The functional and structural characterization of *Trichoderma reesei* dehydrogenase belonging to the PQQ dependent family of Carbohydrate-Active Enzymes Family AA12

Annick Turbe-Doan¹, Eric Record¹, Vincent Lombard², Rajender Kumar³(a), Anthony Levasseur¹(b), Bernard Henrissat²,³ and Marie-Line Garron²#

¹ INRA, Aix-Marseille Université, UMR1163, Biodiversité et Biotechnologie Fongiques, Marseille, France
² Architecture et Fonction des Macromolécules Biologique, UMR 7257 CNRS, Aix Marseille Université, USC 1408, Marseille, France
³ Department of Biological Sciences, King Abdulaziz University, Jeddah, 21589, Saudi Arabia

Present adress:

a present address: Department of Clinical Microbiology, Umeå University, SE-90185 Umeå, Sweden

b Present address: Aix-Marseille Université, Unité de Recherche sur les Maladies Infectieuses et Tropicales Emergentes (URMITE), UM63, CNRS 7278, IRD 198, INSERM U1095, IHU Méditerranée Infection, Pôle des Maladies Infectieuses, Assistance Publique-Hôpitaux de Marseille, Faculté de Médecine, 27 Bd Jean Moulin, 13005 Marseille, France

# Corresponding author: marie-line.garron@univ-amu.fr

Abstract
Pyrrolo-quinoline quinone (PQQ) is an ortho-quinone cofactor of several prokaryotic oxidases. Widely available in diet and necessary for the correct growth of mice, PQQ has been suspected to be a vitamin for eukaryotes. However, no PQQ-dependent eukaryotic enzyme had been identified to use the PQQ until 2014, when a basidiomycete enzyme catalyzing saccharide dehydrogenation using PQQ as a cofactor was characterized and served to define Auxiliary Activity family 12 (AA12). Here we report the biochemical characterization of the AA12 enzyme encoded by the ascomycete *Trichoderma reesei* (TrAA12). Surprisingly, only weak activity against uncommon carbohydrates like L-fucose or D-arabinose was measured. The three-dimensional structure of TrAA12 reveals important similarities with bacterial glucose dehydrogenases (s-GDH). The enzymatic characterization and the structure solved in the presence of calcium confirms the importance of this ion in catalysis, as observed for s-GDH. The structural characterization of TrAA12 was completed by modeling PQQ and L-fucose in the enzyme active site. Based on these results, family AA12 enzymes are likely to have a catalytic mechanism close to that of bacterial s-GDH.

### Importance

The Pyrrolo-quinoline quinone (PQQ) is an important co-factor synthesized by prokaryotes and involved in enzymatic alcohol and sugar oxidation. In eukaryotes, the benefit of PQQ as a vitamin has been suggested but never proved. Recently the first eukaryotic enzyme using PQQ was characterized in the basidiomycete *Coprinopsis cinerea*, demonstrating that fungi are able to use PQQ as an enzyme co-factor. This discovery led to the classification of the fungal PQQ-dependent enzymes in Auxiliary Activity family 12 (AA12) of the Carbohydrate Active Enzymes database (CAZy; www.cazy.org) classification. In the present paper, we report the characterization of ascomycete AA12 enzyme from *Trichoderma reesei* (TrAA12).
Our enzymatic and phylogenetic results show divergence with the only other member of the family characterized, from the basidiomycete *Coprinopsis cinerea*. The crystallographic structure of TrAA12 shows similarities to the global active site architecture of bacterial glucose dehydrogenases suggesting a common evolution between the both families.

**Introduction**

Pyrrolo-quinoline quinone (PQQ) is an *ortho*-quinone cofactor discovered at the end of the 70’s in a bacterial methanol dehydrogenase (1). Until recently, enzymes using PQQ as cofactor were thought to be found only in prokaryotes. Intriguingly, however, the distribution of PQQ is not limited to prokaryotes as it was also found in various plants or animals (2–4). The function of PQQ in eukaryotes is not clear but experiments have shown that PQQ is necessary in diet for a correct development of mice (5). The hypothesis that eukaryotes could use PQQ as a vitamin has been vigorously debated but eukaryotic PQQ-dependent enzymes remained elusive (6, 7).

In bacteria, the PQQ-dependent dehydrogenases are mainly synthesized by Gram negative strains. These enzymes are involved in alcohol dehydrogenation such as methanol dehydrogenases (MDH) or sugar dehydrogenation by glucose dehydrogenases (GDH) (8). GDHs can be divided in two distinct families, namely soluble s-GDHs and membrane-bound m-GDHs, the latter being closely related to the MDH family (8). No three-dimensional structure of m-GDH has been determined yet, but the sequence similarity with MDHs has allowed the modeling of a eight-blade β-propeller domain and the identification of the putative catalytic residues (9). The second family, s-GDH, corresponds to soluble GDHs which fold as a six-blade β-propeller (10). The active site of s-GDH is wide, commensurate with the broad substrate specificity of the enzyme (11) and it was shown that s-GDH use a general base-catalyzed hydride transfer assisted by calcium (12).
Carbohydrate active enzymes (CAZymes) are classified in sequence-based families in the CAZy database (www.cazy.org) (13). A new category of CAZymes, named Auxiliary Activities (AA), was recently introduced in CAZy after the oxidative cleavage of glycosides was established (13–16). In 2014, family AA12 was created after the characterization of a pyranose dehydrogenase (CcPDH) from the basidiomycete Coprinopsis cinerea (17). Interestingly, CcPDH used PQQ as co-factor to oxidize unusual monosaccharides such as D-glucosone or L-fucose (17, 18). The discovery and characterization of CcPDH established that some eukaryotes and in particular fungi do possess enzymes that use PQQ as a cofactor. To better understand the function of AA12 family enzymes in fungi, we performed a phylogenetic analysis of the family along with a functional and structural study of the enzyme from the ascomycete Trichoderma reesei (TrAA12), a fungus heavily utilized as the industrial model for cellulose degradation (19). The structure of TrAA12 reveals interesting structural similarities with bacterial s-GDH. We combined crystallographic structures and molecular docking simulations to explain the binding of PQQ and of the substrate in the catalytic binding site. The conservation of the catalytic machinery indicates that TrAA12 probably use the same general base-catalyzed hydride transfer as observed for s-GDH.

Materials and methods

Strains and culture media

Escherichia coli JM109 (Promega) was used for construction and propagation of vectors, and Aspergillus niger strain D15#26 (lacking the pyrG gene) was used for production of the recombinant protein (20). After cotransformation with vectors containing, respectively, the pyrG gene and the expression cassette containing the TrAA12 encoding gene, A. niger was grown for selection on solid minimal medium (without uridine) containing 70 mM NaNO₃, 7 mM KCl, 11 mM KH₂HPO₄, 2 mM MgSO₄, glucose 1% (wt/vol), and trace elements (500×
stock; 38 mM ZnSO₄, 12.5 mM MnCl₂, 9 mM FeSO₄, 3.55 mM CoCl₂, 3.2 mM CuSO₄, 3.1 mM Na₂MoO₄, 87 mM EDTA). In order to screen the transformants for the production of the recombinant protein, 100 ml of culture medium containing 70 mM NaNO₃, 7 mM KCl, 200 mM Na₂HPO₄, 2 mM MgSO₄, glucose 5% (wt/vol), and trace elements were inoculated with 2.10⁶ spores ml⁻¹ in a 500-ml baffled flask.

