Fungal Genetics and Biology 116 (2018) 24–32

Contents lists available at ScienceDirect

Fungal Genetics and Biology

journal homepage: www.elsevier.com/locate/yfgbi

Regular Articles

# On the origin of vanillyl alcohol oxidases

Gudrun Gygli[a], Ronald P. de Vries[b], Willem J.H. van Berkel[a,*]

[a] Laboratory of Biochemistry, Wageningen University & Research, Stippeneng 4, 6708 WE Wageningen, The Netherlands
[b] Fungal Physiology, Westerdijk Fungal Biodiversity Institute & Fungal Molecular Physiology, Utrecht University, Uppsalalaan 8, 3584 CT Utrecht, The Netherlands

ARTICLE INFO

*Keywords:*
Flavoprotein
Fungus
Phylogeny
Sequence motif
Dehydrogenase

ABSTRACT

Vanillyl alcohol oxidase (VAO) is a fungal flavoenzyme that converts a wide range of para-substituted phenols. The products of these conversions, e.g. vanillin, coniferyl alcohol and chiral aryl alcohols, are of interest for several industries. VAO is the only known fungal member of the 4-phenol oxidising (4PO) subgroup of the VAO/PCMH flavoprotein family. While the enzyme has been biochemically characterised in great detail, little is known about its physiological role and distribution in fungi.

We have identified and analysed novel, fungal candidate VAOs and found them to be mostly present in Pezizomycotina and Agaricomycotina. The VAOs group into three clades, of which two clades do not have any characterised member. Interestingly, bacterial relatives of VAO do not form a single outgroup, but rather split up into two separate clades.

We have analysed the distribution of candidate VAOs in fungi, as well as their genomic environment. VAOs are present in low frequency in species of varying degrees of relatedness and in regions of low synteny. These findings suggest that fungal VAOs may have originated from bacterial ancestors, obtained by fungi through horizontal gene transfer.

Because the overall conservation of fungal VAOs varies between 60 and 30% sequence identity, we argue for a more reliable functional prediction using critical amino acid residues. We have defined a sequence motif P-x-x-x-x-S-x-G-[RK]-N-x-G-Y-G-[GS] that specifically recognizes 4PO enzymes of the VAO/PCMH family, as well as additional motifs that can help to further narrow down putative functions. We also provide an overview of fingerprint residues that are specific to VAOs.

## 1. Introduction

Vanillyl alcohol oxidase (VAO, EC 1.1.3.38) is a covalent flavoenzyme first isolated from the ascomycetous fungus *Penicillium simplicissimum* (de Jong et al., 1992) . VAO is active with a wide range of para-substituted phenols (Fraaije et al., 1995; van den Heuvel et al., 1998). Several VAO reactions produce high-quality aromatic compounds, e.g. vanillin and coniferyl alcohol (see Fig. 1A for an overview). These molecules are of interest for the food, flavour and fragrance industry, exemplified by several patents, e.g. from Mane (Lambert et al., 2007), Rhodia (Gayet et al., 2014) and Unilever (van Berkel et al., 1993).

VAO can also produce chiral aryl alcohols. For instance, the VAO-mediated conversion of 4-ethylphenol results in the formation of (*R*)-1-(4-hydroxyphenyl)ethanol with an enantiomeric excess of 94% (Drijfhout et al., 1998). Interestingly, an engineered variant of VAO was shown to be capable of producing the (*S*)-isomer of 1-(4-hydroxyphenyl)ethanol with an enantiomeric excess of 80% (van den Heuvel et al., 2000a). It has been proposed that the methyl ether 4-(methoxymethyl)phenol is the physiological substrate of VAO, as it is the only known substrate that induces expression of the *vao* gene in *P. simplicissimum* (de Jong et al., 1992). However, little is known about the origin of 4-(methoxymethyl)phenol and the physiological role of VAO. Subcellular localisation studies showed that the *P. simplicissimum* enzyme (*Ps*VAO), together with a co-inducible catalase-peroxidase, is distributed throughout the cytosol and peroxisomes (Fraaije et al., 1998a).

*Ps*VAO is the prototype of a large flavoprotein family, the VAO/PCMH family, together with *p*-cresol methylhydroxylase (*Pp*PCMH, from *Pseudomonas putida* NCIMB 9866, recommended name: 4-methylphenol dehydrogenase (hydroxylating), EC 1.17.99.1). The FAD-binding domains of all enzymes within this family share a common fold (Fraaije et al., 1998b). Within the VAO/PCMH family, VAO and PCMH belong to the 4-phenol oxidising (4PO) subgroup (Ewing et al., 2017a). All 4PO enzymes contain a Tyr-Tyr-Arg triad, which is crucially involved in substrate binding, and therefore is the cause of the selectivity

* Corresponding author.
*E-mail address:* willem.vanberkel@wur.nl (W.J.H. van Berkel).

https://doi.org/10.1016/j.fgb.2018.04.003
Received 18 January 2018; Received in revised form 24 March 2018; Accepted 3 April 2018
Available online 05 April 2018
1087-1845/ © 2018 The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/BY-NC-ND/4.0/).

