compounds of the respective categories A to H is shown. **D:** Substrate profile overlap of *Ps*VAO (red) and *Dc*VAO (blue). The percentage of converted compounds of the respective categories A to H is shown. Substrates were considered transformed when the observed activity on a substrate was higher than the 7.5% of the average activity of all converted substrates from the respective enzyme. In total, 37 substrates were tested which were grouped in eight categories by chemical characteristics. Substrate categories share a common structural motif and differ in other regarding the chain length or *o*-substitution. Categories: *p*-allyl phenol (A), *p*-benzyl alcohol (B), *p*-benzyl amine (C), *p*-benzyl ether (D), *p*-methyl phenol (E), *p*-ethyl phenol (F), *p*-alkyl phenol (G, chain length >2), bicyclic phenol (H). For further details see Table S3.

Table 1. Kinetic data from Michaelis-Menten experiments for *Dc*VAO and *Ps*VAO. The reactions were performed at 25°C in potassium phosphate buffer at pH 7.5. All ether substrates and vanillyl alcohol follow a classical Michaelis-Menten kinetic mechanism while a cooperative effect was observed for substrate eugenol. Further, *Ps*VAO displays substrate inhibition. Subsequently, the kinetic data for eugenol with *Dc*VAO, was fitted with the Hill equation(35) and for *Ps*VAO was fitted with the Licata equation.(36) For further details, see Figure S8 to S12 in the section *Michaelis-Menten kinetics for DcVAO and PsVAO* in the supplementary information.

Substrate	No	Enzyme	Equation	K_M [µM]	k_{cat} [s^{-1}]	n	K_i [mM]
Eugenol	**2**	*Dc*VAO	(2)	60 ± 3	10.3 ± 0.1	2.6 ± 0.2	-
		*Ps*VAO	(4)	67 ± 2	18.7 ± 0.5	4.8 ± 0.6	12.3 ± 4.5
Vanillyl alcohol	**5**	*Dc*VAO	(1)	297 ± 21	4.0 ± 0.1	-	-
		*Ps*VAO	(1)	144 ± 10	6.7 ± 0.1	-	-
Vanillyl methyl ether	**13**	*Dc*VAO	(1)	263 ± 12	1.9 ± 0.01	-	-
		*Ps*VAO	(1)	155 ± 5	7.9 ± 0.1	-	-
Vanillyl ethyl ether	**14**	*Dc*VAO	(1)	198 ± 19	4.7 ±0.2	-	-
		*Ps*VAO	(1)	86 ± 3	2.6 ± 0.02	-	-
Vanillyl butyl ether	**15**	*Dc*VAO	(1)	258 ± 27	2.1 ± 0.1	-	-
		*Ps*VAO	-	n.d.	n.d.	n.d.	n.d.