Expression vector construction and fungal transformation.

The sequence encoding amino acids 24 to 435 of T. reesei family AA12 protein (TrAA12) (GenBank EGR46829.1) was codon-optimized for A. niger, synthesized by GeneArt® Gene Synthesis (Life Technologies, Germany) and cloned into pAN52.3 (GenBank Z32689) by using restriction enzymes NcoI and HindIII. In the expression vector, the A. nidulans glyceraldehyde-3-phosphate dehydrogenase gene (gpdA) promoter, the 5′ untranslated region of the gpdA mRNA, and the A. nidulans trpC terminator were used to drive the expression of the gene coding for TrAA12.

Fungal cotransformations were carried out as described by Punt and van den Hondel using the expression vectors and pAB4-1 containing the pyrG selection marker, in a 5:1 ratio (21, 22). In addition, A. niger D15#26 was transformed with the pyrG gene without the expression vector for control experiment. Cotransformants were selected for uridine prototrophy on selective minimal medium plates (without uridine) and incubated for 5 days at 30°C. In order to screen transformants, two methods were used to screen 17 positive transformants. First, protoplasts were prepared from the transformant spores in Caylase (20 mg ml⁻¹) at 30°C for 3 h. Then, a PCR method was performed on the derived protoplasts by using specific primers (forward primer 5’ AACAGACCAACACCGTGCATCC 3’ and reverse primer 5’ CAACAAAGTTGGTGTGCCAGTGCAG 3’). Secondly, positive transformants were cultured in liquid medium and aliquots (350 µL) were collected after 7 days. Cells were
pelleted by centrifugation (5 minutes at 10,000 × g) and the resulting supernatant was concentrated onto micro-centrifugal units (Spin-X, Corning) with a 10 kDa cut-off, and protein production was confirmed by 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE).

**Phylogenetic analysis**

A phylogenetic analysis was performed on 59 fungal sequences downloaded from Genbank and from the fungal genome portal of the JGI (genome.jgi.doe.gov/programs/fungi/index.jsf). All selected sequences were from published genomes (23, 24). The amino acid sequences were aligned with MAFFT (25). We used the option --maxiterate 1000 and --global pair to obtain high accuracy where --maxiterate # indicates the maximum number of iterative refinement and global pair forces global pairwise alignment. To generate the tree we built a matrix of maximum likelihood distances based on model of substitution LG (26). Finally, the phylogeny reconstruction tree was performed using BIONJ (27).

**Purification of the recombinant protein.**

The best isolate corresponding to the transformant that produced an intense protein band was inoculated in the same conditions as in the screening procedure. The culture was harvested after 8 days of growth, filtered (0.22-μm pore size), and concentrated by ultrafiltration through a polyethersulfone membrane (molecular mass cutoff of 10 kDa) (Millipore). The concentrated fraction was dialyzed against a 50 mM Tris-HCl (pH 7.0) binding buffer, and the purification of His-tagged proteins was performed on HisTrap HP column (GE Healthcare Life Sciences) (28).

**Enzymatic assays and kinetics properties**
Dehydrogenase activity was determined at 40°C in a dye-linked system containing cytochrome c by spectrophotometry at 550 nm (ε=550=28 mM⁻¹ cm⁻¹). The standard reaction condition consisted of cytochrome c (50 μM), sodium acetate buffer (50 mM, pH 4.5), L-fucose (500 mM), PQQ (4 μM), CaCl₂ (1mM), with 4.6 μM of purified enzymes.

The dehydrogenase activity of the purified TrAA12 protein was assayed at various temperatures. In addition, the effect of temperature on enzyme stability was studied by incubating purified enzyme for 30 min at temperatures ranging from 30°C to 75°C. After this treatment, residual enzyme activity was determined under standard conditions. For pH dependency, the dehydrogenase activity was assayed in 50 mM sodium acetate buffer (pH 3.5-5.0) or 50 mM sodium maleate buffer (pH 6-7.0) at 40°C. For pH stability, the purified protein was incubated in the appropriate buffer for 1 hour before the assay.

To reach saturation, the kinetics measurements with L-fucose as the substrate were performed at concentrations ranging from 0.1 to 1M. Experiments were conducted at 40°C in sodium acetate buffer 50mM at pH 5.5, in the standard reaction conditions, 50μM cytochrome c, 1mM CaCl₂, 4μM PQQ. The kinetic parameters of recombinant TrAA12 were determined from a non-linear regression model using the Michaelis-Menten equation. Three replicates were taken for each point.

Crystallization and data collection

The first crystals of TrAA12 were obtained by the sitting drop method in commercial screen Stura (Molecular Limited Dimension, Suffolk, UK) containing 60% PEG550 MME, 0.1M HEPES pH 8.2. This commercial condition was optimized and the best crystallization conditions were 55-60% PEG550 MME, 0.1M HEPES pH 8.0-8.5. The protein concentration for crystallization was 24 mg ml⁻¹ in the purification buffer. The PEG 550 MME concentration was sufficient to protect the crystals during nitrogen flash freezing. Iodide
crystals were obtained by one to five minute soak in reservoir solution added with 1M of potassium iodide followed by a back soak and then freezing. Several soaks were done using calcium, PQQ and L-fucose to obtain a holoenzyme structure. The final concentrations in the drop during the soaking were approximately 1mM CaCl\textsubscript{2}, 5μM PQQ and 500mM L-fucose. Crystals were collected at the French national synchrotron facility, Soleil, on the Proxima 2 beamline for the native dataset, and on the Proxima 1 beamline for the iodide crystal. The crystals containing calcium were collected at ESRF on the ID30B beamline.

**Structure determination and refinement**

The crystals belong to P4\textsubscript{3} 2\textsubscript{1} 2 space group with cell parameters a=b=83 Å, c=143 Å. One protein is present in the asymmetric unit. Single-wavelength Anomalous Dispersion was used to obtain the phase. Data were collected at a wavelength of 1.653 Å. The initial phases were calculated with Phaser suite program (29, 30). Phase improvement was done by Parrot in the CCP4 program suite (30). Only one iodide ion was found in the structure. The first model was built by buccaneer and refined with Refmac5 in the CCP4 program suite (30–32). Coot was used to finalize the model building (33). The other structures were solved by molecular replacement using Refmac5 and coot programs (30, 32, 33). The assignment of the two calcium ions was verified by using the web server CheckMyMetal (34).

