20/03/2014
AEM00490-14-Revised-Draft
Applied & Environmental Microbiology

Directed evolution of Unspecific Peroxygenase from *Agrocybe aegerita*

Patricia Molina-Espeja[1], Eva Garcia-Ruiz[2], David Gonzalez-Perez[1], René Ullrich[3], Martin Hofrichter[3] and Miguel Alcalde[1]

[1] Department of Biocatalysis, Institute of Catalysis, CSIC, 28049 Madrid, Spain.

[2] Department of Chemical and Biomolecular Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

[3] Department of Bio- and Environmental Sciences, TU Dresden – International Institute Zittau, 02763 Zittau, Germany.

[*]Corresponding author: Miguel Alcalde, Department of Biocatalysis, Institute of Catalysis, CSIC, Cantoblanco, 28049 Madrid, Spain.

Phone: +34 915854806, Fax: +34 915854760; Email: malcalde@icp.csic.es.

Running title: directed evolution of unspecific peroxygenase

ABSTRACT

Unspecific peroxygenase (UPO) represents a new type of heme-thiolate enzyme with self-sufficient mono(per)oxygenase activity and many potential applications in organic synthesis. With a view to taking advantage of these properties, we subjected the *Agrocybe aegerita* UPO1 encoding gene to directed evolution in *Saccharomyces cerevisiae*. To promote functional expression, several different signal peptides were fused to the mature protein and the resulting products tested. Over 9,000 clones were screened using an *ad-hoc* dual-colorimetric assay that assessed both peroxidative and oxygen-transfer activities. After 5 generations of directed evolution combined with hybrid approaches, 9 mutations were introduced that resulted in a 3,250-fold total activity improvement with no alteration in protein stability. A breakdown between secretion and catalytic activity was performed by replacing the native signal peptide of the original parental type with that of the evolved mutant: the evolved leader increased functional expression 27-fold whereas a 18-fold improvement in the k_{cat}/K_m for oxygen transfer activity was obtained. The evolved UPO1 was active and highly stable in the presence of organic co-solvents. Mutations in the hydrophobic core of the signal peptide contributed to enhance functional expression up to 8 mg/L, while catalytic efficiencies for peroxidative and oxygen transfer reactions were increased by several mutations in the vicinity of the heme-access channel. Overall, the directed evolution platform described is a valuable point of departure for the development of customized UPOs with improved features and for the study of structure-function relationships.

INTRODUCTION

The unspecific peroxygenase (UPO, EC 1.11.2.1; also known as aromatic peroxygenase, APO) is secreted by the edible mushroom *Agrocybe aegerita* and it belongs to a new type of peroxide-using enzymes that are of considerable interest due to their wide range of potential biotechnological applications (1). UPO is a compact protein with a thiolate axial ligand of the pivotal Fe^{3+} that governs the heme domain. Accordingly, UPO is classified as a member of the heme-thiolate peroxidase (HTP) superfamily, along with chloroperoxidase (CPO) from the ascomycete *Leptoxyphium fumago*, even though CPO is not capable of epoxidizing aromatic rings or hydroxylating alkanes like UPO (2).

The diversity of enzymes in the HTP superfamily is conferred by 2 distinct clusters ("long" and "short" UPOs), which are included in basidiomycetes, ascomycetes and other fungal sequences. UPO-type genes and proteins have been isolated and characterized in *Coprinellus radians*, *Marasmius rotula* and *Coprinopsis cinerea* (3-5). Indeed, to date over 1,000 UPO-like genes have been identified in genetic databases and in basidiomycete genome sequencing, indicating an ancient origin and a widespread distribution of UPO in nature (6, 7). With over 300 positively-tested substrates, UPOs exhibit considerable promiscuity in oxidation reactions, making them potentially attractive industrial biocatalysts. The versatile peroxide-dependent monooxygenase activity of UPO, which is based on a 2-electron oxygenation mechanism (*i.e.,* peroxygenase activity), is of particular significance as selective oxyfunctionalizations are among the most important reactions in organic synthesis (8). The array of oxygen transfer reactions catalyzed by UPO includes bromide oxidation, sulfoxidation, *N*-oxidation, aromatic peroxygenation, double bond epoxidation, hydroxylation of aliphatic

compounds and ether cleavage (2). Fuelled by catalytic amounts of H_2O_2, UPO acts as a self-sufficient monooxygenase through a complex catalytic mechanism that joins the reactive intermediates of heme-peroxidases and P450s (the "peroxide shunt" pathway) (7). Moreover, a UPO oxoiron(IV) protoporphyrin radical cation intermediate (UPO compound I) was recently described and proposed as the main active oxygen species involved in the mono(per)oxygenase activity of UPO (9, 10). Thus, UPO may be considered to be the missing link between P450 monooxygenases and heme-peroxidases (7).

Despite much biotechnological interest in this enzyme, no protein engineering studies have attempted to adapt its unique features to the requirements of specific biotransformation processes in industrial settings. Moreover, endeavors to gain a better understanding of UPO´s complex mechanism of action are hindered by the absence of tools with which to design mutants. Indeed, the successful adaptation of UPO has been hampered by many of the same bottlenecks that have precluded the engineering of CPO for decades, including several obstacles that prevent successful functional expression of HTPs in heterologous hosts (*e.g.,* different codon usage, cumbersome post-translational modifications, and heme-thiolate prosthetic group attachments (11, 12).

Escherichia coli and *Saccharomyces cerevisiae* are the most attractive heterologous hosts in which directed evolution can be performed (13). *S. cerevisiae* is a particularly versatile vehicle for the functional expression and directed evolution of fungal genes involved in lignin modification (including laccases and peroxidases), and it has been used in the directed evolution of versatile peroxidases (VP) for functional expression and stabilization, whereby medium redox potential laccases have been engineered to confer high secretion levels and activity in organic co-solvents (14-16). More recently, this

host has been used in the design of high-redox potential laccases (HRPLs) that are active in human blood and to develop chimeric laccases with combined properties (17-20). The number of protocols developed for the generation of DNA diversity in yeast is steadily increasing and as such, the *in vivo* homologous recombination machinery of this host can be used to enrich mutant libraries (21-23). These strategies have helped to extend the study of *S. cerevisiae* into the fields of synthetic biology and metabolic engineering, highlighting a wide range of potential applications ranging from biofuel production to novel green processes (24-26).

Here, for the first time we describe the use of directed evolution to produce a soluble, active and highly stable form of UPO in *S. cerevisiae*. Several fusion genes were tested to increase initial secretion levels, which were then further optimized by iterative rounds of random mutagenesis, DNA recombination and semi-rational strategies. The enzyme's substrate promiscuity was maintained by simultaneously performing a dual high-throughput screening (HTS) assay to efficiently explore mutant libraries without altering protein stability. The final mutant produced was comprehensively characterized and exhibited markedly improved kinetic properties, secretion and stability over a range of temperatures, as well as in the presence high concentrations of co-solvents.

MATERIAL AND METHODS

Laboratory Evolution: General Aspects

The original parental n-UPO1, and the α-UPO1, α*-UPO1, n*-UPO1 and n*-3F10 fusion genes were constructed as described in the Supplemental Material and Methods. After each round of directed evolution, PCR products were loaded onto a preparative agarose gel and then purified using the Zymoclean Gel DNA Recovery kit (Zymo Research). The recovered DNA

fragments were cloned under the control of the GAL1 promoter of the pJRoC30 expression shuttle vector, using BamHI and XhoI to linearize the plasmid and remove the parent gene. The linearized vector was loaded onto a low melting-point preparative agarose gel and purified as described above. The mutational loads, recombination strategies, library sizes and general conditions for each cycle of evolution are described in **Table S1**. All the primers used in this study are listed in **Table S3**.

First Generation

Four libraries were devised using n-UPO1 and α-UPO1 as the parental types. For each parent, 2 different mutagenic PCR strategies were used: Taq DNA polymerase (Sigma) in the presence of $MnCl_2$ (1-3 mutations/1,000 bp); and the Genemorph II kit (4-9 mutations/1,000 bp) (Stratagene, Mutazyme II). The PCR reaction for Taq/$MnCl_2$ was performed in a final volume of 50 µL containing 3% DMSO, 90 nM RMLN, 90 nM RMLC, 0.3 mM dNTPs (0.075 mM each), 0.01 mM $MnCl_2$, 1.5 mM $MgCl_2$, 0.05 U/µL Taq DNA polymerase and 0.1 ng/µL of the corresponding template. The PCR reactions for Mutazyme II were carried out in a final volume of 50 µL containing 3% DMSO, 0.37 µM RMLN, 0.37 µM RMLC, 0.8 mM dNTPs (0.2 mM each), 0.05 U/µL Mutazyme II and 300 ng of the corresponding initial target template (2,800 ng of pJR-n-*upo*1 and 2,566 ng of pJR-α-*upo*1). Error-prone PCR was performed on a gradient thermocycler (Mycycler, Bio-Rad) using the following parameters: 95ºC for 2 min (1 cycle); 94ºC for 45 s, 53ºC for 45 s, 74ºC for 3 min (28 cycles); 74ºC for 10 min (1 cycle). The PCR products (200 ng) were mixed with the linearized plasmid (100 ng) and transformed into competent *S. cerevisiae* cells using the Yeast Transformation kit (Sigma). To promote *in vivo* ligation, ~50 bp overhangs homologous to the linear vector were designed. Transformed cells were plated on SC drop-out plates and incubated for 3 days at 30 ºC.

160 Colonies containing the whole autonomously replicating vector were selected
161 and subjected to the dual HT-screening assay, and additional re-screening as
162 described in the Supplemental Material and Methods.

163 **Second Generation**

164 The best mutants obtained from the first generation (1A11 and 3C2) were
165 submitted to error prone PCR (Taq/$MnCl_2$ and Mutazyme II) as well as *in vivo*
166 DNA shuffling. The mutagenic rates, the PCR conditions and the thermal
167 cycling program employed were the same as those described for the first
168 generation. Mutated PCR products were mixed with the linearized vector (4:1
169 ratio of PCR products:linearized plasmid) and transformed into competent *S.*
170 *cerevisiae* cells in order to promote *in vivo* DNA shuffling.