A Glu₄₆₆   **B** Glu₄₆₆   **C** Glu₄₆₆   **D** Cys₄₇₀
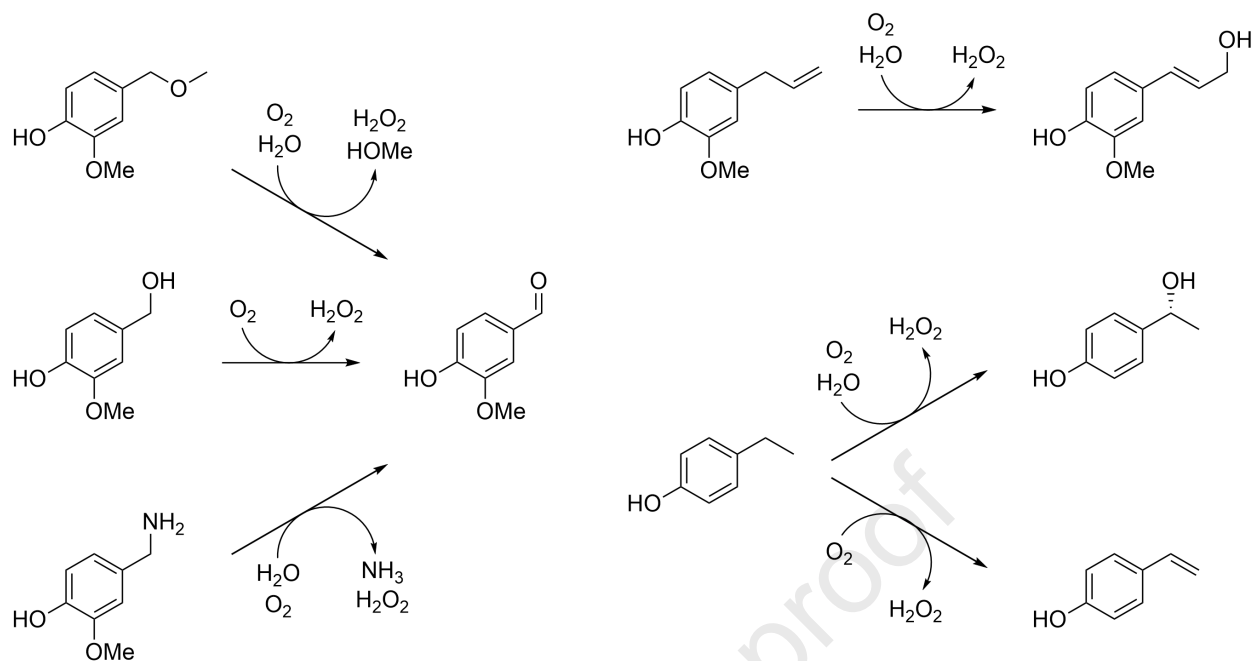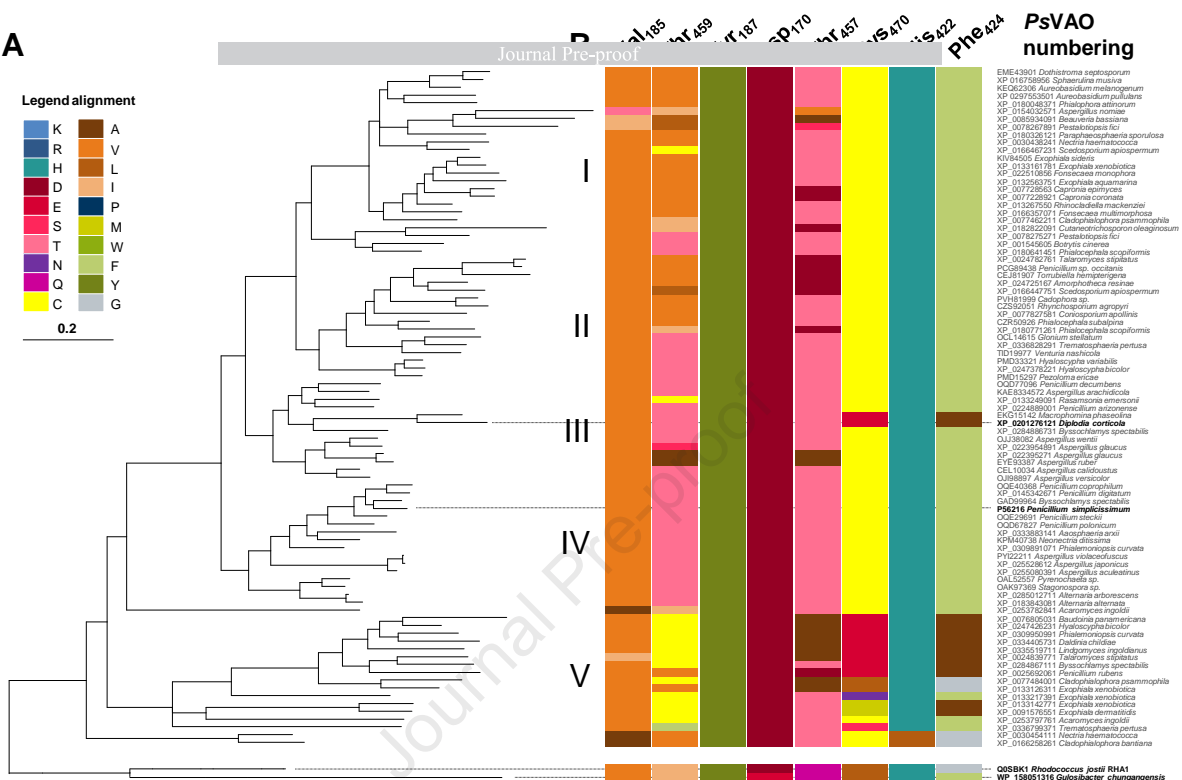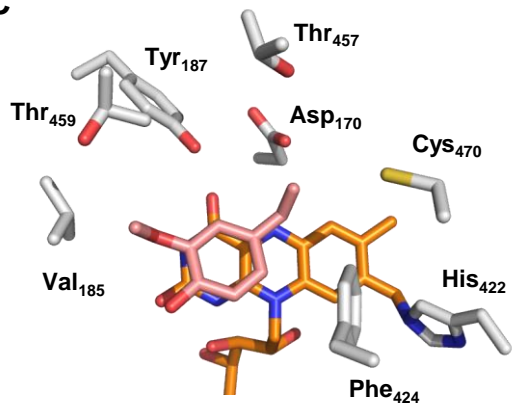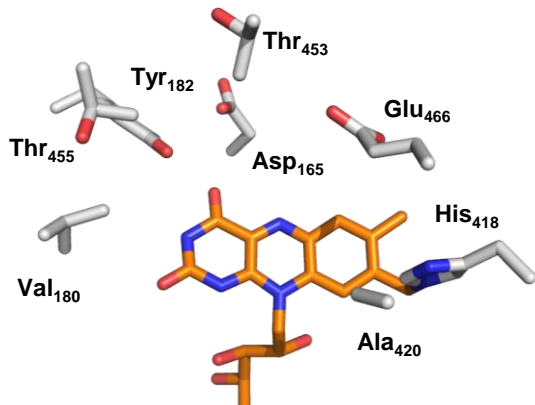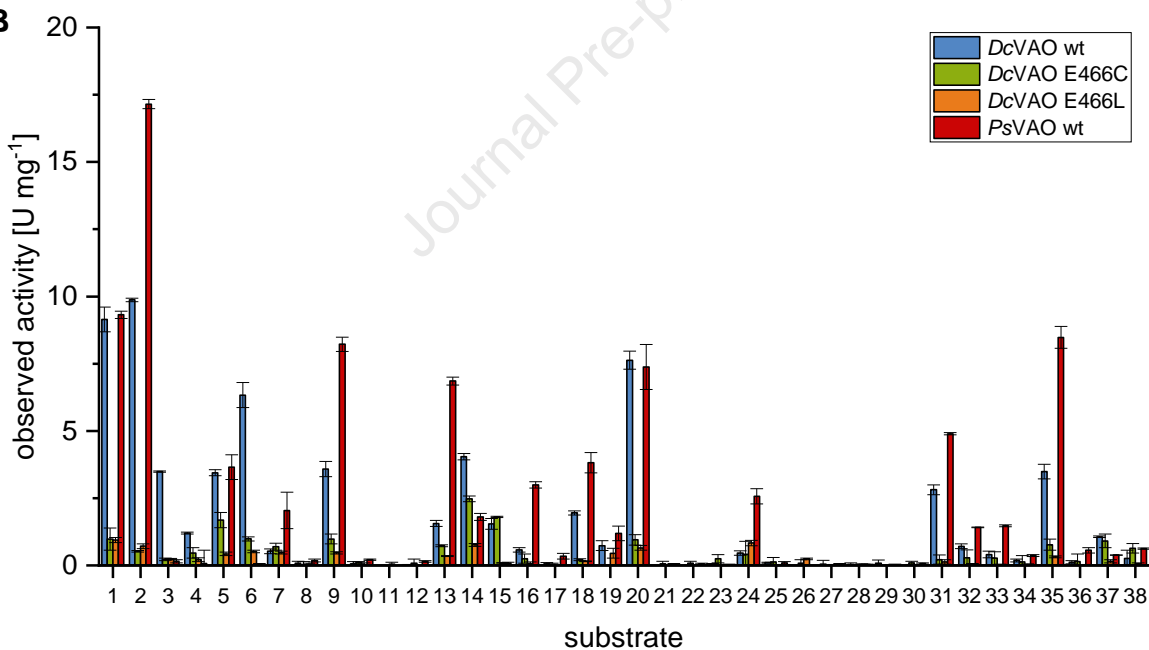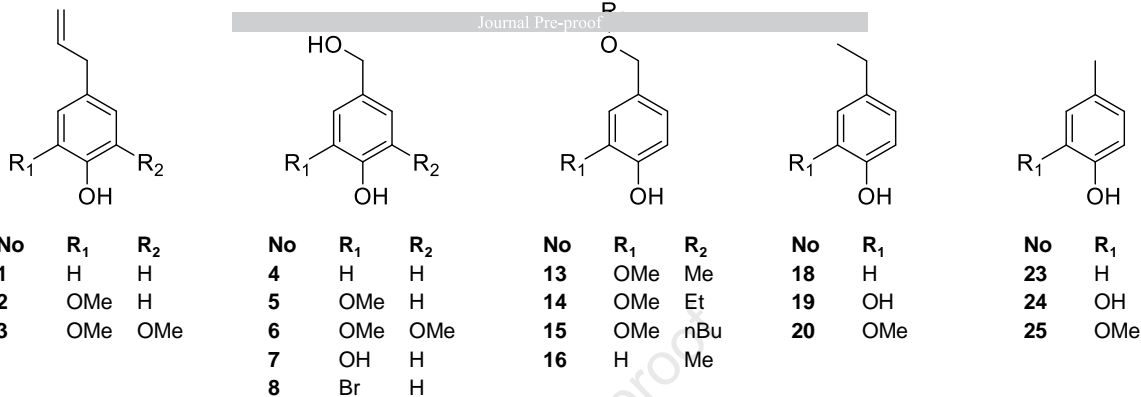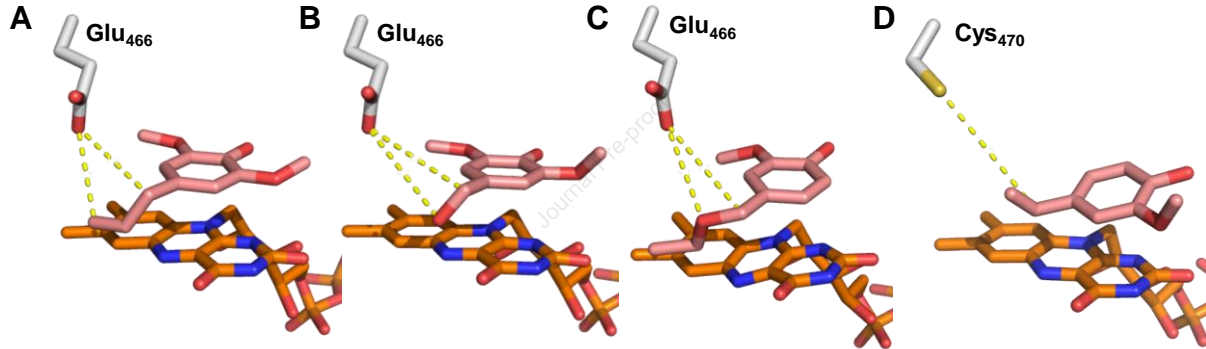