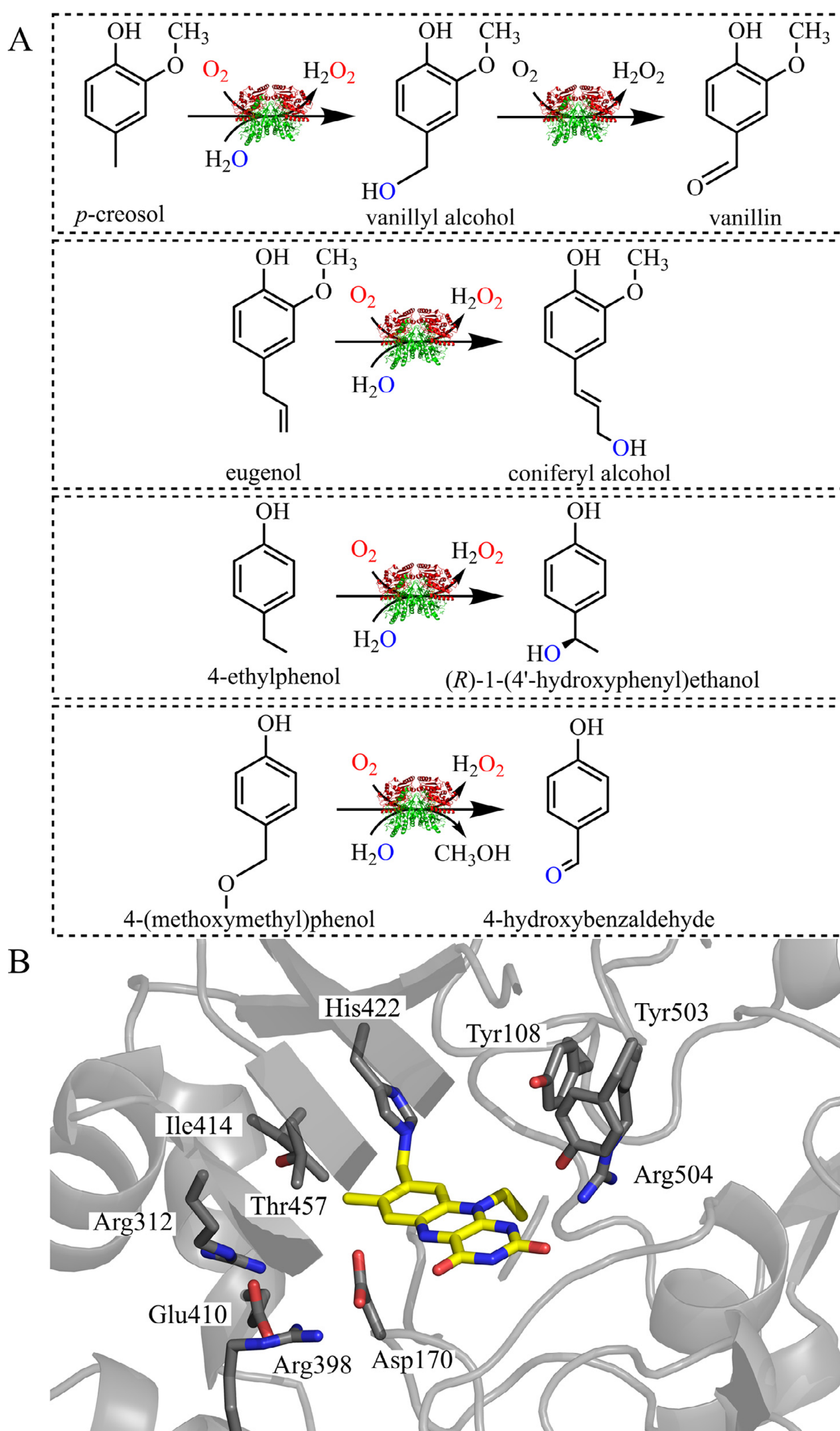


**Fig. 1.** Different reactions catalysed by *Ps*VAO (A) and active site of *Ps*VAO (B). A: *p*-Creosol (4-methylphenol) is converted via vanillyl alcohol (4-hydroxy-3-methoxybenzyl alcohol) to vanillin (4-hydroxy-3-methoxybenzaldehyde). Coniferyl alcohol (4-hydroxy-3-methoxycinnamyl alcohol) is produced from eugenol (2-methoxy-4-allyl-phenol). 4-Ethylphenol is converted to (*R*)-1-(4-hydroxyphenyl)ethanol. The proposed physiological substrate 4-(methoxymethyl)phenol is oxidatively demethylated to 4-hydroxybenzaldehyde. B: Active site of *Ps*VAO (PDBID 1VAO). The FAD cofactor is coloured in yellow and critical amino acid residues are shown in dark grey. For clarity, only the isoalloxazine moiety of the FAD cofactor (which is covalently bound to His422) is shown. See Table 1 for more information about these critical amino acid residues.

of these enzymes for para-substituted phenols (Ewing et al., 2017b). VAO is the only known fungal member of this subgroup, whereas all other known members are bacterial enzymes (Brandt et al., 2001; Cunane et al., 2000; Ewing et al., 2017a; Jin et al., 2007; Leferink et al., 2008; Priefert et al., 1999; Reeve et al., 1989).

Among the 4PO enzymes, *Ps*VAO is characterised in greatest biochemical detail (Gygli et al., 2017). It has been established that the covalent binding of the FAD cofactor of *Ps*VAO is crucial for the redox properties of this enzyme, as it significantly increases its redox potential (Fraaije et al., 1999). This increase in redox potential speeds up the rate of reduction of the FAD by the substrate, and thus increases the overall reaction rate (Fraaije et al., 1999). The same effect has been observed in *Pp*PCMH (Efimov et al., 2004, 2001).

*Ps*VAO has most properties in common with eugenol oxidase from

**Table 1**
Function of critical amino acid residues in *Ps*VAO.

| Residue(s) | Function |
|---|---|
| Tyr108, Tyr503, Arg504 | Substrate binding, substrate deprotonation (Ewing et al., 2017b; Mattevi et al., 1997) |
| His61 | Involved in autocatalytic formation of covalent bond of FAD to His422 (Fraaije et al., 2000) |
| His422 | Covalent bond to FAD (Mattevi et al., 1997) |
| Ile414 | No function known in *Ps*VAO but the corresponding Tyr residue in *Pp*PCMH is covalently binding FAD (Cunane et al., 2000) |
| Asp170 | Involved in autocatalytic formation of covalent bond of FAD to His422; catalytic base; stereoselectivity (Drijfhout et al., 1998; Mattevi et al., 1997; van den Heuvel et al., 2000a, 2000b) |
| Thr457 | Stereoselectivity (Drijfhout et al., 1998; van den Heuvel et al., 2000a) |
| Loop 219–240 (octamerisation loop published: 220–235) | Essential for octamerisation (Ewing et al., 2016) |
| Tyr51, Tyr 408 | Tyrosines suggested to be involved in dioxygen migration to the *re* side of the active site (Gygli et al., 2017) |
| Asp192, Met195, Glu464, His466, (Tyr503) | Proposed to be involved in access of phenolic ligand to *si* side of active site (Gygli et al., 2017) |

the actinobacterium *Rhodoccocus jostii* RHA1 (*Rj*EUGO) (Jin et al., 2007; Nguyen et al., 2016). Both enzymes contain an 8α-($N^3$-histidyl)-FAD as prosthetic group, use oxygen as electron acceptor, and share a considerable overlap in substrate specificity. A main difference, however, is that *Ps*VAO is a homo-octamer (de Jong et al., 1992), while *Rj*EUGO is a homodimer (Jin et al., 2007). A recent protein engineering study showed that an extra surface loop in *Ps*VAO determines this difference in oligomerisation behavior (Ewing et al., 2016). Another remarkable difference between *Ps*VAO and *Rj*EUGO concerns the reactivity with alkylphenols. *Ps*VAO converts a wide range of alkylphenols to the corresponding alcohols or alkenes (van den Heuvel et al., 1998), whereas *Rj*EUGO shows almost no activity with these compounds (Jin et al., 2007).

Because of the conservation of the FAD-binding domain, all 4PO enzymes share at least 25% sequence identity. To reliably identify novel VAOs, their fungal origin can be used, as well as conservation of known critical amino acid residues (see Table 1 and Fig. 1B). Currently, no clear sequence motifs specific to VAO are defined (Dym and Eisenberg, 2001; Garma et al., 2016). Little is known about the occurrence of VAO in fungi, besides that it is produced in *P. simplicissimum*, *Byssochlamys fulva* and *Fusarium monoliforme* (de Jong et al., 1992; Furukawa et al., 1999; van Rooyen, 2012). In *B. fulva*, expression of *vao* is induced by vanillyl alcohol (Furukawa et al., 1999). In *F. monoliforme*, *vao* expression was observed with veratryl alcohol and anisyl alcohol, and to a lesser extent with vanillyl alcohol (van Rooyen, 2012). In *P. simplicissimum*, vanillyl alcohol and eugenol are not inducing expression of *vao*, but 4-(methoxymethyl)phenol, anisyl alcohol and veratryl alcohol are (Fraaije et al., 1997).

Here, we aim to shed light on the occurrence of VAO in fungi. VAO has recently been added to the CAZy (Carbohydrate-Active Enzymes) database, as a separate family of auxiliary activities (AA4) involved in lignin breakdown (Levasseur et al., 2013). We have used this criterion to mine fungal genomes in the Mycocosm database for putative VAOs.

## 2. Results

We have identified candidate VAOs in the Mycocosm database and performed a phylogenetic analysis of them. We have used the bacterial 4POs as outgroup for the analysis of these fungal sequences. We mapped the distribution of VAO homologs in different fungi. We analysed the phylogenetic tree obtained, as well as conservation of critical amino acid residues and cellular localisation of putative VAOs. We analysed the genomic environments of these putative VAOs. We also propose motifs for the future identification of fungal VAOs, as well as bacterial 4POs.

### 2.1. Fungal distribution of VAO homologs

Different classes of fungi possess sequences homologous to the sequence of *Ps*VAO, which are annotated as AA4 in the Mycocosm database. In Basidiomycota only Agaricomycotina were found to contain putative VAOs (see Fig. 2). In Ascomycota, only Pezizomycotina contain putative VAOs: dothideomycetes (29), leotiomycetes (6), sordariomycetes (19), eurotiomycetes (25) and orbiliomycetes (3), see also Fig. 2. Gene duplications appear to have occurred in almost all fungal classes containing VAO homologs. As can be seen in Fig. 2, in the Agaricomycotina only three organisms have multiple copies of *vao* genes, while in Pezizomycotina, 19 organisms have multiple copies of *vao* genes. Most striking is *Aspergillus sydowii*, which contains seven putative VAOs.

The fungi found to carry a *vao* gene are living in very diverse environments, e.g. *Jaapia argillacea* in the northern hemisphere and *Aspergillus glaucus* in the artic regions, while *Eurotium rubrum* is the most common fungal species isolated from the Dead Sea. Other fungi e.g *Gloeophyllum trabeum*, *Dacryopinax primogenitus* and *Glonium stellatum* are wood decayers. This observation does not enable us to draw a conclusion on the native functions of VAOs or fungal 4POs, other than that they are likely involved in a metabolic pathway used by all these fungi.