171 **Third Generation**

172 The best mutant from the second round of evolution (12C12 mutant) was
173 subjected to 2 different processes.

174 i) *In vivo* Assembly of Mutant libraries. A recombined mutant library was built
175 by *in vivo* assembly of mutant libraries constructed with different mutational
176 spectra (IvAM, (27)). Taq/$MnCl_2$ and Mutazyme II libraries were mixed in
177 equimolar amounts and transformed into competent *S. cerevisiae* cells along
178 with the linearized vector as described above (8:1 ratio of mutant
179 library:vector).

180 ii) Focused domain mutagenesis at the signal peptide The 12C12 signal
181 sequence was independently subjected to random mutagenesis by MORPHING
182 (Mutagenic Organized Recombination Process by Homologous *In vivo*
183 Grouping) (28). Mutagenic PCR was prepared in a final volume of 50 µL
184 containing 3% DMSO, 90 nM RMLN, 90 nM Morphing psn apo1 rev, 0.3 mM
185 dNTPs (0.075 mM each), 0.1 mM $MnCl_2$, 1.5 mM $MgCl_2$, 0.05 U/µL Taq
186 polymerase DNA and 0.92 ng/µL template. The amplification parameters were:

187 95ºC for 2 min (1 cycle); 94ºC for 45 s, 50ºC for 45 s, 74ºC for 30 s (28 cycles);
188 and 74ºC for 10 min (1 cycle). The remaining portion of the whole UPO1 gene
189 was amplified by high-fidelity PCR in a final volume of 50 µL containing 3%
190 DMSO, 0.5 µM Morphing psn apo1 dir, 0.5 µM RMLC, 1 mM dNTPs (0.25 mM
191 each), 0.02 U/µL iProof DNA polymerase and 0.2 ng/µL template. High fidelity
192 PCR was carried out on a gradient thermocycler using the following
193 parameters: 98ºC for 30 s (1 cycle); 98ºC for 10 s, 55ºC for 25 s, 72ºC for 45 s
194 (28 cycles); and 72ºC for 10 min (1 cycle). The whole gene was *in vivo*
195 reassembled and recombined by transforming the different PCR products into
196 *S. cerevisiae* competent cells, a process facilitated by ~50 bp overhangs
197 flanking each recombination area. The DNA transformation mixture was
198 composed of linearized plasmid (100 ng) mixed with the mutagenized leader
199 (200 ng) and the mature non-mutagenized protein (200 ng).

200 **Fourth Generation**

201 i) Error-prone PCR and *in vivo* DNA shuffling. Mutagenic PCR reactions were
202 performed separately with mutants I13D3, M5D2 and M4D8. Mutated PCR
203 products were mixed with the linearized vector (at a PCR product:linearized
204 plasmid ratio of 6:1) and transformed into competent *S. cerevisiae* cells to
205 promote *in vivo* DNA shuffling.

206 ii) Site-directed mutagenesis. The I13D3 mutant from the third generation was
207 used as a template to introduce F[12]Y, A[14]V and R[15]G mutations using *In*
208 *Vivo* Overlap Extension (IVOE) (29). Two high-fidelity PCR reactions were
209 performed in a final volume of 50 µL, containing: (i) 3% DMSO, 0.5 µM RMLN,
210 0.5 µM PSN*R, 1 mM dNTPs (0.25 mM each), 0.02 U/µL iProof DNA
211 polymerase and 0.2 ng/µL template; or (ii) 3% DMSO, 0.5 µM PSN*F, 0.5 µM
212 RMLC, 1 mM dNTPs (0.25 mM each), 0.02 U/µL iProof DNA polymerase and
213 0.2 ng/µL template. The following PCR parameters were used for each

reaction: (i) 98°C for 30 s (1 cycle), 98°C for 10 s, 47°C for 25 s, 72°C for 15 s (28 cycles), and 72°C for 10 min (1 cycle); ii) 98°C for 30 s (1 cycle), 98°C for 10 s, 52°C for 25 s, 72°C for 40 s (28 cycles), and 72°C for 10 min (1 cycle). Both PCR products (200 ng each) were mixed with the linearized vector (100 ng) and transformed into *S. cerevisiae* for *in vivo* gene reassembly and cloning. Overlapping areas of ~ 50 bp flanking each segment were created to maximize the efficiency of *in vivo* DNA splicing between fragments.

Fifth Generation

The V57A mutation from 22A10 was introduced into 2A12 by mutational recovery through IVOE. Two high-fidelity PCR reactions were performed in a final volume of 50 µL, containing: (i) 3% DMSO, 0.5 µM RMLN, 0.5 µM 2A12*REV, 1 mM dNTPs (0.25 mM each), 0.02 U/µL iProof DNA polymerase and 0.2 ng/µL template; or (ii) 3% DMSO, 0.5 µM 2A12*DIR, 0.5 µM RMLC, 1 mM dNTPs (0.25 mM each), 0.02 U/µL iProof DNA polymerase and 0.2 ng/µL DNA template. The following PCR parameters were used for each reaction: (i) 98°C for 30 s (1 cycle), 98°C for 10 s, 47°C for 25 s, 72°C for 15 s (28 cycles), and 72°C for 10 min (1 cycle); (ii) 98°C for 30 s (1 cycle), 98°C for 10 s, 52°C for 25 s, 72°C for 35 s (28 cycles), and 72°C for 10 min (1 cycle). Both PCR products (200 ng each) were mixed with the linearized vector (100 ng) and transformed into *S. cerevisiae* for *in vivo* gene reassembly and cloning as described above.

High-throughput Screening

Peroxidative and peroxygenase activities were screened with the help of a dual-HT assay based on the oxidation of ABTS and the hydroxylation of NBD (which is in turn spontaneously cleaved to form the chromophore 4-nitrocatechol), as indicated in the Supplemental Material and Methods. Three consecutive re-

240 screenings were carried out to rule out the selection of false positives, which
241 included a thermostability assay for the estimation of T_{50} values.

242 **Purification and Biochemical characterization**

243 n*-UPO1, the PaDa-I mutant and $_{wt}$UPO1 were produced, purified and
244 biochemically characterized as described in the Supplemental Material and
245 Methods.

246 ## RESULTS AND DISCUSSION

247 **Point of Departure: Construction of Fusion Genes and Design of HTS-**
248 **Assay**

249 The starting point of this study was the cDNA (*upo1* gene) coding for the
250 unspecific peroxygenase from *Agrocybe aegerita* –Accession Nº FM872457- (30,
251 31). This gene encodes a protein of 328 amino acids plus a 43 amino acid
252 signal peptide that directs secretion in *A. aegerita*. To achieve sufficient
253 expression in the heterologous host in order to begin directed evolution,
254 several constructs were prepared that contained the native signal sequence (n-
255 UPO1), the α-factor prepro-leader from *S. cerevisiae* (α-UPO1) and the evolved
256 α-factor prepro-leader (α*-UPO1). The evolved α-factor prepro-leader (α*) was
257 previously engineered in association with a HRPL, achieving functional levels
258 of expression after 8 rounds of laboratory evolution (18). The α* construct
259 contained V10D-N23K-A87K mutations that boosted the expression of other
260 HRPL genes (17, 32, 20) and that may also enhance UPO1 secretion. The
261 secretion in each of the fusion genes was determined in 96-well plate
262 microfermentations (mU-ABTS/L): n-UPO1, 149; α-UPO1, 74; α*-UPO1,
263 negligible. While the use of evolved α-factor prepro-leaders as universal
264 peptides for heterologous expression has been proposed (33), this approach
265 appears to work only when signal sequences are switched between protein

templates with a high degree of sequence identity (as in the case of HRPLs), indicating that they cannot be used in other less related systems.

Several substrates were tested to develop a screening assay for directed UPO evolution. These included benzyl alcohol, veratryl alcohol, 5-nitro-1,3-benzodioxole (NBD), 2,6-dimethoxyphenol (DMP), 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulphonic acid) (ABTS) and *p*-nitro-phenoxy carboxylic acid (*p*NCA). Given the low levels of UPO1 secretion in microtiter plates, only the ABTS oxidation assay was reliable and stable, with a good signal response and a low level of interference in culture broth. This assay was adjusted for the appropriate substrate concentrations (0.3 and 2 mM of ABTS and H_2O_2, respectively) and the optimum pH (4.4). After improving secretion levels (from the second round of evolution onwards), a peroxygenase (oxygen transfer) assay using NBD as substrate was also incorporated into the screening protocol to maintain or even improve mono(per)oxygenase activity. UPO converts NBD into 4-nitrocatechol (yellow) via an initial hydroxylation and the subsequent spontaneous release of formic acid. 4-nitrocatechol can be deprotonated at basic pH values to produce a strong red color (34). Microfermentation conditions were optimized to minimize interferences during screening, evaluating several heme sources (δ-aminolevulinic acid, hemine, hemoglobin), the concentration of $MgSO_4$ (a source of structural Mg^{2+}) and the effect of ethanol on membrane permeability, as well as a range of temperatures, stirring rates and periods of incubation (see Supplemental Material and Methods for details). The heme source chosen for UPO expression (hemoglobin) generated unwanted background activity during screening and it was removed from the expression medium in the last rounds of evolution, as secretion was sufficiently high in its absence. Moreover, the coefficient of variance of the assays was reduced by up to 12% in the final cycles of

evolution due to the stronger activity displayed by the variants. Three consecutive re-screenings were performed to rule out the presence of false positives. To protect UPO stability during the course of the evolution, the T_{50} (temperature at which the enzyme retains 50% of its initial activity after 10 min of incubation) was estimated for each mutant studied during the third re-screening.

Directed Evolution

The total activity of UPO1 was improved ~3,250-fold with respect to the parental type after 5 generations (9,000 clones screened) of directed evolution. Moreover, expression levels of ~8 mg/L were achieved with activity values of 6,500 ABTS-U/L and 1,300 NBD-U/L. A combination of several error-prone PCR strategies with *in vivo* DNA recombination protocols, focused domain mutagenesis on the signal peptide and mutational recovery was performed. In the first generation, both the n-UPO1 and α-UPO1 fusion genes were used as starting points to enhance expression by subjecting the complete constructs to random mutagenesis and recombination. Accordingly, 4 mutant libraries were designed using different DNA polymerases and mutational loads, 2 for each construct. Mutants selected after several consecutive re-screenings were derived exclusively from n-UPO1 libraries. The 2 best variants from this round were 1A11 (L67F) and 3C2 (I248V-F311L) with 13- and 9-fold improvements in total activity compared to n-UPO1, respectively (**Figure 1, Table S1**).