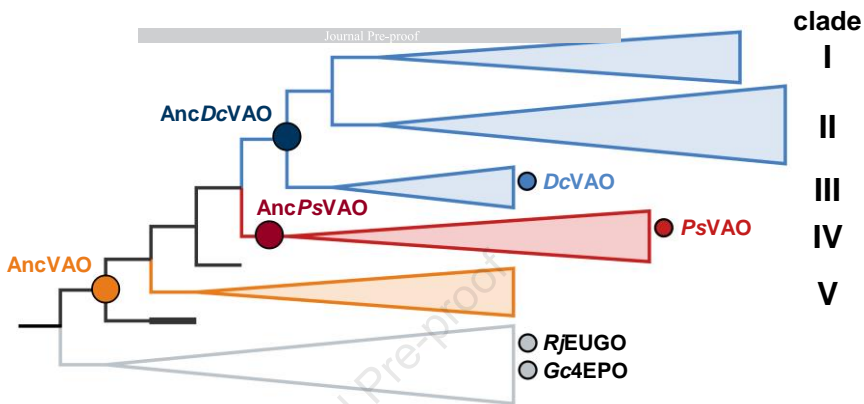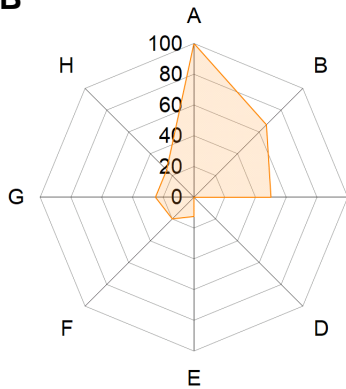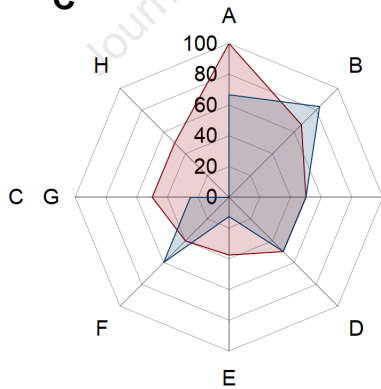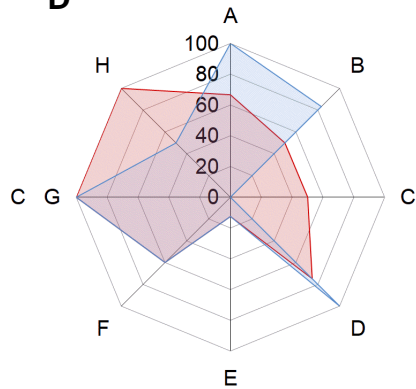
Daniel Eggerichs: Conceptualization, Methodology, Data Curation, Writing - Original Draft, Visualization, Supervision

Nils Weindorf: Methodology, Investigation, Validation, Writing - Review & Editing

Maria Laura Mascotti: Investigation, Data Curation, Writing - Review & Editing, Visualization

Natalie Welzel: Investigation, Validation

Marco W. Fraaije: Conceptualization, Writing - Review & Editing

Dirk Tischler: Conceptualization, Writing - Review & Editing, Funding acquisition, Supervision