The prevalence of VAO sequences in fungal genomes does not follow fungal taxonomy. While several fungal clades do not appear to contain any species with a VAO, there are no clades in which all species contain a VAO. The presence of a VAO in a species also does not appear to be related to life style or biotope as they can be found in some saprobic, plant pathogenic and animal/human pathogenic fungi, but are also frequently absent in other fungi with these life styles. The same applies to the number of VAO copies, which differs for species with a VAO between one and seven and again cannot be linked to biotope or life style.

### 2.2. Cellular localisation

All identified fungal proteins appear to be intracellularly located, since no *N*-terminal signal peptides have been found in any sequence using the SignalP 4.1 Server. A proposed *C*-terminal peroxisomal targeting signal (WKL-COOH instead of the SKL-COOH peroxisomal targeting signal-1 (Fraaije et al., 1998a)) was only present in some of the sequences in the VAO clade, or further changed to YKL-COOH in eurotiomycetes and absent in all other sequences. In the same study, it was found that VAO localises in the cytoplasm as well as peroxisomes in *P. simplicissimum* (Fraaije et al., 1998a). The predictions of the DeepLoc server (Almagro Armenteros et al., 2017) reflect the experimental data, especially the bimodal distribution observed experimentally for VAO: cytoplasmic or peroxisomal localisation is often predicted to be equally likely. With only two exceptions (WP_056693180.1 and Calvi1_515549 (*Calocera viscosa*)), all sequences are predicted to be peroxisomal or cytoplasmic (see supplementary data 3).

In *P. simplicissimum*, VAO is expressed together with a catalase-peroxidase (Fraaije et al., 1998a). At the time of that study, no sequence for this catalase-peroxidase was available, but in the meantime several

 

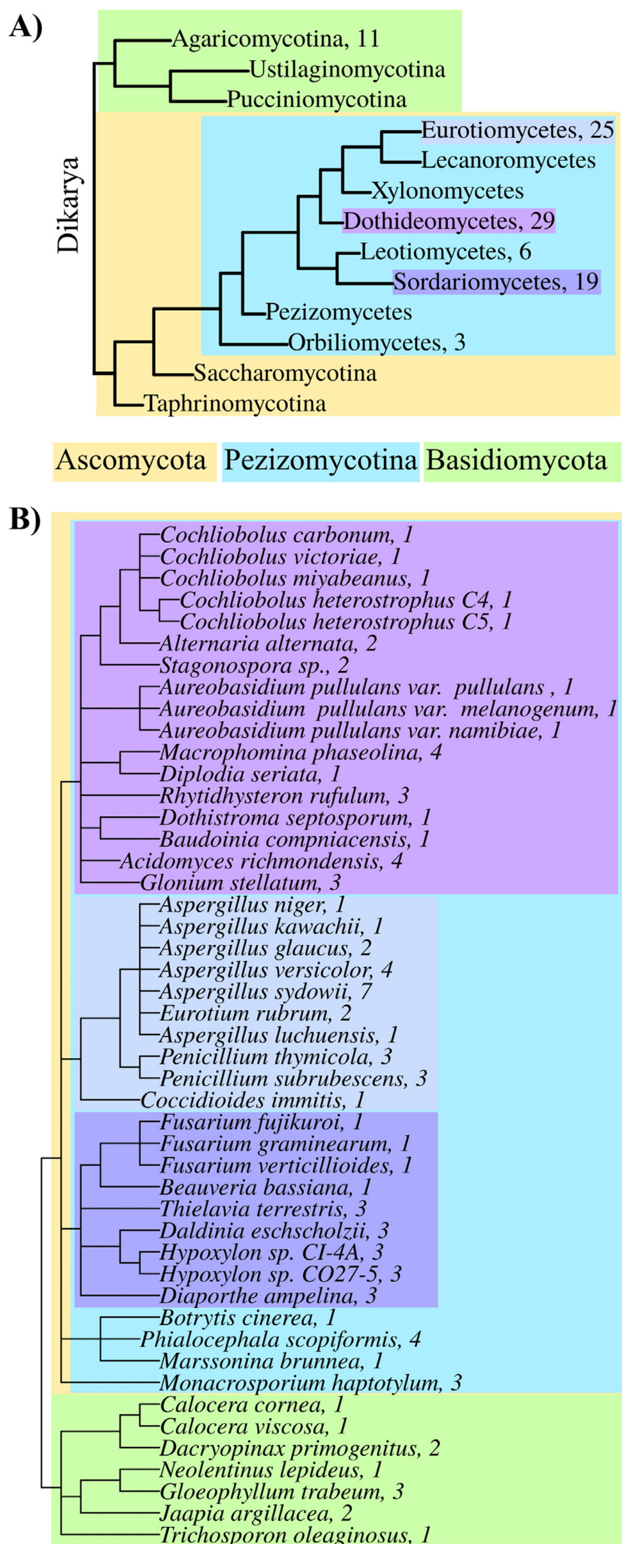


**Fig. 2.** Species-tree showing the occurrence of putative VAO sequences in different fungi. The branch lengths are indicative of evolutionary distances. (A) Overview of all fungal classes, with the number behind the class indicating how many putative VAOs were found in that class. No number means there is no putative VAO present. (B) Details on the fungal classes that were found to contain putative VAOs. The number behind the organism indicates how many putative VAOs were found in that organism. Note that especially the branch lengths for closely related species are not accurate.

sequences and crystal structures from related organisms have been solved e.g. (Carpena et al., 2004; Scherer et al., 2002; Zhao et al., 2013, 2006). Using sequence Q96VT4 from *Aspergillus nidulans*, we searched the genomes of fungi in the red clade (containing *Ps*VAO and sequences most similar to it) for homologs of this catalase-peroxidase. In total, 87 putative catalase-peroxidases were found, including isoforms in several fungi. Ninety percent of these sequences were predicted by the DeepLoc webserver to be localised in the cytoplasm or peroxisome (see supplementary data 4).

### 2.3. Genomic environment of putative VAOs

We analysed the genomic environment of sequences most similar to *Ps*VAO (in the red clade, marked in red in Fig. 3). There is some synteny between closely related species, but a notable lack thereof in a broader view across the fungal kingdom. The fact that VAO is not located in a region of high synteny could indicate that horizontal gene transfer played a role in the origin of fungal VAOs.

### 2.4. Three different subgroups/clades of fungal VAOs

Here we will describe the phylogenetic tree of the *Ps*VAO homologues in more detail (Fig. 3). This tree was created using the bacterial 4POs (*Rj*EUGO and *Pp*PCMH) and the novel fungal VAOs. All sequences were between 502 and 632 amino acids long. Please note that *Rj*EUGO is more closely related to *Ps*VAO than the other bacterial 4POs, forming its own clade (marked in green in Fig. 3). ART40635, WP_068805427, WP_061561479, AAM21269 and WP_056693180 are bacterial sequences that were added to expand the outgroup. Interestingly, these bacterial sequences do not form one outgroup in the tree but two distinct clades. They separate the fungal VAOs into one clade more closely related to PCMH and one clade more closely related to EUGO.

Three main clades can be identified, with 4POs containing a cytochrome c subunit sequence rooting the entire tree (marked in grey in Fig. 3). We describe the phylogenetic tree in counterclockwise order from the grey outgroup on. Please note that all residues are numbered according to the *Ps*VAO numbering.

The blue clade (see Fig. 3 and supplementary Fig. A1) contains sequences, which only share four critical residues with *Ps*VAO, making these four residues (Tyr108, His422, Tyr503 and Arg504) the only ones that are conserved in all fungal VAOs. The only exception is the sequence Diaam1_2(7551, from *Diaporthe ampelina* in the red clade (*vide infra*)). The sequences in the blue clade are more similar to *Pp*PCMH and markedly differ from *Ps*VAO in at least three aspects: (i) they do not contain the octamerisation loop (Ewing et al., 2016), nor do they contain Tyr51, a proposed gatekeeper for oxygen access to the active site (Gygli et al., 2017); (ii) they do not share the three residues Arg312, Arg398, Glu410 which are conserved in most of the other homologs; (iii) some of these sequences contain a short loop between positions 321 and 322, which is absent in all other sequences studied here.

The green clade, containing sequences similar to *Rj*EUGO, separates the sequences in the blue clade from the sequences in the red clade.