The mutations of 1A11 and 3C2 were sufficiently distant (L67F of 1A11 was located at a distance of 181 residues from I248V in the 3C2 mutant) to permit a suitable crossover event to take place in the next cycle of evolution via the yeast *in vivo* recombination machinery. Accordingly, in the second generation these 2 variants were subjected to random mutagenesis and *in vivo* DNA shuffling. As planned, all variants selected from this round combined

L67F-I248V-F311L and they included some extra point mutations in either the mature protein or the signal peptide. The mutant best secreted in this generation (12C12) contained the aforementioned L67F-I248V-F311L mutations as well as A[21]D in the signal peptide. At this stage of evolution, the NBD assay could be incorporated to the screening protocol to measure oxygen transfer activity, thanks to the high levels of secretion observed for the 12C12 mutant. The use of the dual-screening assay based on the NBD:ABTS ratio allowed mutant hits to be selected without jeopardizing activity and variants with improved NBD:ABTS ratios to be identified (see below).

In the third generation, a dual approach was taken. First, mutant libraries with different mutational spectra were assembled *in vivo* by IvAM (27), using the 12C12 mutant (including its signal peptide) as a template (**Figure 1a**). In addition, the signal peptide of 12C12 was subjected to focal mutagenesis by MORPHING (28) in an attempt to enrich the signal peptide in mutations that favor secretion (**Figure 1b**). Libraries landscapes revealed a higher tolerance to mutations in the leader than in the whole UPO1 gene. This is consistent with the observation that mutations in the leader only affect secretion whereas mutations at the whole gene level may also modify catalytic properties. The most promising mutant from the IvAM library, I13D3, contained the new V75I mutation, and it displayed a 70-fold total activity improvement with respect to n-UPO1. By focusing mutational loads in the signal peptide, 3 beneficial mutations were introduced between positions 12 and 15, in 3 independent mutant winners. Mutations F[12]Y, A[14]V and R[15]G in the leader were so close to one other that the likelihood of recombination by *in vitro* or *in vivo* methods was very low. Thus, for the fourth generation we constructed a triple mutant by site-directed mutagenesis using the I13D3 mutant as a template (**Figure 1c**). The resulting mutant (2A12)

contained 4 beneficial mutations F[12]Y-A[14]V-R[15]G-A[21]D in the signal sequence, and it showed markedly improved secretion. Besides, a new round of random mutagenesis and recombination was also performed using I13D3, M5D2 and M4D8 as templates. Although M4D8 was not the best variant in generation 3, it was chosen as the parent due to the mutational redundancy observed at position 4, as this substitution also appeared in the 3B5 mutant from generation 2. From this set of experiments, we identified the 22A10 mutant, generated by a recombination event between M5D2 and I13D3, and containing the new mutation V57A (**Figure 1d**). Finally, the V57A mutation was introduced into 2A12 by mutational recovery, giving rise to the final mutant, PaDa-I.

Biochemical Characterization

Wild-type UPO1 produced homologously by *A. aegerita* ($_{wt}$UPO1) and the PaDa-I mutant secreted in *S. cerevisiae* were purified to homogeneity (Reinheitszahl value (*Rz* (A_{418}/A_{280}) ~ 2) and characterized biochemically (**Table 1, Figure S1**). The average molecular mass measured by MALDI-TOF-MS was 51,100 Da for PaDa-I (*i.e.*, ~5,000 Da higher than that of $_{wt}$UPO1), and the contribution of glycosylation deduced by deglycosylation gels was around 22% for the wild-type and 30% for the mutant (**Figure S1**). *S. cerevisiae* tends to hyperglycosylate foreign proteins up to levels of ~50%, conferring them with increased stability and protection against proteolytic degradation. The glycosylation of $_{wt}$UPO1 observed was exclusively dependent on 6 predicted *N*-glycosylation sites (*O*-glycosylation sites are not described for this enzyme), associated with up to 8 moieties of the high mannose type (30, 35). None of the amino acid substitutions in PaDa-I introduced new glycosylation motifs and thus, the higher sugar content in this mutant may be due to an increased Golgi residence time that leads to the addition of more mannose moieties, as

described in other directed evolution studies in yeast (14, 17). The PaDa-I and $_{wt}$UPO1 proteins had similar spectroscopic characteristics in the Fe^{3+} resting state, both enzymes showing a maximum in the Soret region of around 418 nm, and 2 Q-bands at 570 and 540 nm (**Figure 2A**).

The very weak expression of the parental n-UPO1 in *S. cerevisiae* (~0.007 mg/L) hampered its purification to homogeneity. To overcome this impediment, we constructed a fusion gene in which the native signal peptide of n-UPO1 was replaced by that obtained after 5 cycles of directed evolution in yeast (n*, containing F12Y-A14V-R15G-A21D mutations). Larger amounts of native UPO1 were produced in *S. cerevisiae* from the n*-UPO1 fusion gene, which was then purified to homogeneity. This approach allowed us to make an accurate breakdown of the total activity improvement (TAI) in terms of both specific activity and heterologous functional expression (**Figure 3**). The product of the n*-UPO1 fusion gene showed similar biochemical characteristics to the PaDa-I mutant in terms of molecular mass, the degree of glycosylation and thermal stability. After large scale fermentation, the PaDa-I protein showed a 3,250-fold TAI compared to the parental n-UPO1 (6,500 ABTS-U/L for PaDa-I *vs* 2 ABTS-U/L for n-UPO1), an improvement that was 20-fold higher than that obtained in microtiter plates where stirring conditions and oxygen availability were limiting. The product of n*-UPO1 showed a 27-fold TAI with respect to n-UPO1, reflecting the potency of the evolved signal peptide in promoting UPO1 secretion. The breakdown of the TAI value revealed a 1,114-fold increase in functional expression and a 3.6-fold increase in peroxidative activity (using ABTS) with respect to the parental type. Bearing in mind that n* enhanced functional expression 27-fold, there was a further ~41-fold improvement in functional expression conferred by mutations in the mature UPO1. Secretion levels were significantly enhanced from 0.007 mg/L

401 to 8 mg/L, *i.e.,* to levels similar to those obtained in the original fungus *A.*
402 *aegerita* (31).

403 The pH profiles for peroxidative (with ABTS, DMP) and peroxygenase
404 (with NBD) activities revealed similar shapes and optimum pH values for
405 $_{wt}$UPO1, n*-UPO1 and PaDaI (~4.0, 5.0 and 6.0 for ABTS, DMP and NBD,
406 respectively: **Figure 2C, 2D, and Table 1**). Kinetic constants for peroxidative
407 and peroxygenase activities were assayed using several compounds, along with
408 aryl alcohols with different redox potentials (**Table 2**). The k_{cat} for n*-UPO1
409 expressed in yeast was ~5-fold lower than that of $_{wt}$UPO1, although both
410 showed similar substrate affinities, with the exception of the K_m for H_2O_2 (2-
411 fold lower in n*-UPO1). General differences in protein folding in the
412 heterologous host, particularly those affecting post-translational modifications
413 (*e.g.*, hyperglycosylation), may affect the activity of the recombinant enzyme.
414 These results are consistent with those reported for lignin-modifying enzymes
415 heterologously expressed in yeast, in which a decrease in the k_{cat} was observed
416 of up to 10-fold (36). Both PaDa-I and $_{wt}$UPO1 showed very similar catalytic
417 efficiencies for all the substrates tested, indicating marked improvements in
418 the activity of UPO1 after each round of evolution. Indeed, a notable
419 improvement in the k_{cat} value was detected when n*-UPO1 was compared with
420 PaDa-I (with increases of 10- 9-, 8- and 7-fold for NBD, ABTS, veratryl alcohol
421 and benzyl alcohol, respectively), and these increases accounted for an overall
422 increase in catalytic efficiency: the 18-fold improvement in the k_{cat}/K_m for NBD
423 hydroxylation when comparing n*-UPO1 and PaDa-I was 2.2-times greater
424 than that of $_{wt}$UPO1 (*i.e.*, 700 and 320 $mM^{-1}s^{-1}$ for PaDa-I and $_{wt}$UPO1,
425 respectively). Thus, the dual screening assay used during *in vitro* evolution
426 helped to conserve and even improve both peroxidative and
427 mono(per)oxygenase activities.

Many of the aromatic substrates and oxyfunctionalized compounds converted by UPO are further oxidized by the enzyme, generating a pool of products of varying complexities. Accordingly, for certain applications the removal of UPO's peroxidative activity (1-electron oxidation) may be considered a biotechnological priority. Using the HTS-assay we attempted to uncouple the peroxidation and monooxygenase activities. Taking the NBD:ABTS ratio as a discriminatory factor, the 3F10 variant (T120P mutation) identified in the fourth generation showed a dramatic decrease in peroxidative activity (a 4-fold decrease in the TAI for ABTS, from 132- to 31-fold), while conserving its peroxygenase activity (with a similar TAI for NBD as the parental 2A12, **Figure 1**). The signal peptide was switched in the 3F10 variant, as described for the n*-UPO1 construct (see Supplemental Material and Methods), and the corresponding n*-3F10 construct was produced on a larger scale and compared with both $_{wt}$UPO1 and n*-UPO1. The NBD:ABTS ratios (expressed as percentage) were 19%, 20% and 61% for $_{wt}$UPO1, n*-UPO1 and n*3F10, respectively. During the preliminary characterization of this variant, we detected a dramatic decrease in the thermostability produced by the beneficial but destabilizing T120P mutation (with a 7ºC-decrease in T_{50}), which precluded its purification and further analysis. Although monooxygenase activity is a clear target for directed UPO evolution, our results indicate that the line between peroxidation and mono(per)oxygenase activity in UPO is very fine, and that its catalytic mechanisms are strongly implicated in protein stability (10). This problem may be overcome by including iterative rounds of neutral genetic drift in order to introduce stabilizing mutations into the 3F10 mutant before further evolving its monooxygenase activity (37).

Kinetic thermostability was conserved over the course of evolution, with T_{50} values in the 57-59°C range for all the offspring of the mutants used as

parental types in each round of directed evolution (**Figure 1, 2B**). This effect was due to the screening assay used during evolution, in which destabilizing mutations were excluded from the evolutionary pathway.

