

Functionally Diverse Peroxygenases by AlphaFold2, Design, and Signal Peptide Shuffling

Judith Münch, Niklas Dietz, Shiran Barber-Zucker, Franziska Seifert, Susanne Matschi, Pascal Püllmann, Sarel J. Fleishman,* and Martin J. Weissenborn*



Cite This: ACS Catal. 2024, 14, 4738–4748



Read Online

ACCESS |

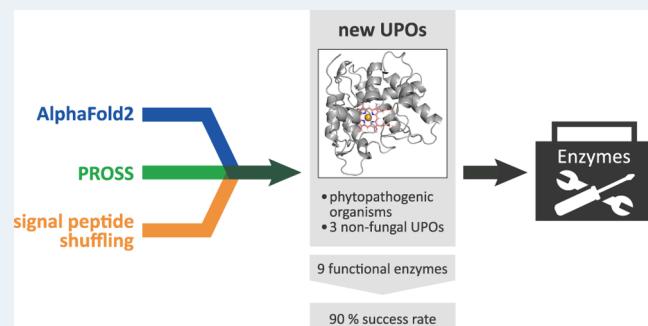
Metrics & More

Article Recommendations

Supporting Information

ABSTRACT: Unspecific peroxygenases (UPOs) are fungal enzymes that attract significant attention for their ability to perform versatile oxyfunctionalization reactions using H₂O₂. Unlike other oxygenases, UPOs do not require additional reductive equivalents or electron transfer chains that complicate basic and applied research. Nevertheless, UPOs generally exhibit low to no heterologous production levels and only four UPO structures have been determined to date by crystallography limiting their usefulness and obstructing research. To overcome this bottleneck, we implemented a workflow that applies PROSS stability design to AlphaFold2 model structures of 10 unique and diverse UPOs followed by a signal peptide shuffling to enable heterologous production. Nine UPOs were functionally produced in *Pichia pastoris*, including the recalcitrant CciUPO and three UPOs derived from oomycetes—the first nonfungal UPOs to be experimentally characterized. We conclude that the high accuracy and reliability of new modeling and design workflows dramatically expand the pool of enzymes for basic and applied research.

KEYWORDS: unspecific peroxygenase, yeast, *Pichia pastoris*, enzyme design, heterologous expression, protein stability



INTRODUCTION

Unspecific peroxygenases (UPOs) are secreted, fungal enzymes that have attracted great interest in recent years.^{1–3} They perform versatile oxyfunctionalization reactions like hydroxylations and epoxidations on a broad substrate scope.⁴ In contrast to P450 monooxygenases, UPOs do not rely on molecular oxygen, NAD(P)H, and electron transport chains but solely require hydrogen peroxide as a cosubstrate, in which the oxygen is already prerduced.^{5–8} They are divided into two major groups: long- and short-type UPOs (group I and II) of approximately 45 and 29 kDa.

The stability of UPOs, their substrate scope,⁴ and turnover numbers (TONs) of up to 900,000 strongly favor them for industrial application.⁹ So far, the main limitation to using UPOs has been their poor and difficult heterologous production. Although databases contain sequence information on >4000 putative UPOs,^{4,8} only about 50 have been heterologously produced since the discovery of this family in 2004.^{10–21} Thus, selecting UPOs for fundamental and applied research is dominated by considerations of heterologous production, limiting the phylogenetic and functional scope of enzymes that have been subjected to research.

As extracellular enzymes, UPOs require an N-terminal signal peptide that coordinates cell trafficking and secretion. The importance of the signal peptide choice regarding the quantity of the secreted target protein has been demonstrated in a

previous work: An *Aae*UPO variant (hereinafter: PaDa-I), which was evolved for higher heterologous production, exhibits nine mutations compared to the wild type, four of them are within the signal peptide. The mutations in the signal peptide alone resulted in a 27-fold improvement in functional production in *Saccharomyces cerevisiae*.¹⁵ Building on this work, our lab pursued an approach that focused on rapidly testing a signal peptide panel derived from yeast organisms, basidiomycetes, ascomycetes, and animals rather than evolving the natural signal peptide. With that approach, the production level in *S. cerevisiae* of the *Aae*UPO variant PaDa-I could be further doubled compared to the evolved signal peptide.¹⁸ This signal peptide shuffling technique further enabled the production of six other UPOs.¹⁸ Further increase in heterologous production was achieved through a combined shuffling of a promotor and signal peptide library in *Pichia pastoris* (syn. *Komagataella phaffii*).¹⁹