**Molecular docking studies**

To estimate the binding mode of the cofactor with the substrate, molecular docking studies were performed using the Autodock Vina Program (35). The cofactor (PQQ) and substrate (L-fucose: α-L-fucopyranose) were built and minimized with Gasteiger charges in UCSF Chimera (36). Further, minimized cofactor and substrate structures were prepared for docking studies applying standard ligand docking protocol. The structure of TrAA12 solved with the active calcium was chosen as the protein model. For protein preparation, all hydrogen atoms
added and non-polar hydrogens were also merged. Kollman united atom charge and atom type parameters were added. The calcium ion present in the active site was assigned a +2 charge. For cofactor and substrate, two separated docking sites were set in such way that included cofactor binding site and active site. Thirty dock poses were generated for each ligand. The best dock conformation was chosen with the lowest docking energy, metal ion distance and superimposition on reported homologue complex structure (PDB ID: 1CQ1) from the bacterium *Acinetobacter calcoaceticus* (37). Further, the interactions of protein-ligand complex, hydrogen bonds and their length were analyzed in UCSF Chimera (36).

**Results**

**Sequence analysis of family AA12**

The phylogenetic analysis of family AA12 performed on 59 fungal sequences (23, 24) subdivided the family in three subgroups (Fig1A). Two groups are directly reflecting taxonomy, i.e. coincide exclusively with *Ascomycota* and *Basidiomycota*, respectively. The third subgroup contains sequences from both phyla. Interestingly, several copies of family AA12 sequences which are encoded by the same genome can be found in separate subgroups of the phylogenetic tree. For example, the ascomycete *Leptosphaeria maculans* codes for three AA12, one is in the *Ascomycota* subgroup whereas the two others are in the mixed subgroup. The two sequences of *L. maculans* that are in the mixed subgroup present a bimodular organization with an AA8 domain fused to the N-terminus of the AA12 domain (Fig1A). The sequence analysis also shows that the AA8 domains and family 1 cellulose-binding modules (CBM1) are the only domains that are appended, together or not, to the AA12 domain. When present, the AA8 domain is always found at the N-terminus of the AA12 domain while the CBM1 is appended to the C-terminus (Fig1B). The AA8 domain is a fungal *b*-type cytochrome which use a heme as cofactor (38). Family AA8 was initially
characterized as a domain in cellobiose dehydrogenases (CDH) where it is appended to an AA3_1 domain (39). In CDH, AA8 is implicated in electron transfer during the dehydrogenation of the substrate by AA3_1 (40). Interestingly AA3_1 proteins present the same modularity as observed for AA12, i.e. AA8 at the N-terminus and a CBM1 at the C-terminus of the AA3_1 domain (Fig1B). The CBM1 module is a small peptide, mainly fungal, identified initially as cellulose binding domain (41). This domain can be appended to polypeptides in various families of Glycoside Hydrolases (GHs), Polysaccharides Lyases (PLs) as well as AAs domains. The bimodular architectures containing AA8 and AA12 are strictly found in the mixed subgroup of the phylogenetic tree (Fig1A). The presence of a signal peptide in most AA12 sequences suggests that these enzymes are secreted. Several AA12 sequences harbor a C-terminal transmembrane domain or a GPI anchor, presumably attaching the protein to the cell outer surface. The AA12 protein of T. reesei belongs to the Ascomycota clade (Fig1A). Unlike TrAA12, CrPDH is a multimodular protein composed of three domains where the AA12 domain is flanked by a AA8 and a CBM1 domain (Fig1B) (17). Among the analyzed sequences, this modularity with three domains is rarely observed and seems to be mainly occurring in the Agaricales order of the Basidiomycota. On the other hand, the bimodular structure containing the AA8 domain and the AA12 cannot be linked to a particular taxonomic clade and is more frequently observed than the AA8-AA12-CBM1 organization (Fig1A).

**Biochemical characterization of TrAA12**

After purification of TrAA12, SDS-PAGE revealed a single band with a molecular mass of approximately 75 kDa (data not shown), much higher than the theoretical 45.3 kDa based on the amino acid sequence from residue 24 to 435. Seven N-glycosylation sites were predicted based on the consensus sequence (Asn-Xaa-Ser/Thr) on Asn25, 94, 147, 184, 228, 306 and...
The glycosylation was confirmed by SDS-PAGE following PNGase treatment (data not shown) and was observed on Asn94, 147, 184, 306 and 341 in the crystallographic structures.

Various monosaccharides were tested to evaluate the enzyme activity and specificity. Several monosaccharides were tested in standard condition requiring calcium and PQQ as cofactor.

No detectable activity against D-glucose or D-fucose was found. TrAA12 oxidizes preferentially L-fucose, and a weak activity was measured against D-arabinose (11.6% relative activity), D-galactose (5.7%), L-arabinose (2.8%) and D-lyxose (2.2%) (Fig2A). The activity against D-glucosone was also tested but the activity of TrAA12 was 10 times lower than that measured with L-fucose. Unfortunately, the reductive activity of D-glucosone on cytochrome c was not stable enough to yield a reliable activity. The kinetic constants determined in steady-state for L-fucose in standard conditions revealed a $K_m$ of 99.9 (± 9.7) mM and a $k_{cat}$ of 0.012 (±0.0003) s$^{-1}$, resulting in a low catalytic efficiency $k_{cat}/K_m$ of 0.119 s$^{-1}$ M$^{-1}$. These kinetic data assume the full-occupancy of the cofactor, which under these experimental conditions remains uncertain. The temperature profile was determined up to 80°C, and TrAA12 activity increased gradually with the temperature between 30 and 80°C (Fig2B). The thermostability of the purified TrAA12 was tested in the range 0 – 75 °C. TrAA12 was stable up to 50°C and its activity decreased by 20% after heating at 60°C for 30 min. No activity was found after 30 min incubation at 70°C (Fig2C). The optimal pH was 4.5.

When the pH was below 3.5 and above 7, only 50% and 10% of the maximum activity were attained respectively (Fig2D). For stability in relation to pH, TrAA12 was stable from pH 3.5 to pH 8. Below pH 3.5, only 50% of the activity could be measured (Fig2E).

**Three-dimensional structures of TrAA12**

Three structures of TrAA12 were solved, namely that of the iodide adduct for phasing, of the native protein and of the catalytic calcium-containing protein (PDB codes 6I1Q, 6H7T and...
The three structures superimposed very well and no major differences in the main chains were observed. The polypeptide chain observable in the crystal structure of TrAA12 covers amino acids Gln24 to Asn435. This domain folds as a 6-bladed β-propeller (Fig3A). This fold is frequently found in various enzymes families where the number of blades varies from 4 to 12 (42). According to the common terminology, each blade of TrAA12 is numbered from 1 to 6 and contains four antiparallel β strands named from A to D (Fig3A). All the structures solved present a constitutive calcium. This calcium stabilizes a long loop which joins sheets C and D of the fourth blade. The calcium ion presents a typical octahedral coordination. Three amino acids are directly involved in the coordination, Glu259 and Gln301 with their side chains and Tyr276 by the main chain. Three water molecules complete the coordination, stabilized mainly by the side chains of Glu238 and Asn26. Of the seven predicted N-glycosylation sites, five are visible in the density maps on Asn94, 147, 184, 306 and 341. The density map quality varies between the structures and the last sugars are built on partial density. The two longer N-glycans, on Asn94 and 306, come close to the symmetric molecule and participate to the crystal packing.