The presence of organic co-solvents is required for many of the transformations mediated by UPO. The activity and stability of $_{wt}$UPO and the PaDa-I mutant were evaluated in the presence of high concentrations of co-solvents with different polarities (with a logP ranging from -0.23 to -1.3) and chemical characteristics (**Figure 4A, B**). Regardless of the enzyme tested, activity was reduced drastically in the presence of increasing concentrations of co-solvents in the following order: ethanol>DMSO>ACN~methanol>acetone. Activity in co-solvents were estimated by measuring the C_{50}, the concentration of co-solvent at which the enzyme shows 50% of the corresponding activity in buffer. The strongest activity was observed in acetone (C_{50} values of 10-12%) and the weakest in ethanol and DMSO (C_{50} ~2%). In terms of stability in co-solvents, both enzymes were very stable at concentrations as high as 50% (v/v), with a half-life of over 48 h and some hyper-activation (**Figure 4C, D**). Under more extreme conditions (concentrations ranging from 60-90%), co-solvents exerted detrimental effects in the following order: DMSO>ethanol>methanol>acetone~ACN. Interestingly, PaDa-I was very stable, retaining ~25% and ~55% of its activity at concentrations of up to 90% (v/v) methanol and ethanol, respectively, when the stability of $_{wt}$UPO1 was negligible (**Figure 4E, F**). The combined effect of hyperglycosylation together with the introduction of stabilizing mutations appears to underlie this resistance.

Mutation Analysis

PaDa-I harbored 9 beneficial mutations (no silent mutations were introduced during evolution), 4 in the signal peptide and 5 in the mature

protein. Five of the mutations (3 in mature protein and 2 in the signal peptide) favored codon usage, which might support secretion (**Table S1**). The 4 mutations in the signal peptide were located in the hydrophobic core of the leader and 3 of these were nearly consecutive (underlined): F[12]Y-A[14]F-R[15]G-A[21]D. Substitutions at positions 12 and 21 enhanced the polarity of this region, while those at positions 14 and 15 had the opposite effect. Overall, these 4 mutations enhanced secretion by up to 27-fold, as seen in the n*-UPO1 fusion gene, **Figure 3**. Subtle differences in the adjustment between the signal recognition particle (SRP) and the evolved signal peptide may benefit secretion, bearing in mind that SRP strongly interacts with the hydrophobic region of the leader. It has been reported that SRP binding to the signal peptide pauses translation at different stages, depending on the nature of the leader (38). In our mutant, this arrest of translation could facilitate proper recognition by the signal peptidase before cleavage and translocation of the nascent UPO polypeptide to the endoplasmic reticulum (39).

Mutations in the mature UPO1 were mapped onto the recently solved crystal structure of UPO1 (35). UPO1 is mostly formed by helical substructures composed of 1 halide binding site, 1 Mg^{2+} binding site and the heme-thiolate domain, with Arg189 and Glu196 forming the acid-base pair for Compound I formation. The funnel-shaped access channel to the substrate binding pocket is 8.5 Å in diameter and mainly covered with aromatic residues. This binding pocket is controlled by a Phe triad (Phe69, Phe121 and Phe199), which is essential for the orientation of (aromatic) substrates. All mutations were conservative in terms of polarity and charge, *i.e.,* non-polar substitutions (V57A, L67F, V75I, I248V and F311L), leading to few apparent changes in terms of H-bond or salt-bridge formation/interruption (**Table S2)**. In fact, mutations were located in hydrophobic environments, in some cases

far from the catalytic site (**Figure S2**). At this point, it is important to note that our aim was to improve total activity while conserving enzyme stability. Thus, we can only speculate as to whether a less rigid directed evolution approach would have unmasked other, less conservative substitutions, at the cost of threatening protein stability (40). The 41-fold enhancement in secretion induced by these mutations may be due to tighter folding during the earliest post-translational stages, which ultimately favors protein stability and secretion. The V57A mutation lies in a helix at the surface of the protein next to the N-terminus (**Figure 5A, B**). According to our model, the replacement of Val57 with a less bulky residue may compress this region between adjacent prolines (at positions 5 and 6 of the N-terminus) and thereby increasing the protein's robustness. The L67F mutation is located in the vicinity of the catalytic pocket, very close to Phe69 of the Phe triad involved in binding aromatic substrates. This mutation may be partially responsible for the kinetic enhancements observed, since after substitution, the aromatic ring of Phe67 is orientated towards the active site (**Figure 5C, D**). The V75I mutation is produced by the substitution of a hydrophobic residue with another slightly larger hydrophobic residue, which may establish new hydrophobic contacts with surrounding residues, thereby favoring protein stability (**Figure 5A, B**). Finally, the I248V and F311L mutations are positioned in the surroundings of the heme channel. The I248V mutation is produced by the replacement of Ile248 at the entrance of the channel to the heme cavity with a Val residue, which could favor the access of bulkier substrates (**Figure 5E, F**). The F311L mutation is located in front of the heme channel with Phe76 in between. The substitution of Phe with Leu enlarges the cavity, which may in turn have beneficial effects on kinetics (**Figure 5A, B**).

Conclusions

Since its discovery 10 years ago, the potential use of UPO in applications ranging from chemical processes (including some relevant industrial transformations such as alkane hydroxylations and olefin epoxidations) to the preparation of *O*- and *N*-dealkylated human drug metabolites, as well as bioremediation (PAH oxidation) and biosensor development, has been studied exhaustively (41-50). For decades, regio- and enantioselective oxyfunctionalization has been a "forbidden territory" for most biocatalysts, except for P450 monooxygenases. However, unlike the latter, UPO is soluble and requires neither expensive cofactors (NAD[P]H) nor auxiliary flavoproteins. Despite these advantages, the lack of suitable directed evolution platforms with which to enhance UPO´s catalytic properties has limited the exploitation of this versatile biocatalyst. The directed evolution process presented here describes for the first time an attractive pathway through which *ad-hoc* UPO variants can be tailored for use in several industrial reactions, such as alkanes hydroxylation, and the transformation of benzene into phenol and naphthalene into naphthol (51, 52).

The evolved UPO1 variant of this study is very active and stable over a wide temperature range, as well as in the presence of a variety of co-solvent types. Easily secreted by yeast, this mutant and any future evolved variants, could be translated to other expression systems for overproduction. Preliminary trials in the methylotrophic yeast *Pichia pastoris* indicate that the evolved enzyme is overproduced 5-fold thanks to the increased density in this host (data not published). Future goals for UPO engineering will include the conversion of the enzyme into an enantioselective self-sufficient mono(per)oxygenase by quenching its peroxidative activity, the improvement of

562 activity in the presence of co-solvents and the enhancement of its
563 oxidative/operational stability in the presence of peroxides. The future
564 combination of directed evolution (including neutral genetic drift) and
565 rational/hybrid design should provide a wealth of information that will help us
566 to better understand the complex mechanism of action as UPO becomes an
567 efficient oxyfunctionalization biocatalyst.

568 **SUPPLEMENTAL INFORMATION**

569 Supplemental information includes 3 Tables, 2 Figures and the Supplemental
570 Material and Methods section.

571 **ACKNOWLEDGEMENTS**

572 This work was supported by European Commission Projects (Peroxicats-FP7-
573 KBBE-2010-4-26537; Indox-FP7-KBBE-2013-7-613549) and the National
574 Project (Evofacel) [BIO2010-19697].

575

576 **REFERENCES**

577 1. **Hofrichter M, Ullrich R**. 2006. Heme-thiolate haloperoxidases: versatile
578 biocatalysts with biotechnological and environmental significance. Appl.
579 Microbiol. Biot. **71**: 276-288.
580 2. **Hofrichter M, Ullrich R, Pecyna MJ, Liers C, Lundell T**. 2010. New
581 and classic families of secreted fungal heme peroxidases. Appl. Microbiol.
582 Biot. **87**: 871-897.
583 3. **Anh DH, Ullrich R, Benndorf D, Svatos A, Muck A, Hofrichter M**.
584 2007. The coprophilous mushroom *Coprinus radians* secretes a
585 haloperoxidase that catalyzes aromatic peroxygenation: Appl. Environ.
586 Microbiol. **73**: 5477-5485.

4. **Gröbe G, Ullrich R, Pecyna MJ, Kapturska D, Friedrich S, Hofrichter M, Scheibner K**. 2011. High-yield production of aromatic peroxygenase by the agaricus fungus *Marasmius rotula*. AMB Express **1**: 31.

5. **Babot ED, del Rio JC, Kalum L, Martinez AT, Gutierrez A**. 2013. Oxyfunctionalization of aliphatic compounds by a recombinant peroxygenase from *Coprinopsis cinerea*: Biotechnol. Bioeng. **110**: 2323-2332.

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

7. **Hofrichter M, Kellner H, Pecyna M, Ullrich R**. 2014. Fungal unspecific peroxygenases: heme-thiolate proteins that combine peroxidase and cytochrome P450 properties. *In*: E. G. Hrycay ed., Monooxygenase, peroxidase and peroxygenase properties and mechanisms of cytochrome P450. Springer, Heidelberg, in press

8. **Pazmino DET, Winkler M, Glieder A, Fraaije MW**. 2010. Monooxygenases as biocatalysts: classification, mechanistic aspects and biotechnological applications. J. Biotechnol. **146**: 9-24.

9. **Wang X, Peter S, Kinne M, Hofrichter M, Groves JT**. 2012. Detection and Kinetic Characterization of a Highly Reactive Heme-Thiolate Peroxygenase Compound I. J. Am. Chem. Soc. **134**: 12897-12900.

10. **Wang X, Peter S, Ullrich R, Hofrichter M, Groves JT**. 2013. Driving Force for Oxygen-Atom Transfer by Heme-Thiolate Enzymes: Angew. Chem. Int. Edit. **52**: 9238-9241.

11. **Conesa A, van de Velde F, van Rantwijk F, Sheldon RA, van den Hondel CAMJ, Punt PJ**. 2001. Expression of the *Caldariomyces fumago* chloroperoxidase in *Aspergillus niger* and characterization of the recombinant enzyme. J. Biol. Chem. **276***:* 17635-17640.

12. **Shaw P, Beckwith J**. 1960. Chloroperoxidase - A Component of the Beta-Ketoadipate Chlorinase System. Federation Proceedings **19**: 47.

13. **Pourmir A, Johannes TW**. 2012 Directed evolution: selection of the host microorganism. Comput. Struct. Biotechnol. J. **2***:* e201209012.