The clade containing *Ps*VAO is marked in red in Fig. 3 (VAO clade). This clade contains sequences that are most similar to *Ps*VAO (see supplementary Fig. A1). All sequences in this clade contain the octamerisation loop, or at least a 2–4 residues shorter variant thereof. Members of the VAO clade all contain Tyr51, Arg312, Arg398 and Glu410 (except the sequence from *Beauveria bassania* (Beaba1), which contains an aspartate at position 312).

The yellow clade (see Fig. 3 and supplementary Fig. A1) contains sequences that share eight critical amino acid residues with *Ps*VAO. This clade can be split into two sub-clades, one without an octamerisation loop (in lighter yellow in Fig. 3) and one with a very short octamerisation loop (in darker yellow, in Fig. 3, formed of 5 instead of 16 residues). It is doubtful whether this loop has the same function as the octamerisation loop in *Ps*VAO (Ewing et al., 2016). Sequences in the

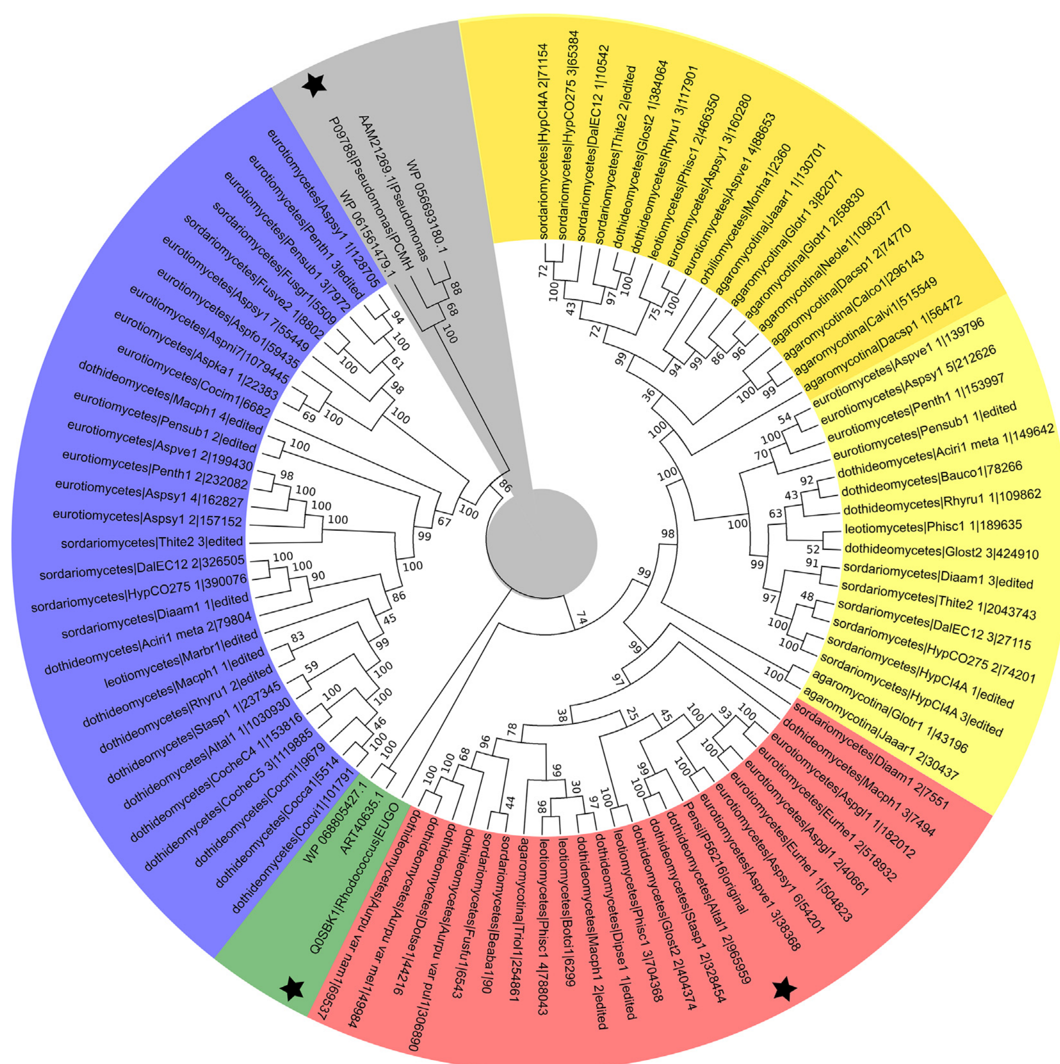


**Fig. 3.** Analysis of the molecular phylogeny of VAO by Maximum Likelihood method. The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model (Jones et al., 1992). The tree with the highest log likelihood (−36578.6019) is shown. Initial tree(s) for the heuristic search were obtained automatically as follows. When the number of common sites was < 100 or less than one fourth of the total number of sites, the maximum parsimony method was used; otherwise BIONJ method with MCL distance matrix was used. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 97 amino acid sequences. All positions containing gaps and missing data were eliminated. There were a total of 429 positions in the final dataset. Evolutionary analyses were conducted in MEGA5 (Tamura et al., 2011). Minor manual modifications were made to the obtained tree.

yellow clade contain His61, Met61 or Phe61. Two sequences contain a cysteine at position 414, Bauco1 (78266, from *Baudoinia compniacensis*) and Aciri1_meta (49642, from *Acidomyces richmondensis*). Otherwise, the conservation of critical amino acid residues in this clade is like in the red clade.

In conclusion, the blue clade contains sequences most similar to bacterial 4POs, while the red and yellow clade contain sequences most similar to *Ps*VAO. *Ps*VAO is found in the red clade, where all sequences are very similar to it. The yellow clade contains sequences without an octamerisation loop (subclade coloured in lighter yellow) and sequences with a very short "octamerisation loop" (subclade coloured in darker yellow). Enzymes whose functions have been biochemically studied are marked with a star.

### 2.5. Conservation of critical amino acid residues

All fungal sequences identified contain the Tyr108-Tyr503-Arg504 triad and His422 with the exception of Diaam1_2 (7551, from *Diaporthe ampelina*), which has phenylalanine residues at positions 503 and 422. His61, which has been found to be involved in the autocatalytic formation of the covalent His422-flavin bond (Fraaije et al., 2000) is not conserved in the sequences, but often replaced by a tyrosine, methionine or phenylalanine. No sequence has a tyrosine residue at position 414 (which covalently binds the FAD cofactor in *Pp*PCMH), but as already indicated above, two sequences of the yellow clade carry a cysteine at that position.

The catalytic base, Asp170 is present in all sequences in the red and yellow clade, with the exception of Diaam1_2 (7551, from *Diaporthe ampelina*), which carries a glutamine at that position. Sequences in the blue clade carry mainly serine or alanine at position 170, and might use an aspartate at position 457 as catalytic base instead. This would resemble the situation in *Pp*PCMH, where an alanine (Ala154) and a glutamate (Glu427) occupy these positions (Cunane et al. 2000). Thr457 is not strictly conserved in either of the three clades, but the majority of sequences carry Thr457 or Asp457. These two facts combined are a good predictor for the enzyme enantioselectivity (van den Heuvel et al., 2000a). Especially in the blue clade, inverted enantioselectivity can be expected, likely producing (S)-1-(4-hydroxyphenyl) ethanol from 4-ethylphenol (van den Heuvel et al., 2000a).

The loop between residues 220 and 237 in *Ps*VAO is present in all

sequences in the VAO clade, and absent in all other sequences. Interestingly a much shorter loop (five residues) is present in some sequences of the yellow clade. This indicates that not all of the sequences identified here will produce octameric VAOs. Considerable variation in the loop itself occurs, but prolines appear to be more conserved than other residues in the loop.

Sequences in the blue clade are more similar to the bacterial 4POs in the sense that they do not share Asp170 with *Ps*VAO, but contain a serine or alanine instead. However, they all share His422 with *Ps*VAO. Like the bacterial sequences, sequences in the blue clade lack residue Tyr51. See also supplementary Fig. A2 for sequence logos of the three fungal clades (yellow, red and blue). Tyr51 was suggested to be involved in dioxygen migration to the active site (together with Tyr408, which is not conserved). These two tyrosines are located in a loop in *Ps*VAO, where the cytochrome c subunit binds to the FAD-subunit in *Pp*PCMH (Gygli et al., 2017).