To understand the enzymatic mechanism, we tried to solve the structure of TrAA12 in presence of the co-factor and of the substrate. Soaking experiments were performed on native crystals in crystallization buffer supplemented by calcium, PQQ and L-fucose. Several crystal soaking experiments were performed and eight data sets were collected. A large additional electron density inside a pocket and a reorganization of the amino acids surrounding this electron density were observed in each of the eight structures (Fig4). Only the best resolution structure was deposed in the PDB and is presented here (PDB code 6I1T). Based on the difference map peak intensity, a part of the additional electron density was assigned to a second calcium ion. The plausibility of calcium was confirmed by the web server CheckMyMetal (34). Without the cofactor, the coordination sphere is uncompleted but
predicted to be octahedral. The deviation observed in the angle geometry (gRMSD) is 13.9° which is only 0.4° higher than the acceptable value for this parameter. The coordination distances measured and compared to the Cambridge Structural Database (CSD) are in agreement with a calcium ion (Fig3B). Finally, after refinement, the calcium presents a B-factor of 33.9 Å² which is coherent with the average of 38.4 Å² for its environment. The calcium is directly coordinated by the side chain of Asp242, the carbonyl of the Ser240 and completed by two water molecules. A third water molecule completes this well-organized network (Fig3B). This organization is conserved in the native crystal where the calcium ion is replaced by a water molecule. The addition of the catalytic calcium and its coordinating water molecules were not sufficient to fully fit the additional density. The remainder of the density was too large and too dense to be filled only by water molecules (Fig4) but too small to correspond to the entire PQQ molecule, presumably indicating a weak and partial occupancy of the site (Fig4). In addition to PQQ, several molecules were tested to occupy the extra electron density but no satisfactory result was obtained. As the occupancy was too weak to place the PQQ molecule without ambiguity, the PDB file was deposited only with the catalytic calcium but without PQQ (PDB code 6I1T). Even if the additional density cannot be clearly assigned to PQQ, the molecule that enters the active site induced the movement and reorganization of the neighboring amino acids (Fig4). The superimposition of the catalytic calcium structure with the native one shows a movement of the loop spanning residues Thr353 to Thr357. This shift involves the reorientation of several side chains and the creation of new hydrogen bonds (Fig4). Asp354 presents the largest difference with a displacement of 1.55 Å of its Cα and a change of the orientation of its side chain. The new side chain position of Asp354 is stabilized by two new hydrogen bonds with Asp243 (Fig4). These movements partially close a part of the binding site and bring the side chains of Asp242 and Thr353 near the putative substrate binding site.
Structural homologs were searched with the DALI server (43). The closest structures (with Z-scores around 30) were from the bacterial soluble glucose dehydrogenase family (s-GDH). s-GDHs use PQQ as their cofactor and utilize a calcium-assisted mechanism to oxidize mono or disaccharides to the corresponding lactones (44). Several structures of s-GDHs are available in the PDB but only the structure of the *Acinetobacter calcoaceticus* enzyme (PDB code 1CQ1) was solved with the catalytic calcium, the PQQ and the substrate (12). The superimposition of *TrAA12* with *A. calcoaceticus* s-GDH shows that the global fold is well conserved (RMSD at 2.4Å, alignment done on 284 amino acids with 17% of sequence identity). Three calcium sites were identified in the *A. calcoaceticus* s-GDH structure but only two sites are strictly conserved in the other s-GDH structures (45, 46). Interestingly, the two sites are the same than those observed in the *TrAA12* structure. The first calcium is the constitutive calcium which has a well conserved coordination. Except for Gln301 in *TrAA12* which is replaced by water molecule in s-GDH, all the others amino acids involved in the calcium coordination, directly or indirectly via water molecules, are strictly conserved. The second calcium is involved in catalysis and interacts with PQQ in the binding site of s-GDHs. The coordination of the second calcium of *TrAA12* is comparable to that observed in *A. calcoaceticus* s-GDH (Fig5). In the two structures, the loop joining sheet B to sheet C of the fourth blade is important for the coordination of the calcium (Fig3A). In the *A. calcoaceticus* s-GDH structure, Gly247 and Pro248 in this loop stabilize the calcium ion. In *TrAA12*, Ser240 adopts the same role and position as Gly247 in *A. calcoaceticus* s-GDH. Because *TrAA12* has a longer loop than *A. calcoaceticus* s-GDH, the amino acid which follows directly Ser240 is not well positioned to interact with the calcium like the Pro248 in s-GDH. Instead, the long loop of *TrAA12* allows the side chain of Asp242 to complete the calcium...
coordination and replace one of the water molecules observed in the calcium coordination of *A. calcoaceticus* s-GDH (Fig5).

**Molecular docking of PQQ and L-fucose on TrAA12**

Because no density was observed for L-fucose and only a partial density was obtained for PQQ in our crystals, molecular docking studies were initiated to confirm the binding position of PQQ and to estimate the putative position of the substrate. The docking simulation was done in two steps. Firstly, PQQ was docked and the best complex protein-cofactor was kept. Then the best substrate position was calculated using the protein-cofactor complex. The best docked position of PQQ is close to the catalytic calcium. The PQQ molecule occupies a large flat cavity in front of the calcium pocket (Fig6A). The oxygen (O7A) of the carboxylate, the nitrogen of the amine N6 and the oxygen (O5) of the ketone are positioned at 2.9, 3 and 2.5 Å respectively from the calcium. These distances are compatible with electrostatic interactions which could be involved in the calcium stabilization and activation. The PQQ co-factor is surrounded mainly by basic amino acids. The side chains of Arg57, His153, Arg220, His330 and His350 are well positioned (with distances less than or close to 3 Å) for electrostatic and/or hydrogen bonds with the carboxylates or the carbonyl groups of PQQ. Hydrogen bonds with two acidic residues (Asp221 and Asp242) and the amino groups of the main chain of Thr353 and Asp354 complete the interaction (Fig6B). The floor of the PQQ cavity is constituted by the side chains of Asn239 and Ser332. On the other side, the PQQ molecule is totally exposed to the solvent and no amino acid hinders substrate accessibility (Fig6A). In conclusion, the docking simulation suggests that the putative substrate binding site is located above the PQQ molecule in the same cavity. The L-fucose residue docks onto the PQQ surface by hydrophobic stacking at around 4 Å distance. The sugar molecule could be stabilized by a hydrogen bond between the hydroxyl group at C2 and the side chain of Asp152. In L-fucose, the C4 hydroxyl group is in axial position and points down toward the
PQQ co-factor. This orientation allows the establishment of a hydrogen bond with the carboxylic function at C9 of PQQ. Finally, the methyl group of Thr353 comes close (3.5 Å) to the methyl group of L-fucose (Fig6B).