In addition to poor secretion, low heterologous production could be due to marginal protein stability.²² Production under

Received: February 7, 2024

Accepted: February 28, 2024

Published: March 13, 2024



ACS Publications

© 2024 The Authors. Published by
American Chemical Society

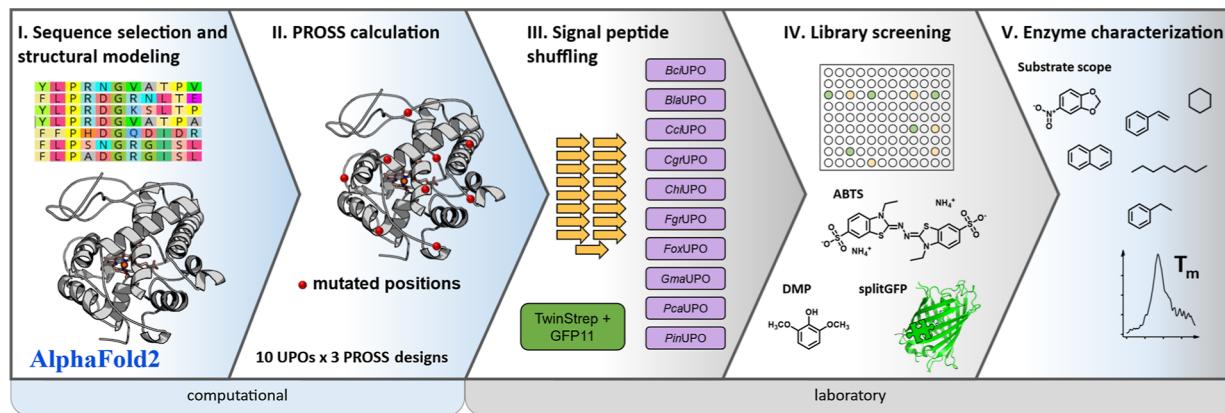


Figure 1. Overview of the different steps of the work protocol.

non-native conditions such as overexpression or the use of heterologous expression systems can lead to improperly folded proteins or aggregation²³ resulting in low protein yields. Increasing native-state stability may consequently lead to improved yields. Campaigns to enhance stability and production levels by introducing beneficial mutations based on directed evolution have been successful.^{24,25} However, the high labor intensity of this iterative approach renders it challenging to use when multiple enzymes are targeted. Furthermore, these approaches require detectable starting levels of secretion and activity in the relevant host, but some UPOs exhibit none. To address such challenges in heterologous production, in recent years, several algorithms have been developed to design stabilizing mutations.^{26–29} Among them is the Protein Repair One-Stop Shop (PROSS) algorithm.^{23,30} PROSS combines phylogenetic analysis with Rosetta atomistic calculations to design multipoint mutants with a favorable native-state energy. In dozens of previous studies, stability increases were accompanied by gains in functional production levels after a single design calculation and experimental screening of 3–5 constructs per protein target.^{31,32} These studies include, among others, the successful design of challenging oxygenases, such as high-redox potential laccases and versatile peroxidases (VPs).^{31,32}

Until recently, however, PROSS was limited only to the small fraction of enzymes for which crystallographic structures are available. With only four experimentally determined UPO structures, this class of enzymes is not amenable to atomistic design calculations.^{17,33–35} In a previous study dedicated to improve the functional expression of VPs using trRosetta, a legacy AI-based structure predictor, and PROSS, three of 11 enzymes that could not be functionally expressed in yeast before were functionally produced and characterized.³¹

The dramatic recent improvement in AI-based ab initio structure prediction methods such as AlphaFold2 has enabled predictions that are almost as accurate as those obtained from X-ray crystallography.^{36,37} These groundbreaking developments could enable the use of AlphaFold2 to provide structures for PROSS designs and empower researchers to unleash the full potential of improving protein production and stability directly from sequence with no recourse to experimental data. In the current study, we combine the much more accurate AlphaFold2 predictor with PROSS and signal peptide shuffling to 10 UPOs, achieving unprecedented levels of success with nine enzymes exhibiting functional heterologous expression compared with only one of the 10

wild-type enzymes exhibiting limited functional expression (Figure 1).