### 2.6. Motifs for VAOs

*Ps*VAO and the other members of the 4PO subgroup, as well as members of the VAO/PCMH family all share the same FAD-binding fold, for which a motif has been proposed (Dym and Eisenberg, 2001). Their specific mode of FAD binding however differs: they bind their FAD cofactor non-covalently, mono-covalently, or bi-covalently (Ewing et al., 2017a). Members that bind their FAD cofactor via the C8α position of the isoalloxazine ring do not always use the same residue or even the same domain (Heuts et al., 2009). This fact can be used to distinguish between the different members at the sequence level (Garma et al., 2016). Expanding on this idea of function-related patterns or motifs, it should also be possible to distinguish between 4PO enzymes with different activities based on patterns or motifs of active site residues.

To define a motif specific to VAOs, we expanded the sequence motifs proposed by Dym and Eisenberg (Dym and Eisenberg, 2001) with the present knowledge of critical amino acid residues (see Fig. 4). The well-known dinucleotide-binding motif (Wierenga et al., 1983) is absent in VAO, as these enzymes do not contain a Rossmann fold, but have their own, distinct fold, which is shared by all members of this family. We defined motifs based on Fig. 4, as summarised in Table 2. Some of these motifs are specific to VAO, and others to the known members of the 4PO subgroup.

## 3. Discussion

Fungi and bacteria separated about 4090 million years ago (MYA). This makes it remarkable that *Rj*EUGO still clusters more closely with VAO than with the bacterial 4POs. To better understand this, we looked at *Rj*EUGO more closely. This enzyme originates from *Rhodoccocus jostii* RHA1 (Jin et al., 2007; Nguyen et al., 2016), which is an actinobacterium. The other bacterial members of the 4PO subgroup all come from *Pseudomonas* strains, which makes them proteobacterial. Actinobacteria and Proteobacteria separated 3169 MYA. It is interesting to note that these speciation events all took place before the great oxidation event, so before atmospheric oxygen levels increased permanently (2400 to 2100 MYA, (Lyons et al., 2014)). The absence of atmospheric oxygen at the time of this separation makes it likely that the common ancestor of all 4POs was not a true oxidase.

Candidate fungal VAOs are present in a variety of fungi, but predominantly in the class of Pezizomycotina and Agaricomycotina. None of the earlier lineages appear to contain any copies of this gene, suggesting that it originated after the dikarya evolved as a separate group from those earlier lineages. However, in those groups of fungi that contain species with a VAO (e.g. Sordariomycetes, Dothidiomycetes, Eurotiomycetes, Leotiomycetes, Agaricomycotina), the presence of VAO genes is restricted to a small number of species, often not closely related. The broad diversity of species that contains VAO genes could indicate that VAO is an old enzyme that was frequently lost in fungal species, likely due to it no longer being essential. However, the phylogenetic analysis of fungal VAOs also points to an alternative explanation. As the VAO tree only partially follows the taxonomic relationships between the species, this suggests the possibility that several of these species obtained the VAO gene through horizontal transfer, possibly from bacteria. This would explain the low frequency of presence of VAO in fungal genomes, as well as the large phylogenetic diversity of the fungal VAO candidates.

Interestingly, among the species that contain candidate VAO genes, the number varies from one to seven copies, suggesting that those species with multiple copies have really capitalized on this enzymatic function and therefore likely require it for life in their biotope. An interesting example of this is *Aspergillus sydowii* that has the highest number of VAO genes detected (seven) and has both a terrestrial and a marine life style, in which it has been implicated as a major causal agent of coral bleach (Ein-Gil et al., 2009).

Another factor speaking in favour of horizontal transfer as the origin of fungal VAOs is the non-conserved genomic region in which the genes have been detected. This would be consistent with introduction by horizontal transfer, while an ancient fungal origin would more likely result in some conservation/synteny of the neighbouring genes of the VAO genes, which we could not detect.

Based on the conservation of critical amino acid residues, we are confident that the sequences we have identified will produce enzymes with VAO-like activity, with the exception of the putative VAO from *Diaporthe ampelina*. This sequence might be a pseudogene. Sequences in the blue clade are likely to have modified activities, especially with changed enantioselectivity. We argue that it is more reliable to use critical amino acid residues than overall sequence identity to predict the enzymatic function of a sequence. Sequence identities between *Ps*VAO and the novel VAOs in the blue clade are only between 30 and 40%, but most critical amino acid residues are strictly conserved. It is therefore crucial to use these critical amino acid residues when annotating novel VAO sequences. We have also included the critical amino acid residues into the motifs we defined.

The functional role of the identified motifs is likely structural: The residues in the 4PO-motif-1 form a loop, which is in close contact with the FAD cofactor. It is important to note that one of the tyrosines crucially involved in substrate deprotonation (Tyr108) is located in this loop. This motif is therefore likely linked to proper orientation and stabilisation of the position of Tyr108. The 4PO-motif-5 also forms a loop containing the other two residues involved in substrate deprotonation, Tyr503 and Arg504. Proper orientation of these residues is also the likely reason for the conservation of this motif.

Searching the Swissprot database with the defined motifs shows that, as could be expected, the shorter motifs are not specific enough to identify 4POs from the database. However, 4PO-motif-1 is highly selective and only yields sequences of *Ps*VAO and *Pp*PCMH. The sequence of *Rj*EUGO is not a reviewed sequence entry in the Swissprot database and therefore not a hit. We then expanded our search to GenBank, UniProt, RefSeq and PDBSTR. Only the 4PO-motif-1 was sufficiently selective, giving 848 hits. These sequences are (automatically) annotated as VAOs, PCMHs, glycolate oxidases, alcohol oxidases, aryl alcohol oxidases, hypothetical proteins, FAD-binding oxidoreductases, FAD binding domain-containing proteins and uncharacterized proteins. It is highly likely that the annotations of these sequences are erroneous and that they are actually all 4POs or even VAOs. For example, one sequence annotated only as FAD-binding oxidoreductase is actually *Rj*EUGO (WP_011595933.1). *Ps*VAO is also annotated as an aryl-alcohol oxidase in many organisms, including Y15627 from *P. simplicissimum*. This annotated aryl-alcohol oxidase is identical to P56216, the original sequence of VAO. Calling VAO an aryl-alcohol oxdiase is not technically wrong, but leads to confusion because the enzymes with EC number 1.1.3.7 that are defined as aryl-alcohol oxidases are part of the glucose-methanol-choline oxidoreductase family (Ferreira et al., 2009).