Discussion

Family AA12 was created in 2014 after the biochemical characterization of the *Coprinopsis cinerea* pyranose dehydrogenase, *CePDH* (17). Whereas PQQ was considered as a strictly prokaryotic cofactor, the discovery of family AA12 demonstrated that some eukaryotes, at least fungi, can use PQQ for monosaccharide dehydrogenation processes. Despite a poor sequence identity (<20%), *TrAA12* presents important structural similarities with bacterial s-GDHs, including the functional and the structural calcium sites. The PQQ molecule could not be built in the *TrAA12* structure because only a partial density was observed. Docking simulations placed the PQQ molecule in a position consistent with the extra density observed in crystal structure (Fig4). This extra density could correspond to the carboxylic group (C7), the nitrogen (N6) of the pyridine cycle and may be completed by the oxygen of the ketone (C5). This part of the cofactor is stabilized by the interaction with the calcium, providing a possible explanation to the residual density observable only around the calcium ion. The remainder of the PQQ molecule could be more mobile around this hinge. The predicted PQQ binding site is comparable to that observed in *A. calcoaceticus* s-GDH where the co-factor is stabilized mainly by electrostatic bonds involving well conserved basic amino acids (Fig5).

As mentioned above, the structure solved in presence of calcium and PQQ shows small loops movement and sides chains reorientation. The loop going from Thr353 to Thr357 comes closer to the extra density that could correspond to the carboxylic group (C7) of PQQ. In this position, the carboxylic group of PQQ can establish hydrogens bonds with the amines of the main chains of Thr353 and Asp354, which would explain the movement of the loop. This reorganization results in a narrower substrate binding site (Fig4).
In the s-GDH family, the oxidation mechanism involves a direct hydride transfer between substrate C1 and the C5 of the PQQ followed by a general based-catalyzed proton abstraction (37). Based on the quaternary structure of *A. calcoaceticus* s-GDH, His144 is the best candidate catalytic base, functionally assisted by calcium and Arg228 to enhance the reactivity of the C5-O5 bond of PQQ (12, 47). The orientation of the L-fucose, found by docking simulation on *TrAA12*, is the same as that observed for the D-glucose in the structure of *A. calcoaceticus* s-GDH (PDB code 1CQ1) (37). The proton at C1 position of the L-fucose is well oriented and points down towards C5 of PQQ, which would be coherent with a catalytic mechanism similar to that of s-GDH. Moreover, the equivalent of His144 in *A. calcoaceticus* s-GDH, His153, is close enough to the anomeric carbon of L-fucose (3 Å) to act as the general base. In addition to the catalytic histidine, the two amino acids surrounding the PQQ ketones (Arg228, Asn229 in *A. calcoaceticus* s-GDH) and potentially important for the catalytic activity are also conserved in *TrAA12* (Arg220, Asp221). If *TrAA12* operates with the same catalytic mechanism as s-GDH, then His153 would be the catalytic base. However, the phylogenetic analysis shows that this histidine is not strictly conserved in family AA12 (Fig7). In more than 25% of the sequences, this histidine is not conserved and is frequently replaced by a leucine. Interestingly, the conservation rate of His153 is directly linked to the subgroup classification. More than 93% of the sequences of the mixed subgroup harbor the conserved putative catalytic histidine whereas the rate drops to 78% for the sequences belonging to the *Ascomycota* and to only 30% for the *Basidiomycota* sequences. In comparison, the couple Arg/Asx (Arg220, Asp221 in *TrAA12*) which completes the PQQ interaction with the reactive ketones is more conserved (around 94% for each) (Fig7). The conservation of Arg/Asx dyad in family AA12 suggests an implication of these amino acids in catalysis. The structural similarities observed between *TrAA12* and s-GDH suggest that the two families have a common evolutionary ancestor. However, the high sequence divergence
around the putative catalytic residue in some subgroups raises questions on the conservation of the catalytic mechanism and/or of the substrate specificity/selectivity.

So far family AA12 is the first and the only known eukaryotic family using PQQ as cofactor. However, fungi do not possess the machinery to synthesize PQQ. This was previously also observed for some bacterial strains. For example, E. coli is not able to synthesize PQQ, but produces a PQQ-dependent dehydrogenase (48). The wide natural availability of PQQ allows these enzymes -that are extracellular- to find the cofactor directly in the medium (49). For E. coli the PQQ dehydrogenase is involved in its periplasmic respiratory network. The capacity to use PQQ in electron transfer system increases the bioenergetics options of the organisms in various environments (46). The metabolic pathway in which fungal family AA12 would be involved is unclear at present. Moreover, the substrates oxidized by family AA12 enzymes are uncommon and naturally infrequent. The oxidoreduction mechanism used by lytic polysaccharide monooxygenases (LPMOs) to degrade recalcitrant polysaccharides requires an electron transfer system (50). LPMOs can accept electrons from various donors. This versatility is utilized by fungi to breakdown the biomass at various steps of the degradation (50). Among the electron donors of LPMOs, several families of Auxiliary Activities are suspected to play a role. The best characterized one is the cellobiose dehydrogenase (CDH) which has a modular structure comprising a N-terminal family AA8 domain appended to a family AA3 domain, sometimes followed by a C-terminal CBM1 domain (51). Functional and structural data have shown electron transfer from domain AA3 to domain AA8 of CDH, followed by LPMO reduction to initiate oxygen activation (52). Interestingly, some AA12 enzymes display a modular organization similar to that of CDH, as CcPDH. Our phylogenetic analysis shows that the evolutionary pressure was different on modular AA12 proteins, since regardless of the phyla, multimodular enzymes are gathered in a separate subgroup and present a higher conservation rate of the putative catalytic histidine. A recent publication has
confirmed that family AA12 enzymes, associated with a AA8 domain, can activate AA9 LPMOs (53). It is thus possible that multimodular AA12, as CcPDH, could represent one of the links of the extracellular electrons transfer chain that regulates LPMO activity independently of changes in substrate composition. However, this hypothesis cannot be generalized to the entire AA12 family as the majority of AA12 sequences are non-modular. Moreover, experimentally TrAA12 diverges from CcPDH with a low enzymatic rate and different specificity. TrAA12 seems also to have a weaker affinity for the PQQ than previously reported for CcPDH.