14. **Bulter T, Alcalde M, Sieber V, Meinhold P, Schlachtbauer C, Arnold FH**. 2003. Functional expression of a fungal laccase in *Saccharomyces cerevisiae* by directed evolution. Appl. Environ. Microbiol. **69***:* 987-995.

15. **Garcia-Ruiz E, Gonzalez-Perez D, Ruiz-Dueñas FJ, Martinez AT, Alcalde M**. 2012. Directed evolution of a temperature-, peroxide- and alkaline pH-tolerant versatile peroxidase. Biochem. J. **441**: 487-498.

16. **Zumarraga M, Bulter T, Shleev S, Polaina J, Martinez-Arias A, Plou FJ, Ballesteros A, Alcalde M**. 2007. In vitro evolution of a fungal laccase in high concentrations of organic cosolvents. Chem. Biol. **14**: 1052-1064.

17. **Camarero S, Pardo I, Cañas A, Molina P, Record E, Martinez AT, Martinez MJ, Alcalde M**. 2012. Engineering Platforms for Directed Evolution of Laccase from *Pycnoporus cinnabarinus*: Appl. Environ. Microbiol. **78**: 1370-1384.

18. **Mate D, Garcia-Burgos C, Garcia-Ruiz E, Ballesteros AO, Camarero S, Alcalde M**. 2010. Laboratory Evolution of High-Redox Potential Laccases. Chem. Biol. **17**: 1030-1041.

19. **Mate DM, Gonzalez-Perez D, Falk M, Kittl R, Pita M, De Lacey AL, Ludwig R, Shleev S, Alcalde M**. 2013. Blood Tolerant Laccase by Directed Evolution. Chem. Biol. **20**: 223-231.

20. **Pardo I, Vicente AI, Mate DM, Alcalde M, Camarero S**. 2012. Development of chimeric laccases by directed evolution. Biotechnol. Bioeng. **109**: 2978-2986.

21. **Gonzalez-Perez D, Garcia-Ruiz E, Alcalde M**. 2012. *Saccharomyces cerevisiae* in directed evolution: An efficient tool to improve enzymes. Bioengineered **3**: 172-177.

22. **Pirakitikulr N, Ostrov N, Peralta-Yahya P, Cornish VW**. 2010. PCRless library mutagenesis via oligonucleotide recombination in yeast. Protein Sci. **19**: 2336-2346.

23. **Shao ZY, Zhao H, and Zhao HM**. 2009. DNA assembler, an *in vivo* genetic method for rapid construction of biochemical pathways. Nucleic Acid Res. **37***:* 2.

24. **Da Silva NA, Srikrishnan S**. 2012. Introduction and expression of genes for metabolic engineering applications in *Saccharomyces cerevisiae*. Fems Yeast Res. **12**: 197-214.

664 25. **Hong KK, Nielsen J**. 2012. Metabolic engineering of *Saccharomyces*
665 *cerevisiae*: a key cell factory platform for future biorefineries. Cell. Mol. Life
666 Sci. **69**: 2671-2690.
667 26. **Krivoruchko A, Siewers V, Nielsen J**. 2011. Opportunities for yeast
668 metabolic engineering: Lessons from synthetic biology. Biotechnol. J. **6**:
669 262-276.
670 27. **Zumarraga M, Camarero S, Shleev S, Martínez-Arias A, Ballesteros**
671 **A, Plou FJ, Alcalde M**. 2008. Altering the laccase functionality by *in vivo*
672 assembly of mutant libraries with different mutational spectra. Proteins **71**:
673 250-260.
674 28. **Gonzalez-Perez D, Molina-Espeja P, Garcia-Ruiz E, Alcalde M**. 2014.
675 Mutagenic organized recombination process by homologous in vivo grouping
676 (MORPHING) for directed enzyme evolution. Plos One **9**, 3: e90919.
677 29. **Alcalde M**. 2010. Mutagenesis protocols in *Saccharomyces cerevisiae* by
678 *in vivo* overlap extension, p 3-15 . *In*: J. Bramman (ed.), *In vitro* mutagenesis
679 protocols 3rd ed. Methods in Molecular Biology 634. Humana Press, Totowa,
680 New Jersey.
681 30. **Pecyna MJ, Ullrich R, Bittner B, Clemens A, Scheibner K, Schubert**
682 **R, Hofrichter M**. 2009. Molecular characterization of aromatic
683 peroxygenase from *Agrocybe aegerita*. Appl. Microbiol. Biot. **84**: 885-897.
684 31. **Ullrich R, Nüske J, Scheibner K, Spantzel J, Hofrichter M**. 2004.
685 Novel haloperoxidase from the agaric basidiomycete *Agrocybe aegerita*
686 oxidizes aryl alcohols and aldehydes. Appl. Environ. Microb. **70**: 4575-
687 4581.
688 32. **Mate D, Garcia-Ruiz E, Camarero S, Alcalde M**. 2011. Directed
689 Evolution of Fungal Laccases. Curr. Genomics. **12**: 113-122.

690 33. **Rakestraw J, Sazinsky SL, Piatesi A, Antipov E, Wittrup K**. 2009.
691 Directed Evolution of a Secretory Leader for the Improved Expression of
692 Heterologous Proteins and Full-Length Antibodies in *Saccharomyces*
693 *cerevisiae.* Biotechnol. Bioeng. **103**: 1192-1201.
694 34. **Poraj-Kobielska M, Kinne M, Ullrich R, Scheibner K, Hofrichter M**.
695 2012. A spectrophotometric assay for the detection of fungal peroxygenases.
696 Anal. Biochem., **421**: 327-329.
697 35. **Piontek K, Strittmatter E, Ullrich R, Gröbe G, Pecyna MJ, Kluge M,**
698 **Scheibner K, Hofrichter M, Plattner DA**. 2013. Structural basis of
699 substrate conversion in a new aromatic peroxygenase: P450 functionality
700 with benefits. J. Biol. Chem. **288**: 34767-3476.
701 36. **Mate DM, Garcia-Ruiz E, Camarero S, Shubin VV, Falk M, Shleev S,**
702 **Ballesteros AO, Alcalde M**. 2013. Switching from blue to yellow: altering
703 the spectral properties of a high redox potential laccase by directed
704 evolution. Biocatal. Biotransfor. **31**: 8-21.
705 37. **Goldsmith M, Tawfik, DS**. 2013. Enzyme Engineering by Targeted
706 Libraries. Method Enzymol. 523: 257-283.
707 38. **Nothwehr SF, Gordon JI**. 1990. Targeting of proteins into the
708 eukaryotic secretory pathway: signal peptide structure/function
709 relationships. Bioessays **12**: 479-484.
710 39. **Romanos MA, Scorer CA, Clare JJ**. 1992. Foreign gene expression in
711 yeast: a review. Yeast **8**: 423-488.
712 40. **Bloom JD, Labthavikul ST, Otey CR, Arnold FH**. 2006. Protein
713 stability promotes evolvability. Proc. Natl. Acad. Sci. USA. **103**: 5869-5874.
714 41. **Aranda E, Ullrich R, Hofrichter M**. 2010. Conversion of polycyclic
715 aromatic hydrocarbons, methyl naphthalenes and dibenzofuran by two
716 fungal peroxygenases: Biodegradation **21**: 267-281.

42. **Barkova K, Kinne M, Ullrich R, Hennig L, Fuchs A, Hofrichter M**. 2011. Regioselective hydroxylation of diverse flavonoids by an aromatic peroxygenase. Tetrahedron **67**: 4874-4878.

43. **Gutierrez A, Babot ED, Ullrich R, Hofrichter M, Martinez AT, del Rio JC**. 2011. Regioselective oxygenation of fatty acids, fatty alcohols and other aliphatic compounds by a basidiomycete heme-thiolate peroxidase. Arch. Biochem. Biophys. 514: 33-43.

44. **Kinne M, Zeisig C, Ullrich R, Kayser G, Hammel KE, Hofrichter M**. 2010. Stepwise oxygenations of toluene and 4-nitrotoluene by a fungal peroxygenase. Biochem. Biophys. Res. Co. **397**: 18-21.

45. **Kinne M, Poraj-Kobielska M, Aranda E, Ullrich R, Hammel KE, Scheibner K, Hofrichter M**. 2009. Regioselective preparation of 5-hydroxypropranolol and 4 '-hydroxydiclofenac with a fungal peroxygenase. Bioorg. Med. Chem. Lett. **19**: 3085-3087.

46. **Kluge M, Ullrich R, Scheibner K, Hofrichter M**. 2012. Stereoselective benzylic hydroxylation of alkylbenzenes and epoxidation of styrene derivatives catalyzed by the peroxygenase of *Agrocybe aegerita*. Green Chem. **14**: 440-446.

47. **Peng, L, Wollenberger U, Kinne M, Hofrichter M, Ullrich R, Scheibner K, Fischer A, Scheller FW.** 2010. Peroxygenase based sensor for aromatic compounds. Biosens. Bioelectron. 26: 1432-1436.

48. **Peter S, Kinne M, Ullrich R, Kayser G, Hofrichter M**. 2013. Epoxidation of linear, branched and cyclic alkenes catalyzed by unspecific peroxygenase. Enzyme Microb. Technol. **52**: 370-376.