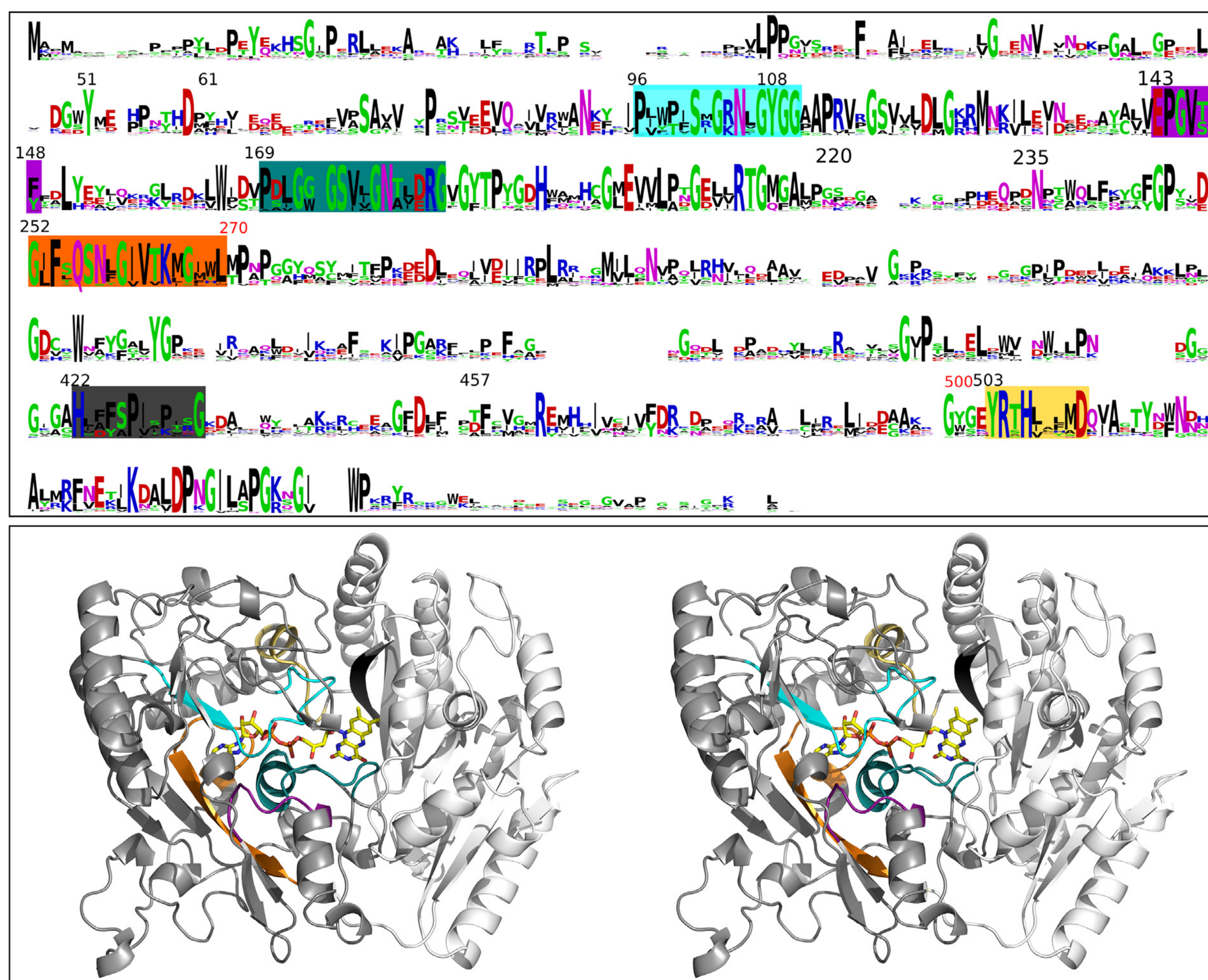


**Fig. 4.** Sequence logos of fungal VAOs and their location in the structure (PDBID 2VAO). Top: coloured boxes indicate the different motifs. Cyan highlights the 4PO-motif-1, purple the 4PO-motif-2, turquoise the 4PO-motif-3, orange the 4PO-motif-4 and yellow the 4PO-motif-5. Turquoise also highlights the VAO-motif-1, which overlaps with the 4PO-motif-3 and grey the VAO-motif-2. Gaps are introduced by the removed bacterial outgroup. Bottom: Stereo view of the structure of VAO with the same colours as above indicating the location of the different motifs defined in this study (see Table 2 for details). The FAD cofactor is shown as sticks and coloured in yellow. The cap domain (residues 270–500, interrupting the FAD binding domain) is shown in white and the FAD binding domain in grey.

We recommend using the 4PO-motif-1 for a knowledge-based search of novel 4POs or to improve the annotation of these sequences. The additional use of the other motifs will help to further narrow down putative functions. In addition to these motifs, several fingerprint residues can be used to identify VAOs in a given set of sequences (see Table 1). Including the original *Ps*VAO sequence (P56216.1) is highly recommended for ease of finding these fingerprints.

Cellular localisation of the candidate VAOs and a co-expressed catalase-peroxidase was predicted to be peroxisomal or cytoplasmic. This is in agreement with experimental data available on both enzymes from *P. simplicissimum* (Fraaije et al., 1998a, 1996). It is likely that fungi that intracellularly express highly active VAO also co-express a

**Table 2**
Motifs for 4POs and their starting position in *Ps*VAO, defined based on Fig. 4 (in PROSITE format). Bold letters indicate a critical amino acid residue.

| | |
|---|---|
| 4PO-motif-1; start 96 | P-x-x-x-x-S-x-G-[RK]-N-x-G-**Y**-G-[GS] |
| 4PO-motif-2; start 143 | E-P-G-V-[TS]-[FY] |
| 4PO-motif-3; start 178 | G-N-x-x-x-x-G |
| 4PO-motif-4; start 252 | G-x-F-x-Q-x-x-x-G-x-x-x-K-x-G-x-x-L |
| 4PO-motif-5; start 503 | **Y**-**R**-x-x-(x)-x-x-x-D |
| VAO-motif-1; start 169 | P-x-x-G-x-(x)-G-S-x-x-G-N-x-x-x-x-G |
| VAO-motif-2; start 422 | **H**-x-x-x-x-P-x-x-x-x-x-G |

catalase-peroxidase to prevent accumulation of toxic levels of hydrogen peroxide. However, the catalase-peroxidase genes we identified in the genomes of the red clade are not located in the immediate environment of the *vao* gene.

VAO and the other members of the 4PO subgroup are the only known members of the VAO/PCMH family that covalently bind their FAD cofactor via a residue in the cap domain. The other family members bind FAD either mono-covalently or bi-covalently via (a) residue(s) in the FAD-binding domain, or non-covalently. Tyrosines, cysteines and histidines are involved in the covalent binding modes (Cunane et al., 2000; Huang et al., 2005; Mattevi et al., 1997). Histidines can bind FAD via an 8α-$N^1$-histidyl (to the FAD-binding domain) or 8α-$N^3$-histidyl bond. *Ps*VAO uses His422 of the cap domain to install the 8α-($N^3$-histidyl)-FAD bond (Mattevi et al., 1997).

Interestingly, many putative VAOs do not contain His61 which is involved in the autocatalytic incorporation of FAD (Fraaije et al., 2000), but all contain His422 and none contain Tyr414. For putative VAOs with Tyr61 this residue may be able to have the same function. But for other putative VAOs, the autocatalytic incorporation of FAD, as observed for *Ps*VAO, is unlikely, or another mechanism is in place to activate His422. The absence of Tyr414 also negates the option of a 8α-(*O*-tyrosyl)-FAD, although two sequences with a cysteine at position 414 might have a 8α-(*S*-cysteinyl)-FAD. Some of these putative VAOs might bind their FAD cofactor non-covalently.

## 4. Conclusion

In conclusion, we have identified sequences, which represent with a high likelihood novel, fungal VAOs. Genomic analysis revealed presence of VAO in only a small subset of fungal species that are present in several fungal orders. Based on phylogeny and synteny analysis, it seems more likely that fungi that possess VAOs obtained these through horizontal transfer, possibly from bacteria, rather than that many fungi have lost this gene through evolution. We have defined a specific sequence motif (P-x-x-x-x-S-x-G-[RK]-N-x-G-Y-G-[GS]) that recognizes 4PO enzymes, that can be used in combination with several fingerprint residues to identify novel VAOs.

Several properties of the members of the 4PO subgroup remain confusing. Their substrate specificities overlap, but *Rj*EUGO is unable to convert alkylphenols. *Pp*PCMH and other bacterial 4POs as well as the fungal *Ps*VAO are able to convert these molecules enantioselectively. Their fungal or bacterial origin cannot be used to functionally separate them as the bacterial *Rj*EUGO is a true oxidase like *Ps*VAO, but the other bacterial 4POs are not. Also, these enzymes are induced by different molecules (Fraaije et al., 1997; Otani et al., 2014; Reeve et al., 1989), and no common inducer or physiological substrate has been identified (yet). Biochemical characterisation of the identified fungal sequences will help to shed light on whether there are “fungal EUGOs” or “fungal PCMHs” and additional work to identify and characterise bacterial members of the 4PO subgroup will also help to illuminate whether there are “bacterial VAOs”. More work is certainly needed to disentangle these enzymes from each other.

## 5. Methods

VAO is currently the only member of the Auxiliary Activity Family 4 (AA4) of the CAZy database. In CAZy, there are 21 eukaryotic and 131 bacterial sequences classified as AA4, presenting a clear bias against fungal sequences.

A manual search of the MYCOCOSM database (April 2017) revealed 89 putative VAOs. Gene models of enzymes annotated as CAZy-AA4 in the MYCOCOSM database were analysed. Sequences from non-published genomes were removed, greatly reducing the amount of sequences. Protein sequences and genomes were downloaded and gene models were manually corrected where necessary. Sequences with obvious internal STOP codons were removed. If unresolvable problems were present with one sequence from an organism, all sequences from that organism were removed. Sequences were aligned using clustal omega. The verified alignment was used to build a phylogenetic tree using MEGA (version 5.2). The raw data of the alignment and the tree can be found in the supplementary Data (supplementary data 1 and supplementary data 2).