In conclusion, the clades observed on the phylogenetic tree of family AA12 may reflect different functional activities within these subfamilies. Compared to CcPDH, further investigations are needed to identify the physiological substrate of TrAA12. Nevertheless, the important structural conservation observed between TrAA12 and the bacterial dependent PQQ enzymes comforts our hypothesis that TrAA12 requires PQQ as cofactor.

Acknowledgements. The authors are grateful to Pierre Legrand and Nicolas Foos for their useful help to collect and solve the structure of TrAA12 at the Proxima 1 beamline of French facility Soleil synchrotron.

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

References

1. Salisbury SA, Forrest HS, Cruse WB, Kennard O. 1979. A novel coenzyme from bacterial primary alcohol dehydrogenases. Nature 280:843–844.
2. Bishop A, Paz MA, Gallop PM, Karnovsky ML. 1994. Methoxatin (PQQ) in guinea-pig neutrophils. Free Radic Biol Med 17:311–320.

3. Kumazawa T, Sato K, Seno H, Ishii A, Suzuki O. 1995. Levels of pyrroloquinoline quinone in various foods. Biochem J 307 (Pt 2):331–333.

4. Mitchell AE, Jones AD, Mercer RS, Rucker RB. 1999. Characterization of pyrroloquinoline quinone amino acid derivatives by electrospray ionization mass spectrometry and detection in human milk. Anal Biochem 269:317–325.

5. Steinberg FM, Gershwin ME, Rucker RB. 1994. Dietary pyrroloquinoline quinone: growth and immune response in BALB/c mice. J Nutr 124:744–753.

6. Felton LM, Anthony C. 2005. Biochemistry: role of PQQ as a mammalian enzyme cofactor? Nature 433:E10; discussion E11-12.

7. Kasahara T, Kato T. 2003. Nutritional biochemistry: A new redox-cofactor vitamin for mammals. Nature 422:832.

8. Anthony C. 2004. The quinoprotein dehydrogenases for methanol and glucose. Arch Biochem Biophys 428:2–9.

9. Cozier GE, Anthony C. 1995. Structure of the quinoprotein glucose dehydrogenase of Escherichia coli modelled on that of methanol dehydrogenase from Methylobacterium extorquens. Biochem J 312 (Pt 3):679–685.

10. Oubrie A, Rozeboom HJ, Kalk KH, Duine JA, Dijkstra BW. 1999. The 1.7 A crystal structure of the apo form of the soluble quinoprotein glucose dehydrogenase from Methylobacterium extorquens.
Acinetobacter calcoaceticus reveals a novel internal conserved sequence repeat. J Mol Biol 289:319–333.

11. Olsthoorn AJ, Duine JA. 1998. On the mechanism and specificity of soluble, quinoprotein glucose dehydrogenase in the oxidation of aldose sugars. Biochemistry (Mosc) 37:13854–13861.

12. Oubrie A, Rozeboom HJ, Kalk KH, Olsthoorn AJ, Duine JA, Dijkstra BW. 1999. Structure and mechanism of soluble quinoprotein glucose dehydrogenase. EMBO J 18:5187–5194.

13. Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. 2014. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42:D490–D495.

14. Levasseur A, Drula E, Lombard V, Coutinho PM, Henrissat B. 2013. Expansion of the enzymatic repertoire of the CAZy database to integrate auxiliary redox enzymes. Biotechnol Biofuels 6:41.

15. Quinlan RJ, Sweeney MD, Lo Leggio L, Otten H, Poulsen J-CN, Johansen KS, Krogh KBRM, Jørgensen CI, Tovborg M, Anthonsen A, Tryfona T, Walter CP, Dupree P, Xu F, Davies GJ, Walton PH. 2011. Insights into the oxidative degradation of cellulose by a copper metalloenzyme that exploits biomass components. Proc Natl Acad Sci U S A 108:15079–15084.

16. Vaaje-Kolstad G, Westereng B, Horn SJ, Liu Z, Zhai H, Sørlie M, Eijsink VGH. 2010. An oxidative enzyme boosting the enzymatic conversion of recalcitrant polysaccharides. Science 330:219–222.
17. Matsumura H, Umezawa K, Takeda K, Sugimoto N, Ishida T, Samejima M, Ohno H, Yoshida M, Igarashi K, Nakamura N. 2014. Discovery of a eukaryotic pyrroloquinoline quinone-dependent oxidoreductase belonging to a new auxiliary activity family in the database of carbohydrate-active enzymes. PloS One 9:e104851.

18. Takeda K, Matsumura H, Ishida T, Samejima M, Ohno H, Yoshida M, Igarashi K, Nakamura N. 2015. Characterization of a novel PQQ-dependent quinohemoprotein pyranose dehydrogenase from Coprinopsis cinerea classified into auxiliary activities family 12 in carbohydrate-active enzymes. PloS One 10:e0115722.

19. Bischof RH, Ramoni J, Seiboth B. 2016. Cellulases and beyond: the first 70 years of the enzyme producer Trichoderma reesei. Microb Cell Factories 15:106.

20. Gordon CL, Archer DB, Jeenes DJ, Doonan JH, Wells B, Trinci AP, Robson GD. 2000. A glucoamylase::GFP gene fusion to study protein secretion by individual hyphae of Aspergillus niger. J Microbiol Methods 42:39–48.

21. van Hartingsveldt W, Mattern IE, van Zeijl CM, Pouwels PH, van den Hondel CA. 1987. Development of a homologous transformation system for Aspergillus niger based on the pyrG gene. Mol Gen Genet MGG 206:71–75.

22. Punt PJ, van den Hondel CA. 1992. Transformation of filamentous fungi based on hygromycin B and phleomycin resistance markers. Methods Enzymol 216:447–457.

23. Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martínez AT, Otillar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Copeland A, Coutinho PM, de Vries RP, Ferreira P, Findley K, Foster B, Gaskell J, Glotzer D,
Górecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, Kohler A, Kües U, Kumar TKA, Kuo A, LaButti K, Larrondo LF, Lindquist E, Ling A, Lombard V, Lucas S, Lundell T, Martin R, McLaughlin DJ, Morgenstern I, Morin E, Murat C, Nagy LG, Nolan M, Ohm RA, Patyshakuliyeva A, Rokas A, Ruiz-Dueñas FJ, Sabat G, Salamov A, Samejima M, Schmutz J, Slot JC, St John F, Stenlid J, Sun H, Sun S, Syed K, Tsang A, Wiebenga A, Young D, Pisabarro A, Eastwood DC, Martin F, Cullen D, Grigoriev IV, Hibbett DS. 2012. The Paleozoic origin of enzymatic lignin decomposition reconstructed from 31 fungal genomes. Science 336:1715–1719.