RESULTS AND DISCUSSION

Sequence Selection, Design Calculations, and Library Construction. We chose eight short-type UPOs derived from phytopathogens (Table 1), as UPOs from phytopathogenic

Table 1. Overview of the UPOs Studied in This Work

enzyme	original organism	type	division/class
BciUPO	<i>Botrytis cinerea</i> ^a	short-type	Ascomycota
BlaUPO	<i>Bremia lactucae</i> ^a	short-type	Oomycota
CciUPO	<i>Coprinopsis cinerea</i>	long-type	Basidiomycota
CgrUPO	<i>Colletotrichum graminicola</i> ^a	short-type	Ascomycota
ChiUPO	<i>Colletotrichum higginsianum</i> ^a	short-type	Ascomycota
FgrUPO	<i>Fusarium graminearum</i> ^a	short-type	Ascomycota
FoxUPO	<i>Fusarium oxysporum</i> ^a	short-type	Ascomycota
GmaUPO	<i>Galerina marginata</i>	long-type	Basidiomycota
PcaUPO	<i>Phytophthora cactorum</i> ^a	short-type	Oomycota
PinUPO	<i>Phytophthora infestans</i> ^a	short-type	Oomycota

^aPhytopathogenic organism.

fungi may be involved in coping with plant defensive compounds.^{38,39} Gaining access to those enzymes could prove to be a crucial step in understanding their role in the natural environment. Three of these UPOs originate from nonfungal oomycetes, which are more closely related to algae than fungi, further expanding the scope and novelty of our study as these would be the first functionally characterized UPOs from nonfungal origin.^{40,41} In addition to these short-type UPOs, we subjected two challenging long-type UPOs derived from basidiomycetes to prove the generality of the protocol (Figure 2). In previous work, we used signal peptide shuffling to successfully produce an UPO derived from *Galerina marginata* in *S. cerevisiae*; yet, the recombinant enzyme did not exhibit detectable activity.¹⁸ Another UPO derived from *G. marginata*, which exhibits 94% sequence identity to the one we selected, was shown recently (after the start of our work) to be functionally expressed in *P. pastoris*.¹⁴ Throughout this work, we use the name *GmaUPO* referring to the first published *GmaUPO*.¹⁸ *CciUPO* has so far only been produced in *Aspergillus oryzae*^{16,42} and not in any fast-growing microbe but holds substantial scientific interest for its relatively high activity levels.^{11,43–47}