49. **Poraj-Kobielska M, Atzrodt J, Holla W, Sandvoss M, Groebe G, Scheibner K, Hofrichter M**. 2013. Preparation of labeled human drug

743 metabolites and drug-drug interaction-probes with fungal peroxygenases.
744 Journal of Labelled Compd Rad. **56**: 513-519.
745 50. **Poraj-Kobielska M, Kinne M, Ullrich R, Scheibner K, Kayser G,**
746 **Hammel KE, Hofrichter M**. 2011. Preparation of human drug metabolites
747 using fungal peroxygenases. Biochem. Pharmacol. **82**: 789-796.
748 51. **Karich A, Kluge M, Ullrich R, Hofrichter M**. 2013. Benzene
749 oxygenation and oxidation by the peroxygenase of *Agrocybe aegerita.* AMB
750 Express **3**, 5.
751 52. **Kluge M, Ullrich R, Dolge C, Scheibner K, Hofrichter M**. 2009.
752 Hydroxylation of naphthalene by aromatic peroxygenase from *Agrocybe*
753 *aegerita* proceeds via oxygen transfer from H_2O_2 and intermediary
754 epoxidation. Appl. Microb. Biotech. **81**: 1071-1076.
755
756
757
758
759
760
761
762
763
764
765
766

FIGURE LEGENDS

Figure 1. Route for the directed evolution of UPO1 towards functional expression and improved activity. New mutations are depicted as stars and accumulated mutations as squares. Mutations in the mature PaDa-I mutant and their origin are highlighted in boxes. The signal peptide is represented in dark red and the mature protein in red. In the parent n-UPO1, the glycosylation sites (Asn11, Asn141, Asn161, Asn182, Asn286, Asn295) are represented as blue stars, the Phe triad (Phe69, Phe121, Phe199) involved in the binding of aromatic substrates is marked with green arrows, and the acid-base pair for peroxide cleavage (Glu196 and Arg189) is indicated with blue arrows. TAI represents the improvement in UPO1 activity detected in *S. cerevisiae* microcultures for each mutant compared with the parental n-UPO1. Thermostability (T_{50}) was estimated from culture supernatants (see also **Figure 2B**). The breakdown of the TAI into specific activity and expression is shown in **Figure 3**. Dashed arrows indicate the parental types used for each round of evolution. In the 3rd generation, (**a**) indicates the offspring obtained by IvAM of 12C12, and (**b**) indicates the offspring obtained by MORPHING in the signal peptide of 12C12. In the 4th generation, (**c**) indicates the triple mutant at the signal peptide constructed using I13D3 as a template and (**d**) the offspring obtained by mutagenic PCR and shuffling of parents I13D3, M5D2 and M4D8. n.d., not determined; n.m., not measurable. See also **Table S1**.

Figure 2. Biochemical characterization. (**A**) Spectroscopic characteristics of $_{wt}$UPO1 (dashed line) and the PaDa-I mutant (solid line). (**B**) Thermostability (T_{50}) of PaDa-I and different parental types. Each point represents the mean and standard deviation of 3 independent experiments. (**C, D**) pH activity profiles for $_{wt}$UPO1 (white triangles) and PaDa-I (black squares). Activities were measured in 100 mM citrate/phosphate/borate buffer at different pH values

794 with 2 mM H_2O_2 and 0.3 mM ABTS (**C**) or with 1 mM H_2O_2 and 1 mM NBD (**D**).
795 UPO1 activity was normalized to the optimum activity value and each point
796 represents the mean and standard deviation of 3 independent experiments.

797 **Figure 3. Breakdown of specific activity and functional expression.** A
798 fusion gene containing the evolved signal peptide (n*) attached to the native
799 mature UPO1 was engineered. The n-UPO1, n*-UPO1 and PaDa-I variants
800 were produced on a large scale and their TAIs measured. The n*-UPO1 and
801 PaDa-I were purified and their specific activities calculated. The n* enhanced
802 functional expression ~27-fold, whereas mutations in mature PaDa-I resulted
803 in a ~120-fold increase in total activity. The 3,250-fold increase in the total
804 activity of PaDa-I was broken down as a 3.6-fold improvement in specific
805 activity and a 1,114-fold improvement in functional expression.

806 **Figure 4. Activity and stability in organic co-solvents**. The relative
807 activities in of $_{wt}$UPO1 (**A**) and the PaDa-I mutant (**B**) in organic co-solvents
808 were assessed with 2 mM H_2O_2 and 0.3 mM ABTS in 100 mM sodium
809 phosphate/citrate buffer [pH 4.4] containing the corresponding concentration
810 of co-solvent. (**C, D**) The stability of $_{wt}$UPO1 (**C**) and the PaDa-I mutant (**D**) after
811 incubation for 48 h in 50% organic co-solvents was assessed by incubating
812 enzyme samples in 100 mM potassium phosphate buffer [pH 7.0] containing
813 50% (v/v) organic co-solvent in screw-cap vials. After 48 h, aliquots were
814 removed and analysed in an activity assay with 2 mM H_2O_2 and 0.3 mM ABTS
815 in 100 mM sodium phosphate/citrate buffer [pH 4.4]. (**E, F**) The stabilities of
816 $_{wt}$UPO1 (**E**) and the PaDa-I mutant (**F**) at high concentrations of organic co-
817 solvents were assessed after a 5 h incubation in increasing concentrations of
818 co-solvents, incubating enzyme samples at 20ºC in 100 mM potassium
819 phosphate buffer [pH 7.0] containing increasing concentrations (v/v) of
820 organic co-solvent (60-90%). After 5 h, aliquots were removed and analysed in

821 the activity assay, as described above. Residual activities were expressed as a
822 percentage of the original activity at the corresponding concentration of
823 organic co-solvent.

824 **Figure 5. Mutations in evolved UPO1.** A molecular model using as template
825 the *A. aegerita* crystal structure (PDB code 2YOR) was prepared to map the
826 mutations. Details are shown of the 5 mutations (in green) in the PaDa-I
827 mutant (**B, D, F**) compared with the corresponding residues (in yellow) in the
828 native UPO1 (**A, C, E**). Dashed lines indicate distances (in Å) from the
829 surrounding residues. Phe residues delimiting the active site are highlighted in
830 pink, and the Cys36 axial ligand in light blue. The Fe^{3+} of heme is shown as a
831 red sphere and the structural Mg^{2+} as a pink sphere. See also **Table S2,**
832 **Figure S2**.

Table 1. Biochemical features of wild-type and evolved UPO variants.

Biochemical and Spectroscopy features	$_{wt}$UPO1	PaDa-I mutant
MW (Da)[1]	46,000	52,000
MW (Da)[2]	n.d.	51,100
MW (Da)[3]	35,942	35,914
Glycosylation degree (%)	22	30
Thermal stability, T_{50} (ºC)[4]	53	55
pI	4.9-5.7	5.5
Optimum pH for ABTS	4.0	4.0
Optimum pH for DMP	7.0	6.0
Optimum pH for NBD	6.5	6.0
R_Z, (A_{410}/A_{280})	2.4	1.8
Soret region (nm)	420	418
CT1 (nm)	572	570
CT2 (nm)	540	537

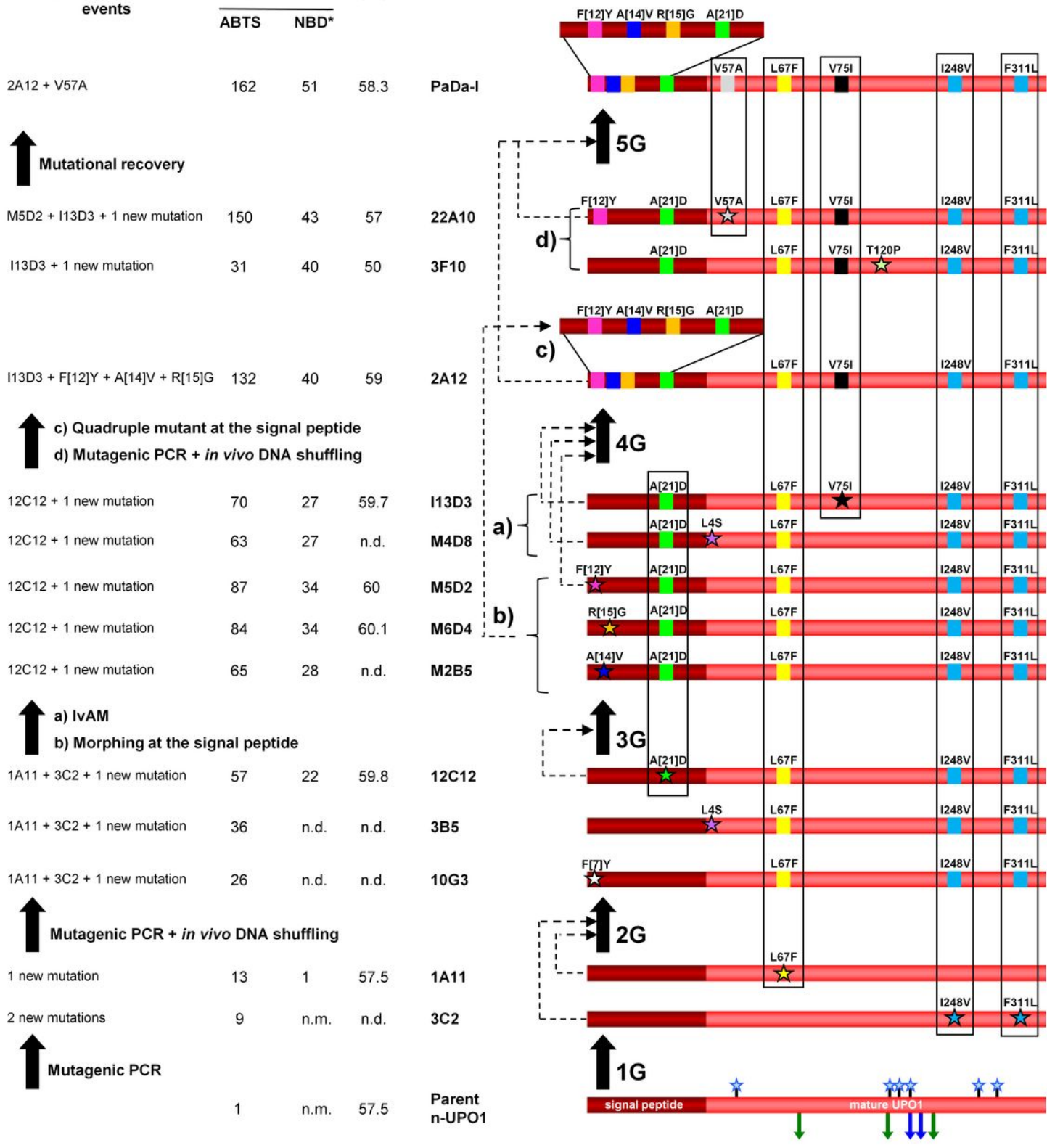
[1] Estimated by SDS-PAGE; [2]estimated by MALDI-TOF mass spectrometry; [3]estimated from amino acid composition. [4]Estimated from purified variants. n.d. not determined. $_{wt}$UPO1, UPO1 wild-type expressed in *A. aegerita*; PaDa-I mutant, ultimate variant of the whole evolution process *in S. cerevisiae* (containing the evolved signal peptide (n*) plus the evolved UPO1). See also **Fig. S1**.