Kohler A, Kuo A, Nagy LG, Morin E, Barry KW, Buscot F, Canbäck B, Choi C, Cichocki N, Clum A, Colpaert J, Copeland A, Costa MD, Doré J, Floudas D, Gay G, Girlanda M, Henrissat B, Herrmann S, Hess J, Högborg N, Johansson T, Khouja H-R, LaButti K, Lahrmann U, Levasseur A, Lindquist EA, Lipzen A, Marmeisse R, Martino E, Murat C, Ngan CY, Nehls U, Plett JM, Pringle A, Ohm RA, Perotto S, Peter M, Riley R, Rineau F, Ruytinx J, Salamov A, Shah F, Sun H, Tarkka M, Tritt A, Veneault-Fourrey C, Zuccaro A, Mycorrhizal Genomics Initiative Consortium, Tunlid A, Grigoriev IV, Hibbett DS, Martin F. 2015. Convergent losses of decay mechanisms and rapid turnover of symbiosis genes in mycorrhizal mutualists. Nat Genet 47:410–415.

Katoh K, Kuma K, Toh H, Miyata T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511–518.

Le SQ, Gascuel O. 2008. An improved general amino acid replacement matrix. Mol Biol Evol 25:1307–1320.

Gascuel O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol Biol Evol 14:685–695.
Porath J, Carlsson J, Olsson I, Belfrage G. 1975. Metal chelate affinity chromatography, a new approach to protein fractionation. Nature 258:598–599.

McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. 2007. Phaser crystallographic software. J Appl Crystallogr 40:658–674.

Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel EB, Leslie AGW, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA, Powell HR, Read RJ, Vagin A, Wilson KS. 2011. Overview of the CCP4 suite and current developments. Acta Crystallogr D Biol Crystallogr 67:235–242.

Cowtan K. 2006. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta Crystallogr D Biol Crystallogr 62:1002–1011.

Murshudov GN, Vagin AA, Dodson EJ. 1997. Refinement of macromolecular structures by the maximum-likelihood method. Acta Crystallogr D Biol Crystallogr 53:240–255.

Emsley P, Cowtan K. 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr D Biol Crystallogr 60:2126–2132.

Zheng H, Chordia MD, Cooper DR, Chruszczyz M, Müller P, Sheldrick GM, Minor W. 2014. Validation of metal-binding sites in macromolecular structures with the CheckMyMetal web server. Nat Protoc 9:156–170.

Trott O, Olson AJ. 2010. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J Comput Chem 31:455–461.
36. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem 25:1605–1612.

37. Oubrie A, Rozeboom HJ, Kalk KH, Olsthoorn AJ, Duine JA, Dijkstra BW. 1999. Structure and mechanism of soluble quinoprotein glucose dehydrogenase. EMBO J 18:5187–5194.

38. Hallberg BM, Bergfors T, Bäckbro K, Pettersson G, Henriksson G, Divne C. 2000. A new scaffold for binding haem in the cytochrome domain of the extracellular flavocytochrome ceblobiose dehydrogenase. Struct Lond Engl 1993 8:79–88.

39. Henriksson G, Pettersson G, Johansson G, Ruiz A, Uzcategui E. 1991. Cellobiose oxidase from Phanerochaete chrysosporium can be cleaved by papain into two domains. Eur J Biochem 196:101–106.

40. Igarashi K, Yoshida M, Matsumura H, Nakamura N, Ohno H, Samejima M, Nishino T. 2005. Electron transfer chain reaction of the extracellular flavocytochrome ceblobiose dehydrogenase from the basidiomycete Phanerochaete chrysosporium. FEBS J 272:2869–2877.

41. Johansson G, Ståhlberg J, Lindeberg G, Engström Å, Pettersson G. 1989. Isolated fungal cellulose terminal domains and a synthetic minimum analogue bind to cellulose. FEBS Lett 243:389–393.

42. Kopec KO, Lupas AN. 2013. β-Propeller blades as ancestral peptides in protein evolution. PloS One 8:e77074.
43. Holm L, Rosenström P. 2010. Dali server: conservation mapping in 3D. Nucleic Acids Res 38:W545–W549.

44. Duine JA, Frank J, van Zeeland JK. 1979. Glucose dehydrogenase from Acinetobacter calcoaceticus: a “quinoprotein.” FEBS Lett 108:443–446.

45. Sakuraba H, Yokono K, Yoneda K, Watanabe A, Asada Y, Satomura T, Yabutani T, Motonaka J, Ohshima T. 2010. Catalytic properties and crystal structure of quinoprotein aldose sugar dehydrogenase from hyperthermophilic archaeon Pyrobaculum aerophilum. Arch Biochem Biophys 502:81–88.

46. Southall SM, Doel JJ, Richardson DJ, Oubrie A. 2006. Soluble aldose sugar dehydrogenase from Escherichia coli: a highly exposed active site conferring broad substrate specificity. J Biol Chem 281:30650–30659.

47. Oubrie A, Rozeboom HJ, Dijkstra BW. 1999. Active-site structure of the soluble quinoprotein glucose dehydrogenase complexed with methylhydrazine: a covalent cofactor-inhibitor complex. Proc Natl Acad Sci U S A 96:11787–11791.

48. Matsushita K, Arents JC, Bader R, Yamada M, Adachi O, Postma PW. 1997. Escherichia coli is unable to produce pyrroloquinoline quinone (PQQ). Microbiol Read Engl 143 (Pt 10):3149–3156.

49. Stites TE, Mitchell AE, Rucker RB. 2000. Physiological Importance of Quinoenzymes and the O-Quinone Family of Cofactors. J Nutr 130:719–727.

50. Kracher D, Scheiblbrandner S, Felice AKG, Breslmayr E, Preims M, Ludwika K, Haltrich D, Ejsink VGH, Ludwig R. 2016. Extracellular electron transfer systems fuel cellulose oxidative degradation. Science 352:1098–1101.
51. Phillips CM, Beeson WT, Cate JH, Marletta MA. 2011. Cellobiose dehydrogenase and a copper-dependent polysaccharide monooxygenase potentiate cellulose degradation by Neurospora crassa. ACS Chem Biol 6:1399–1406.

52. Tan T-C, Kracher D, Gandini R, Sygmund C, Kittl R, Haltrich D, Hällberg BM, Ludwig R, Divne C. 2015. Structural basis for cellobiose dehydrogenase action during oxidative cellulose degradation. Nat Commun 6.