Sequences originating from ascomycota and basidiomycota were identified, with the majority of sequences in the ascomycota phylum (78, vs 11 in basidiomycota). Gene models for 15 of these 89 uncharacterised fungal sequences were manually corrected. Organisms from different classes in the two phyla contained putative VAOs. The original sequence of *Ps*VAO was added to the 89 putative VAOs, as were the bacterial sequences of characterized 4POs, giving a total of 97 sequences, for which an alignment, and subsequently a phylogenetic tree, was built using the clustal omega webserver (https://www.ebi.ac.uk/Tools/msa/clustalo/) and the MEGA5 program (Tamura et al., 2011).

Sequence logos were created using the weblogo webserver ((Crooks et al., 2004), http://weblogo.berkeley.edu/) and edited with inkscape (https://inkscape.org/en/). The genomic environment was manually extracted from the MYCOCOSM database, only selecting ORFs on the same strand as the putative VAO. Visualisation was performed with inkscape. The DeepLoc webserver was used to analyse the probable localisation of putative VAOs ((Almagro Armenteros et al., 2017), http://www.cbs.dtu.dk/services/DeepLoc-1.0/). Estimated times of speciation events of the last common ancestor of organisms containing 4POs were median times given by http://timetree.org/. The proposed motifs were tested using http://www.genome.jp/tools/motif/MOTIF2.html, on 14.11.2017. Catalase-peroxidases (katGs) were identified in fungi in the red clade by blasting with Q96VT4. Cellular localisation of sequences of the best hits was analysed using the DeepLoc server ((Almagro Armenteros et al., 2017), http://www.cbs.dtu.dk/services/DeepLoc-1.0/).

## Acknowledgements

Funding: This work was supported by the European Union through the INDOX project (FP7-KBBE-2013-7-613549).

The authors state no conflict of interest.

## Conflict of interest

The authors state no conflict of interest.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found in the online version, at https://doi.org/10.1016/j.fgb.2018.04.003.

## References

Almagro Armenteros, J.J., Sønderby, C.K., Sønderby, S.K., Nielsen, H., Winther, O., 2017. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics 33, 3387–3395. http://dx.doi.org/10.1093/bioinformatics/btx431.

Brandt, K., Thewes, S., Overhage, J., Priefert, H., Steinbüchel, A., 2001. Characterization of the eugenol hydroxylase genes (*ehyA/ehyB*) from the new eugenol-degrading *Pseudomonas* sp. strain OPS1. Appl. Microbiol. Biotechnol. 56, 724–730. http://dx.doi.org/10.1007/s002530100698.

Carpena, X., Melik-Adamyan, W., Loewen, P.C., Fita, I., 2004. Structure of the C-terminal domain of the catalase-peroxidase KatG from *Escherichia coli*. Acta Crystallogr. Sect. D 60, 1824–1832. http://dx.doi.org/10.1107/S0907444904020621.

Crooks, G.E., Hon, G., Chandonia, J.-M., Brenner, S.E., 2004. WebLogo: a sequence logo generator. Genome Res. 14, 1188–1190. http://dx.doi.org/10.1101/gr.849004.

Cunane, L.M., Chen, Z.-W., Shamala, N., Mathews, F.S., Cronin, C.N., McIntire, W.S., 2000. Structures of the flavocytochrome *p*-cresol methylhydroxylase and its enzyme-substrate complex: gated substrate entry and proton relays support the proposed catalytic mechanism. J. Mol. Biol. 295, 357–374. http://dx.doi.org/10.1006/jmbi.1999.3290.

de Jong, E., van Berkel, W.J.H., van der Zwan, R.P., de Bont, J.A.M., 1992. Purification and characterization of vanillyl-alcohol oxidase from *Penicillium simplicissimum*. Eur. J. Biochem. 208, 651–657. http://dx.doi.org/10.1111/j.1432-1033.1992.tb17231.x.

Drijfhout, F.P., Fraaije, M.W., Jongejan, H., van Berkel, W.J.H., Franssen, M.C.R., 1998. Enantioselective hydroxylation of 4-alkylphenols by vanillyl alcohol oxidase. Biotechnol. Bioeng. 59, 171–177. http://dx.doi.org/10.1002/(SICI)1097-0290(19980720)59:2<171::AID-BIT5>3.0.CO;2-E.

Dym, O., Eisenberg, D., 2001. Sequence-structure analysis of FAD-containing proteins. Protein Sci. 10, 1712–1728. http://dx.doi.org/10.1110/ps.12801.

Efimov, I., Cronin, C.N., Bergmann, D.J., Kuusk, V., McIntire, W.S., 2004. Insight into covalent flavinylation and catalysis from redox, spectral, and kinetic analyses of the R474K mutant of the flavoprotein subunit of *p*-cresol methylhydroxylase. Biochemistry 43, 6138–6148. http://dx.doi.org/10.1021/bi035772x.

Efimov, I., Cronin, C.N., McIntire, W.S., 2001. Effects of noncovalent and covalent FAD binding on the redox and catalytic properties of *p*-cresol methylhydroxylase. Biochemistry 40, 2155–2166. http://dx.doi.org/10.1021/bi001644m.

Ein-Gil, N., Ilan, M., Carmeli, S., Smith, G.W., Pawlik, J.R., Yarden, O., 2009. Presence of *Aspergillus sydowii*, a pathogen of gorgonian sea fans in the marine sponge *Spongia obscura*. Isme J. 3, 752–755. http://dx.doi.org/10.1038/ismej.2009.18.

Ewing, T.A., Fraaije, M.W., Mattevi, A., van Berkel, W.J.H., 2017a. The VAO/PCMH flavoprotein family. Arch. Biochem. Biophys. 632, 104–117. http://dx.doi.org/10.1016/j.abb.2017.06.022.

Ewing, T.A., Gygli, G., van Berkel, W.J.H., 2016. A single loop is essential for the octamerization of vanillyl alcohol oxidase. FEBS J. 283, 2546–2559. http://dx.doi.org/10.1111/febs.13762.

Ewing, T.A., Nguyen, Q.-T., Allan, R.C., Gygli, G., Romero, E., Binda, C., Fraaije, M.W., Mattevi, A., van Berkel, W.J.H., 2017b. Two tyrosine residues, Tyr-108 and Tyr-503, are responsible for the deprotonation of phenolic substrates in vanillyl-alcohol oxidase. J. Biol. Chem. 292, 14668–14679. http://dx.doi.org/10.1074/jbc.M117.778449.

Ferreira, P., Hernandez-Ortega, A., Herguedas, B., Martínez, Á.T., Medina, M., 2009. Aryl-alcohol oxidase involved in lignin degradation: a mechanistic study based on steady and pre-steady state kinetics and primary and solvent isotope effects with two alcohol substrates. J. Biol. Chem. 284, 24840–24847. http://dx.doi.org/10.1074/jbc.M109.011593.

Fraaije, M.W., Pikkemaat, M., van Berkel, W.J.H., 1997. Enigmatic gratuitous induction of the covalent flavoprotein vanillyl-alcohol oxidase in *Penicillium simplicissimum*. Appl. Environ. Microbiol 63, 435–439.

Fraaije, M.W., Roubroeks, H.P., Hagen, W.R., Van Berkel, W.J.H., 1996. Purification and characterization of an intracellular catalase-peroxidase from *Penicillium simplicissimum*. Eur. J. Biochem. 235, 192–198. http://dx.doi.org/10.1111/j.1432-1033.1996.00192.x.

Fraaije, M.W., Sjollema, K.A., Veenhuis, M., van Berkel, W.J.H., 1998a. Subcellular localization of vanillyl-alcohol oxidase in *Penicillium simplicissimum*. FEBS Lett. 422, 65–68. http://dx.doi.org/10.1016/S0014-5793(97)01605-0.

Fraaije, M.W., van Berkel, W.J.H., Benen, J.A.E., Visser, J., Mattevi, A., 1998b. A novel oxidoreductase family sharing a conserved FAD-binding domain. Trends Biochem. Sci. 23, 206–207. http://dx.doi.org/10.1016/S0968-0004(98)01210-9.