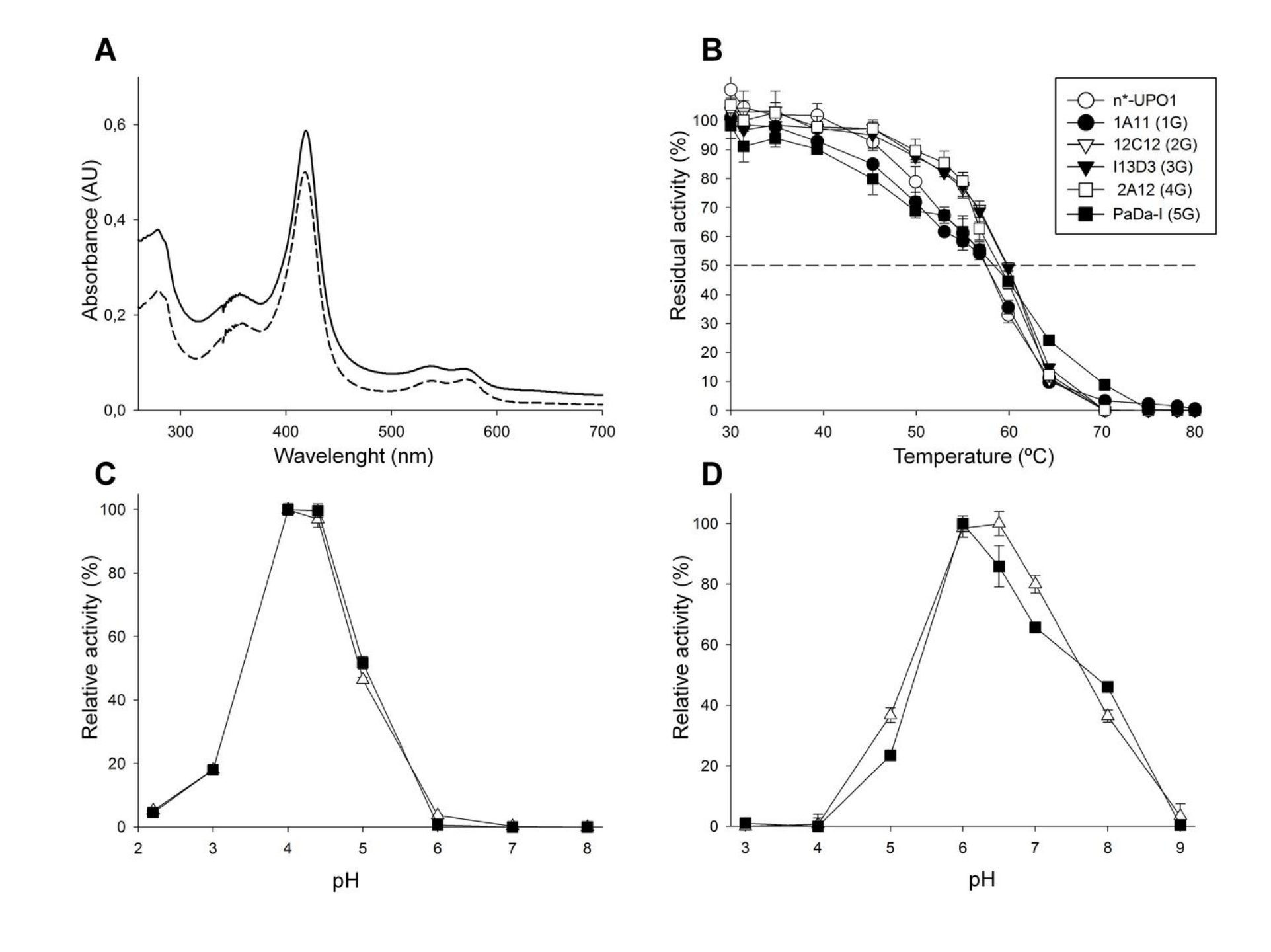
Table 2. Kinetic parameters of wild-type, recombinant and evolved UPO variants

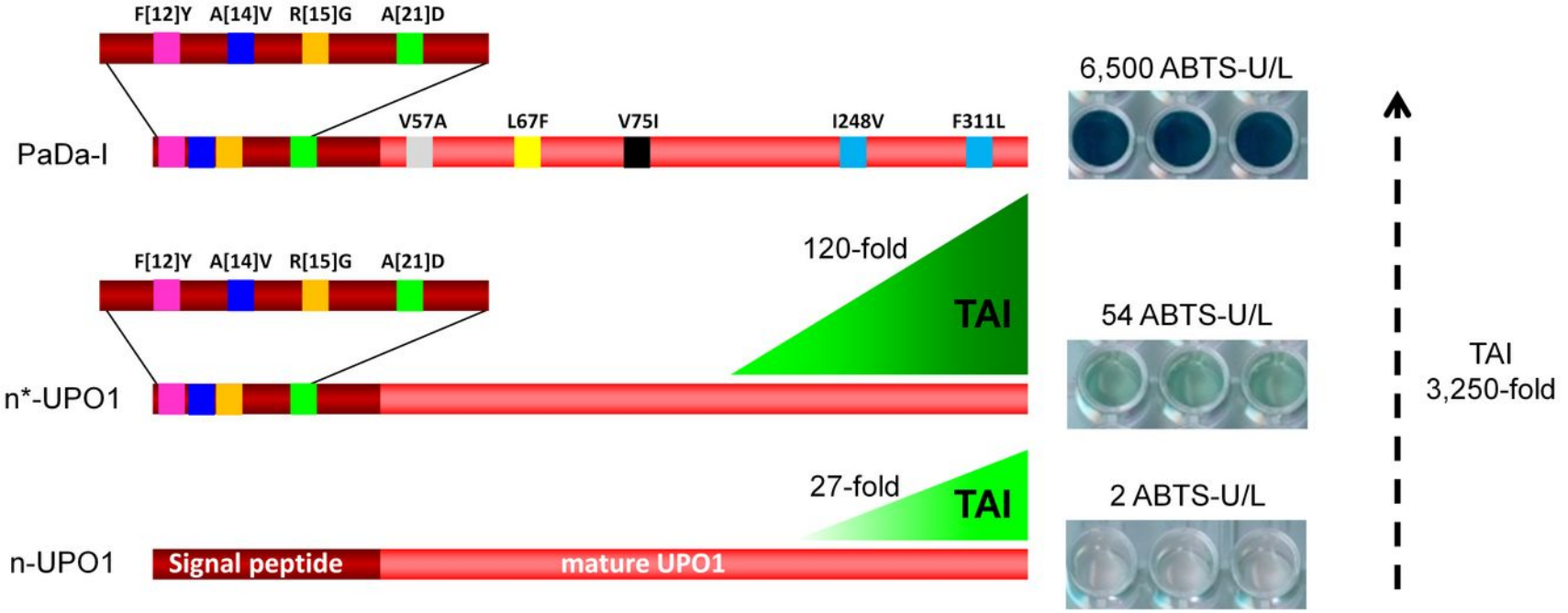
Substrate	Kinetics constants	$_{wt}$UPO1	n*-UPO1	PaDa-I
ABTS	K_m (mM)	0.025 ± 0.002	0.027 ± 0.005	0.048 ± 0.004
	k_{cat} (s^{-1})	221 ± 6	45.0 ± 2.7	395 ± 13
	k_{cat}/K_m (mM^{-1} s^{-1})	8,800 ± 692	1,600 ± 37	8,200 ± 598
NBD	K_m (mM)	0.684 ± 0.207	0.782 ± 0.352	0.483 ± 0.095
	k_{cat} (s^{-1})	219 ± 25	31.7 ± 6.1	338 ± 22
	k_{cat}/K_m (mM^{-1} s^{-1})	320 ± 64	38.0 ± 11	700 ± 99
Benzyl alcohol	K_m (mM)	1.90 ± 0.11	1.10 ± 0.23	2.47 ± 0.32
	k_{cat} (s^{-1})	329 ± 7	44.8 ± 3.1	307 ± 15
	k_{cat}/K_m (mM^{-1} s^{-1})	174 ± 7	41.0 ± 6.3	124 ± 11
Veratryl alcohol	K_m (mM)	5.20 ± 0.31	5.30 ± 0.82	7.9 ± 0.7
	k_{cat} (s^{-1})	88 ± 2	15.2 ± 1.1	121 ± 5
	k_{cat}/K_m (mM^{-1} s^{-1})	17 ± 0.7	2.9 ± 0.25	15 ± 0.9
H_2O_2	K_m (mM)	1.37 ± 0.16	0.69 ± 0.20	0.49 ± 0.06
	k_{cat} (s^{-1})	290 ± 15	40.9 ± 3.8	238 ± 8
	k_{cat}/K_m (mM^{-1} s^{-1})	211 ± 15	59.0 ± 12.3	500 ± 42

*ABTS kinetic constants for UPO1 were estimated in 100 mM sodium citrate/phosphate buffer pH 4.4 containing 2 mM H_2O_2; for the rest of the substrates in 100 mM potassium phosphate buffer pH 7.0 containing 2 mM H_2O_2 (benzyl and veratryl alcohols) or 1 mM H_2O_2 (NBD). H_2O_2 kinetic constants were estimated using benzyl alcohol as reducing substrate at the corresponding saturated conditions. $_{wt}$UPO1, UPO1 wild-type expressed in *A. aegerita*; n*-UPO, native UPO1 fused to the evolved signal peptide for secretion in *S. cerevisiae*; PaDa-I mutant, ultimate variant of the whole evolution process *in S. cerevisiae* (containing the evolved signal peptide plus the evolved UPO1).