53. Várnai A, Umezawa K, Yoshida M, Eijsink VGH. 2018. The pyrroloquinoline-quinone dependent pyranose dehydrogenase from Coprinopsis cinerea (CcPDH) drives lytic polysaccharide monooxygenase (LPMO) action. Appl Environ Microbiol.
Table 1. Crystallographic data collection and refinement statistics.

|                      | TrAA12-KI | TrAA12  | TrAA12-Ca2+ |
|----------------------|-----------|---------|-------------|
| **Data collection**  |           |         |             |
| Beamline             | Soleil proxima 1 | Soleil proxima 2 | ESRF-ID30B |
| Wavelength (Å)       | 1.6531    | 0.9801  | 0.9677      |
| Space group          | P4₁ 2₁ 2  | P4₁ 2₁ 2 | P4₁ 2₁ 2    |
| Cell dimension       | 82.78Å 82.78Å 140.75Å | 82.97Å 82.97Å 140.71Å | 82.95Å 82.95Å 146.9Å |
| Resolution range*    | 45.00 - 1.99 (2.05-1.99) | 46.9-2.1 (2.21-2.1) | 45.8-1.8 (1.89-1.8) |
| Completeness (%)     | 99.5 (93.93) | 100 (100) | 99.8 (100) |
| Unique reflections   | 32304     | 29499   | 45806       |
| R_{merge}           | 0.101 (0.705) | 0.109 (1.163) | 0.051 (0.883) |
| R_{pim}              | 0.03 (0.248) | 0.042 (0.441) | 0.019 (0.335) |
| Mean <I/σI>          | 20.4 (4.2) | 11.1 (1.9) | 21.5 (2.3) |
| B-factor from Wilson plot (Å²) | 33.5 | 39.5 | 30.9 |
| **Refinement statistics** |           |         |             |
| R_{cryst}            | 0.185 (0.271) | 0.204 (0.319) | 0.195 (0.328) |
| R_{free}             | 0.222 (0.298) | 0.256 (0.355) | 0.233 (0.399) |
| N° of free reflections | 1701    | 1460    | 2383        |
| Protein atoms        | 2988      | 2987    | 3018        |
| Other ligand atoms   | 129       | 118     | 105         |
| Solvent atoms        | 135       | 105     | 136         |
| r.m.s deviations from target values | | | |
| Bond lengths (Å)     | 0.013     | 0.017   | 0.013       |
| Bond angles (°)      | 1.81      | 1.99    | 1.75        |
| Chiral volumes (Å³)  | 0.137     | 0.109   | 0.079       |
| Average B-factors (Å²) | 35       | 44.9    | 38          |

Ramachandran plot statistics (%)
| Residues in favoured regions | 95.95 | 95.71 | 96.53 |
|----------------------------|-------|-------|-------|
| Residues in allowed regions | 3.29  | 3.54  | 2.73  |
| Residues in outlier regions | 0.76  | 0.76  | 0.74  |

PDB code | 6I1Q | 6H7T | 6I1T |
|---------|------|------|------|

* Throughout the table the values in parentheses apply for the outermost resolution shell

\[ R_{merge} = \frac{\sum hkl \sum i (I_{i(hkl)} - \langle I_{hkl} \rangle)}{\sum hkl \sum i I_{i(hkl)}} \]

\[ R_{merge} = \sum hkl \sqrt{\frac{1}{n - 1} \sum i |I_{i(hkl)} - \langle I_{hkl} \rangle|} / \sum hkl \sum i I_{i(hkl)} \]

\[ R_{cryst} = \sum hkl |F_o| - |F_c| / \sum hkl |F_o| \]

\[ R_{cryst} = \frac{\sum hkl |F_o| - |F_c|}{\sum hkl |F_o|} \]

where \( I \) is an individual reflection measurement and \( <I> \) is the mean intensity for symmetry related reflection.
Figures legend

Fig. 1 Phylogenetic tree of the AA12 family. (A) The phylogenetic tree is subdivided in three subgroups, colored orange in the Basidiomycota group, blue in the Ascomycota group and green in the mixed group. The proteins were identified by their Genbank accession numbers or their JGI protein IDs. The red and blue asterisks correspond to the TrAA12 and the CcPDH sequences. The black arrows indicate the multimodular sequences having at least both AA12 and AA8 domains. The three sequences from to Leptosphaeria maculans are highlighted in yellow. (B) The schematic modularity of TrAA12, CcPDH and CDH from Neurospora crassa (Uniprot Q7RXM0). Signal peptides are colored in purple, linkers in grey, AA12 domains in dark blue, AA8 in yellow, AA3 in orange and CBM1 in green.

Fig. 2 Enzymatic characterization of TrAA12. (A) Specificity of TrAA12 against various substrates in presence of 500mM of the electron donor. The activities are given relative to L-fucose. (B-E) The effects of temperature and pH on the activity of the purified TrAA12. Various temperatures ranging from 30°C to 80 °C and pH values (3.5 to 7) were tested under standard conditions. The effects of temperature (C) and pH (E) on the stability of the purified protein after 1 hour in 100mM tartrate buffer (pH 2-3), 100mM sodium acetate buffer (pH 3.5-6), 100mM HEPES (pH 7-8). All assays were performed with L-fucose as substrate.

Fig. 3 Crystallographic structure of TrAA12. (A) The structure of TrAA12 presented is the one solved with the two calcium ions (PDB 6I1T), which are represented by green spheres. The structure is coloured from blue (N-terminus) to red (C-terminus). Glycosylation observed in the structure are represented as pink sticks (B) The coordination network stabilizing the putative catalytic calcium. The amino acids are in stick representation and the red spheres
represent the water molecules. The black dotted lines represent electrostatics forces and
distances were measured by the program Coot (33). The electron density observable on the
2FoFc map after refinement is represented in mesh contoured at 1.5 σ.

Fig. 4 Superimposition of the native and the calcium bound structures. The native
structure is coloured in green (PDB code 6H7T) and the structure with catalytic calcium in
cyan (PDB code 6I1T). The docked PQQ is represented in stick and coloured in pink. The
residual electron density observable on the FoFc map after refinement is represented in mesh
contoured at 3 σ.

Fig. 5 Structural similarity. Superimposition of TrAA12 docking structure with that of s-
GDH from Acinetobacter calcoeticus in presence of PQQ (PDB code 1CQ1). The calcium
ions of the two structures are perfectly superimposed and represented as green spheres, amino
acids and PQQ are in stick representation, TrAA12 is coloured in cyan and s-GDH in orange.

Fig. 6 Molecular docking of PQQ and L-fucose. (A) The overall structure of TrAA12 in
presence of the catalytic calcium is represented in cyan surface. The molecules docked in the
putative active site are represented in stick and coloured in pink for PQQ and in green for L-
fucose. The calcium ion and the water molecules are observable at the bottom of the cavity
whereas PQQ is largely exposed to the solvent. (B) The amino acids potentially involve in
PQQ and substrate interaction are shown in stick representation. The black dotted lines
represent electrostatics forces and distances were measured by the program Coot (33). An
asterisk indicates the carbone 1 of the L-fucose.
Fig. 7 **Sequence alignment.** The red and blue asterisks correspond to the *TrAA12* and the *CcPDH* sequences. The black bars present the putative catalytic amino acids, His153, Arg220 and Asn221 in *TrAA12*. The amino acids are colored based on the percentage of identity.
Fig 1.

---

A

Basidiomycota

Ascomycota

Mixed group

B

* TrAA12

* CcPDH

CDH from Neurospora crassa
Fig 4.
Fig5.
Fig. 6.