Fraaije, M.W., van den Heuvel, R.H.H., van Berkel, W.J.H., Mattevi, A., 2000. Structural analysis of flavinylation in vanillyl-alcohol oxidase. J. Biol. Chem. 275, 38654–38658. http://dx.doi.org/10.1074/jbc.M004753200.

Fraaije, M.W., van den Heuvel, R.H.H., van Berkel, W.J.H., Mattevi, A., 1999. Covalent flavinylation is essential for efficient redox catalysis in vanillyl-alcohol oxidase. J. Biol. Chem. 274, 35514–35520. http://dx.doi.org/10.1074/jbc.274.50.35514.

Fraaije, M.W., Veeger, C., van Berkel, W.J.H., 1995. Substrate specificity of flavin-dependent vanillyl-alcohol oxidase from *Penicillium simplicissimum*: evidence for the production of 4-hydroxycinnamyl alcohols from 4-allylphenols. Eur. J. Biochem. 234, 271–277. http://dx.doi.org/10.1111/j.1432-1033.1995.271_c.x.

Furukawa, H., Wieser, M., Morita, H., Sugio, T., Nagasawa, T., 1999. Purification and characterization of vanillyl-alcohol oxidase from *Byssochlamys fulva* V107. J. Biosci. Bioeng. 87, 285–290. http://dx.doi.org/10.1016/S1389-1723(99)80033-4.

Garma, L.D., Medina, M., Juffer, A.H., 2016. Structure-based classification of FAD binding sites: a comparative study of structural alignment tools. Proteins Struct. Funct. Bioinform. 84, 1728–1747. http://dx.doi.org/10.1002/prot.25158.

Gayet, H., Revelant, D., Vibert, M., 2014. Method for the purification of natural vanillin. US2017204039.

Gygli, G., Lucas, M.F., Guallar, V., van Berkel, W.J.H., 2017. The ins and outs of vanillyl alcohol oxidase: identification of ligand migration paths. PLOS Comput. Biol. 13, 1–27. http://dx.doi.org/10.1371/journal.pcbi.1005787.

Heuts, D.P.H.M., Scrutton, N.S., McIntire, W.S., Fraaije, M.W., 2009. What's in a covalent bond? FEBS J. 276, 3405–3427. http://dx.doi.org/10.1111/j.1742-4658.2009.07053.x.

Huang, C.-H., Lai, W.-L., Lee, M.-H., Chen, C.-J., Vasella, A., Tsai, Y.-C., Liaw, S.-H., 2005. crystal structure of glucooligosaccharide oxidase from *Acremonium strictum:* a novel flavinylation of 6-S-cysteinyl, 8α-N1-histidyl FAD. J. Biol. Chem. 280, 38831–38838. http://dx.doi.org/10.1074/jbc.M506078200.

Jin, J., Mazon, H., van den Heuvel, R.H.H., Janssen, D.B., Fraaije, M.W., 2007. Discovery of a eugenol oxidase from *Rhodococcus* sp. strain RHA1. FEBS J. 274, 2311–2321. http://dx.doi.org/10.1111/j.1742-4658.2007.05767.x.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Bioinformatics 8, 275–282. http://dx.doi.org/10.1093/bioinformatics/8.3.275.

Lambert, F., Zucca, J., Mane, J., 2007. Production of ferulic acid, coniferyl alcohol and/or natural vanillin, comprises bioconversion of eugenol by a bacterium belonging to genus Streptomyces comprising at least a nucleotide sequence. FR2912758.

Leferink, N.G.H., Heuts, D.P.H.M., Fraaije, M.W., van Berkel, W.J.H., 2008. The growing VAO flavoprotein family. Arch. Biochem. Biophys. 474, 292–301. http://dx.doi.org/10.1016/j.abb.2008.01.027.

Levasseur, A., Drula, E., Lombard, V., Coutinho, P.M., Henrissat, B., 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol. Biofuels 6. http://dx.doi.org/10.1186/1754-6834-6-41.

Lyons, T.W., Reinhard, C.T., Planavsky, N.J., 2014. The rise of oxygen in Earth's early ocean and atmosphere. Nat. J. 506, 307–315. http://dx.doi.org/10.1038/nature13068.

Mattevi, A., Fraaije, M.W., Mozzarelli, A., Olivi, L., Coda, A., 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 5, 907–920. http://dx.doi.org/10.1016/S0969-2126(97)00245-1.

Nguyen, Q.-T., de Gonzalo, G., Binda, C., Rioz-Martínez, A., Mattevi, A., Fraaije, M.W., 2016. Biocatalytic properties and structural analysis of eugenol oxidase from *Rhodococcus jostii* RHA1: a versatile oxidative biocatalyst. ChemBioChem 17, 1359–1366. http://dx.doi.org/10.1002/cbic.201600148.

Otani, H., Lee, Y.-E., Casabon, I., Eltis, L.D., 2014. Characterization of *p*-hydroxycinnamate catabolism in a soil actinobacterium. J. Bacteriol. 196, 4293–4303. http://dx.doi.org/10.1128/JB.02247-14.

Priefert, H., Overhage, J., Steinbüchel, A., 1999. Identification and molecular characterization of the eugenol hydroxylase genes (*ehyA/ehyB*) of *Pseudomonas* sp. strain HR199. Arch. Microbiol. 172, 354–363. http://dx.doi.org/10.1007/s002030050772.

Reeve, C.D., Carver, M.A., Hopper, D.J., 1989. The purification and characterization of 4-ethylphenol methylenehydroxylase, a flavocytochrome from *Pseudomonas putida* JD1. Biochem. J. 263, 431–437. http://dx.doi.org/10.1042/bj2630431.

Scherer, M., Wei, H., Liese, R., Fischer, R., 2002. *Aspergillus nidulans* catalase-peroxidase gene (*cpeA*) is transcriptionally induced during sexual development through the transcription factor StuA. Eukaryot. Cell 1, 725–735. http://dx.doi.org/10.1128/EC.1.5.725-735.2002.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731–2739. http://dx.doi.org/10.1093/molbev/msr121.

van Berkel, W.J.H., de Jong, E., Fraaije, M.W., 1993. Enzymatic process for producing 4-hydroxy-cinnamyl alcohols. US5721125.

van den Heuvel, R.H.H., Fraaije, M.W., Ferrer, M., Mattevi, A., van Berkel, W.J.H., 2000a. Inversion of stereospecificity of vanillyl-alcohol oxidase. Proc. Natl. Acad. Sci. 97, 9455–9460. http://dx.doi.org/10.1073/pnas.160175897.

van den Heuvel, R.H.H., Fraaije, M.W., Laane, C., van Berkel, W.J.H., 1998. Regio- and stereospecific conversion of 4-alkylphenols by the covalent flavoprotein vanillyl-alcohol oxidase. J. Bacteriol. 180, 5646–5651.

van den Heuvel, R.H.H., Fraaije, M.W., Mattevi, A., van Berkel, W.J.H., 2000b. Asp-170 is crucial for the redox properties of vanillyl-alcohol oxidase. J. Biol. Chem. 275, 14799–14808. http://dx.doi.org/10.1074/jbc.275.20.14799.

van Rooyen, N., 2012. Identification, cloning and heterologous expression of fungal vanillyl-alcohol oxidases. University of the Free State.

Wierenga, R.K., Drenth, J., Schulz, G.E., Huber, R., 1983. Comparison of the three-dimensional protein and nucleotide structure of the FAD-binding domain of p-hydroxybenzoate hydroxylase with the FAD- as well as NADPH-binding domains of glutathione reductase. J. Mol. Biol. 167, 725–739. http://dx.doi.org/10.1016/S0022-2836(83)80106-5.

Zhao, X., Hersleth, H.-P., Zhu, J., Andersson, K.K., Magliozzo, R.S., 2013. Access channel residues Ser315 and Asp137 in *Mycobacterium tuberculosis* catalase-peroxidase (KatG) control peroxidatic activation of the pro-drug isoniazid. Chem. Commun. 49, 11650–11652. http://dx.doi.org/10.1039/C3CC47022A.

Zhao, X., Yu, H., Yu, S., Wang, F., Sacchettini, J.C., Magliozzo, R.S., 2006. Hydrogen peroxide-mediated isoniazid activation catalyzed by *Mycobacterium tuberculosis* catalase-peroxidase (KatG) and its S315T mutant. Biochemistry 45, 4131–4140. http://dx.doi.org/10.1021/bi051967o.