DNA diversity strategies and mutagenic/recombination events	TAI (in fold)		T_{50} (°C)	MUTANT
	ABTS	NBD*		
2A12 + V57A	162	51	58.3	PaDa-I
Mutational recovery				
M5D2 + I13D3 + 1 new mutation	150	43	57	22A10
I13D3 + 1 new mutation	31	40	50	3F10
I13D3 + F[12]Y + A[14]V + R[15]G	132	40	59	2A12
c) Quadruple mutant at the signal peptide **d) Mutagenic PCR + *in vivo* DNA shuffling**				
12C12 + 1 new mutation	70	27	59.7	I13D3
12C12 + 1 new mutation	63	27	n.d.	M4D8
12C12 + 1 new mutation	87	34	60	M5D2
12C12 + 1 new mutation	84	34	60.1	M6D4
12C12 + 1 new mutation	65	28	n.d.	M2B5
a) IvAM **b) Morphing at the signal peptide**				
1A11 + 3C2 + 1 new mutation	57	22	59.8	12C12
1A11 + 3C2 + 1 new mutation	36	n.d.	n.d.	3B5
1A11 + 3C2 + 1 new mutation	26	n.d.	n.d.	10G3
Mutagenic PCR + *in vivo* DNA shuffling				
1 new mutation	13	1	57.5	1A11
2 new mutations	9	n.m.	n.d.	3C2
Mutagenic PCR				
	1	n.m.	57.5	Parent n-UPO1

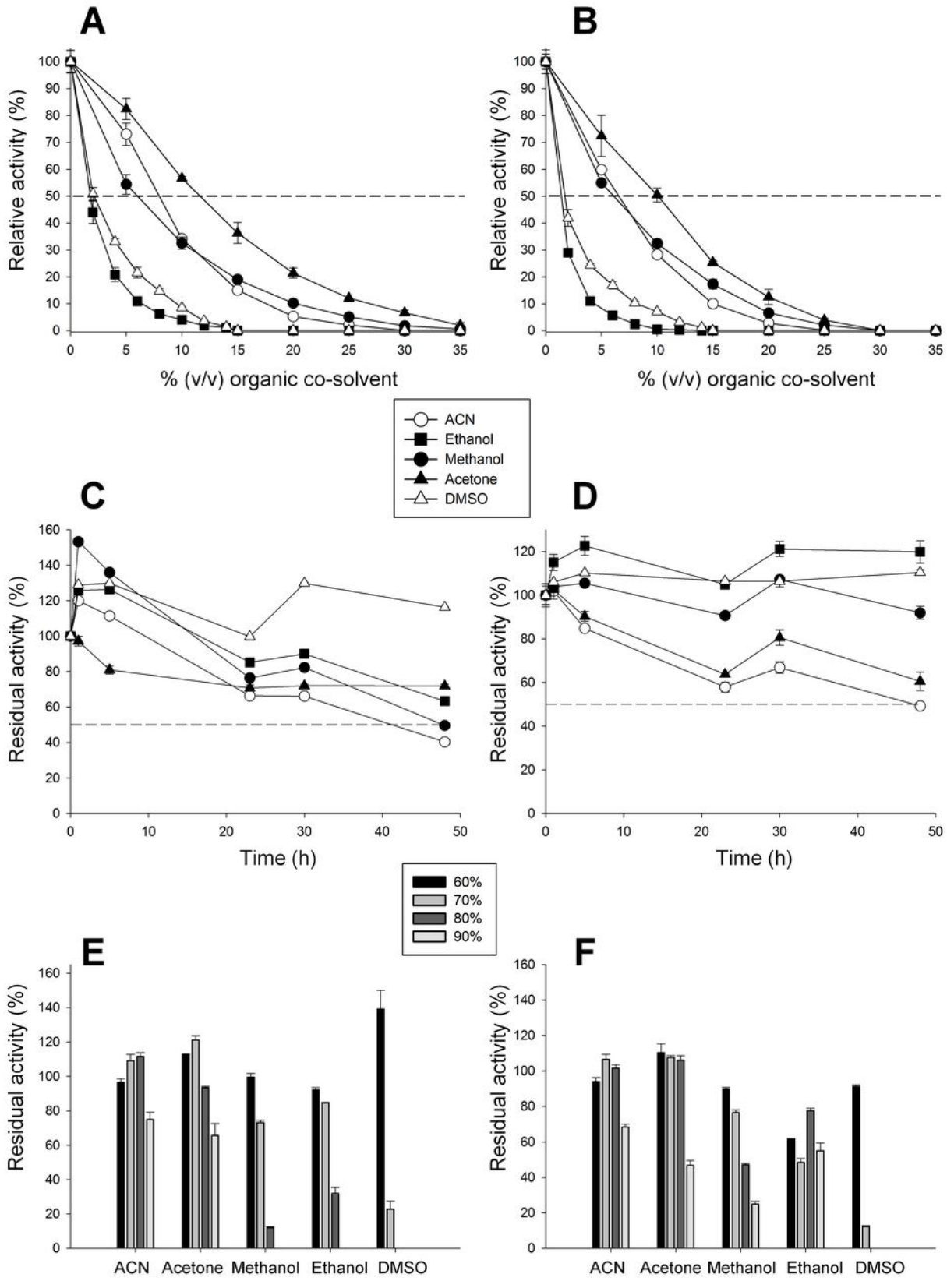


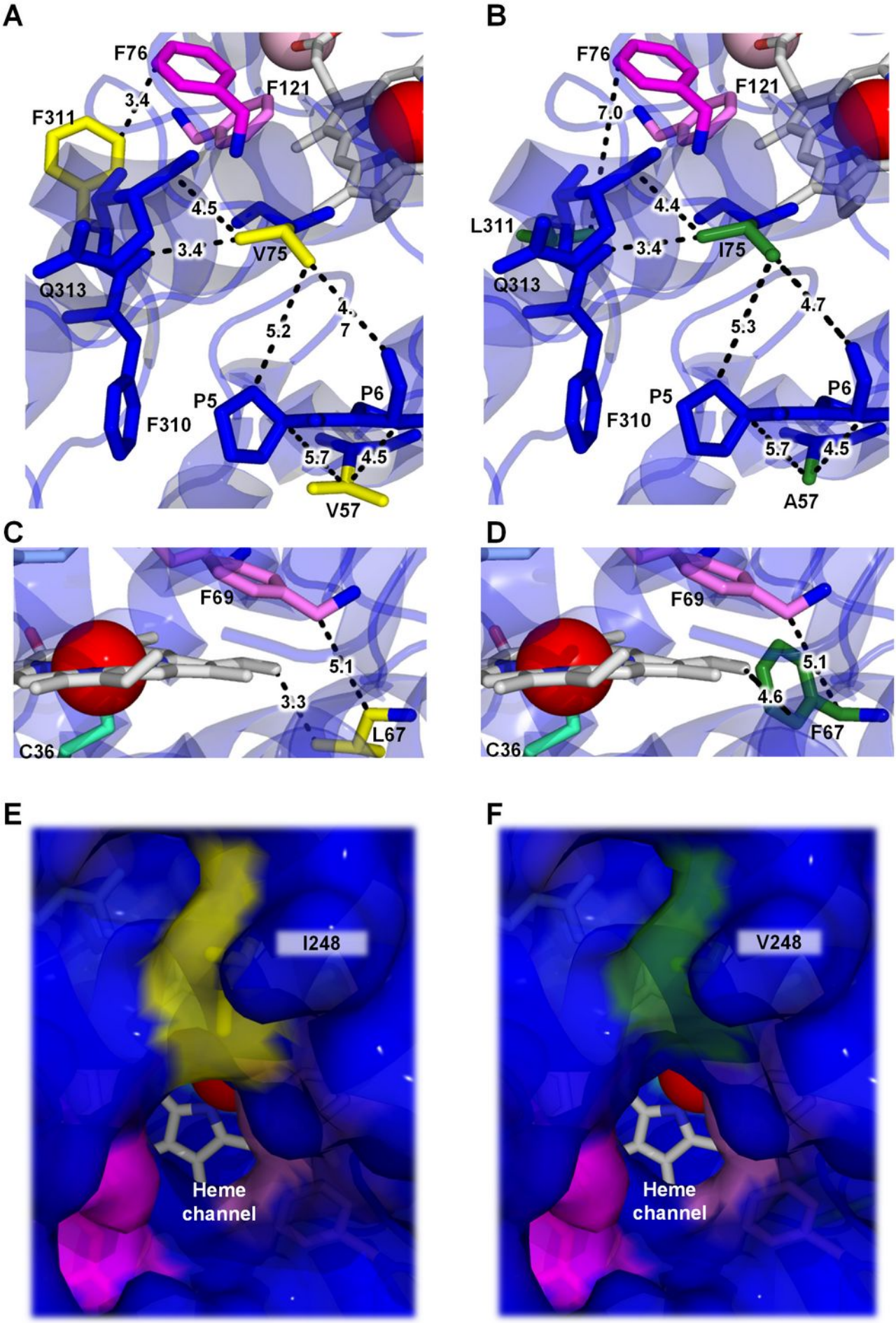

A
Absorbance (AU)
0,0
0,2
0,4
0,6
300
400
500
600
700
Wavelenght (nm)
B
Residual activity (%)
0
10
20
30
40
50
60
70
80
90
100
30
40
50
60
70
80
Temperature (ºC)
n*-UPO1
1A11 (1G)
12C12 (2G)
I13D3 (3G)
2A12 (4G)
PaDa-I (5G)
C
Relative activity (%)
0
20
40
60
80
100
2
3
4
5
6
7
8
pH
D
Relative activity (%)
0
20
40
60
80
100
3
4
5
6
7
8
9
pH



	U/L (crude extract)	U/mg	Secretion levels (mg/L)	Improvement breakdown		
				Total activity improvement TAI (fold increase)[b]	Specific activity (fold)	Expression (fold)
n-UPO1	2	n.m.	0.007[a]	1	1	1
n*-UPO1	54	233	0.2	27	1	27
PaDa-I mutant	6,500	828	7.8	3,250	3.6	1,114

[a]Secretion levels for n-UPO1 were calculated assuming the same specific activity as n*-UPO1. [b]Total activities (U/L) and TAI values (in fold) are calculated from large scale fermentation experiments. Activities were assessed in 100 mM sodium citrate/phosphate buffer pH 4.4 containing 0.3 mM ABTS and 2 mM H_2O_2.


A
B
Relative activity (%)
% (v/v) organic co-solvent
ACN
Ethanol
Methanol
Acetone
DMSO
C
D
Residual activity (%)
Time (h)
60%
70%
80%
90%
E
F
ACN
Acetone
Methanol
Ethanol
DMSO


A
F76
F121
F311
3.4
4.5
3.4
V75
Q313
5.2
4.7
P5
P6
F310
5.7
4.5
V57
B
F76
F121
7.0
4.4
L311
3.4
I75
Q313
5.3
4.7
P5
P6
F310
5.7
4.5
A57
C
F69
5.1
3.3
L67
C36
D
F69
5.1
4.6
F67
C36
E
I248
Heme
channel
F
V248
Heme
channel