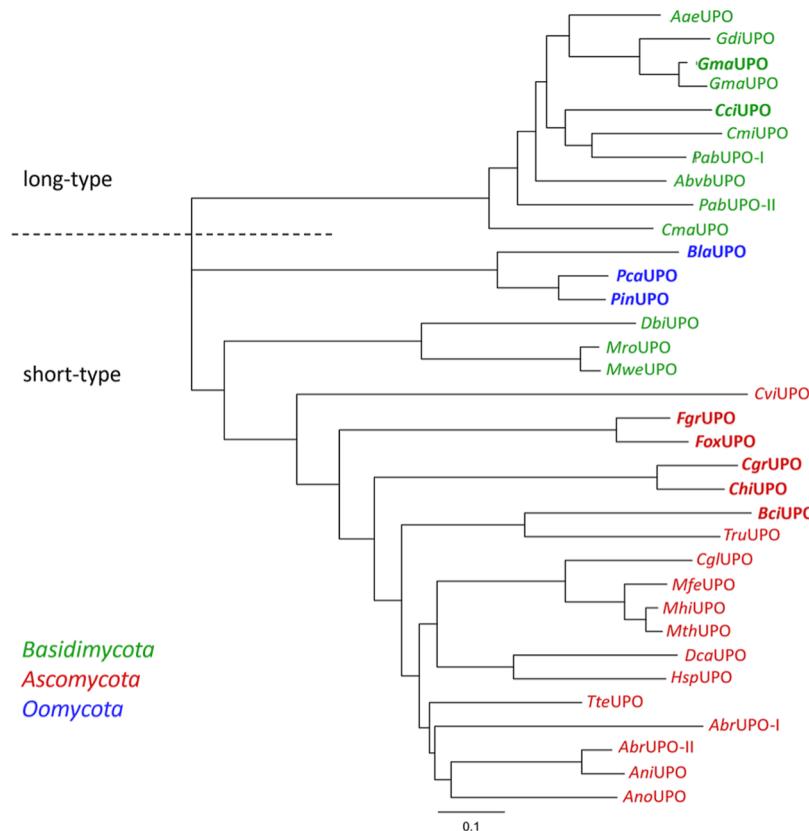


Figure 2. Neighbor-joining phylogenetic tree of selected previously known and heterologously produced UPOs (before September 2023) and our new UPO targets using Jukes–Cantor genetic distances. Basidiomycota (green), Ascomycota (red), and Oomycota (blue) (Table S3). UPOs in bold font were examined in this work. The dotted lines separate UPO sequences of group I and II (short and long UPOs). Generated with Geneious Prime 2023.1.2 (Biomatters Inc., Auckland, New Zealand).

The ten selected UPO sequences were modeled using AlphaFold2 (AF2; October 2021 release) and relaxed via the AF2 suite using AMBER. All models show high predicted reliability (average pLDDT score >90% for the best model). The model with the highest pLDDT score out of five calculated models for each target was subjected to the PROSS stability design.

PROSS combines phylogenetic analysis and Rosetta atomistic calculations to restrict design choices at each position according to their likelihood to occur in evolution and to stabilize the native state. In the final step, Rosetta combinatorial design is used to compute up to nine designs per starting structure with varying numbers of stabilizing mutations. As AF2 does not model ligands, we used the previously determined structures of *MroUPO* (pdb:SFUK) and PaDa-I (pdb:SOXU) as templates for determining the positions involved in the essential heme and magnesium (Mg^{2+})⁴⁸ binding. All AF2 models aligned well with the reference crystal structures (long UPOs root-mean-square deviation (RMSD) to SOXU: <0.5 Å; short UPOs RMSD to SFUK: <1.1 Å). Furthermore, the amino acids that coordinate heme and Mg^{2+} exhibit similar constellations as those in the reference crystallographic structures (Figure S18). These amino acids were not allowed to mutate or change conformations during design calculations. The design was further restricted in the substrate tunnel, putative N-glycosylation sites, regions that structurally vary between the five calculated AF2 models, regions with low AF2-predicted accuracy (pLDDT <90% or 5 Å from these residues), and

segments that are aligned to only a few sequence homologues. Finally, three designs for each enzyme were selected with different mutational loads of low (7–13 mutations, on average 4% of the sequence), medium (14–22 mutations, on average 7% of the sequence), and high (21–34 mutations, on average 11% of the sequence) for experimental testing (Table S1).

Seventeen different signal peptides⁴⁸ were fused to each of the 30 enzyme variants (three PROSS designs for each of the 10 UPO enzymes) leading to 510 different designs that were analyzed. All signal peptides and utilized vectors are available via Addgene (Yeast Secrete and Detect Kit #1000000166).

Functional Expression of Stabilized UPOs. Each variant was analyzed for its activity toward the two peroxidase substrates: (i) ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)) and (ii) DMP (2,6-dimethoxyphenol). Additionally, secretion was detected via a split-GFP assay⁴⁹ (Figures S1–S10).

PROSS designs of nine of the 10 UPO targets showed detectable secretion via the split-GFP assay (Figure 3). Of these nine, seven were active toward at least one of the substrates under screening conditions with five displaying activity toward ABTS and seven toward DMP (Figure 3). Solely, *FoxUPO* exhibited no secretion or activity.

Secretion and activity greatly varied among the PROSS variants. For *CgrUPO* and *FgrUPO*, the highest secretion levels were seen in the PROSS designs with the lowest mutational load. For *BciUPO*, *CciUPO*, *ChiUPO*, and *PinUPO*, the highest secretion levels were achieved with a medium mutational load. *BlaUPO*, *GmaUPO*, and *PcaUPO* showed

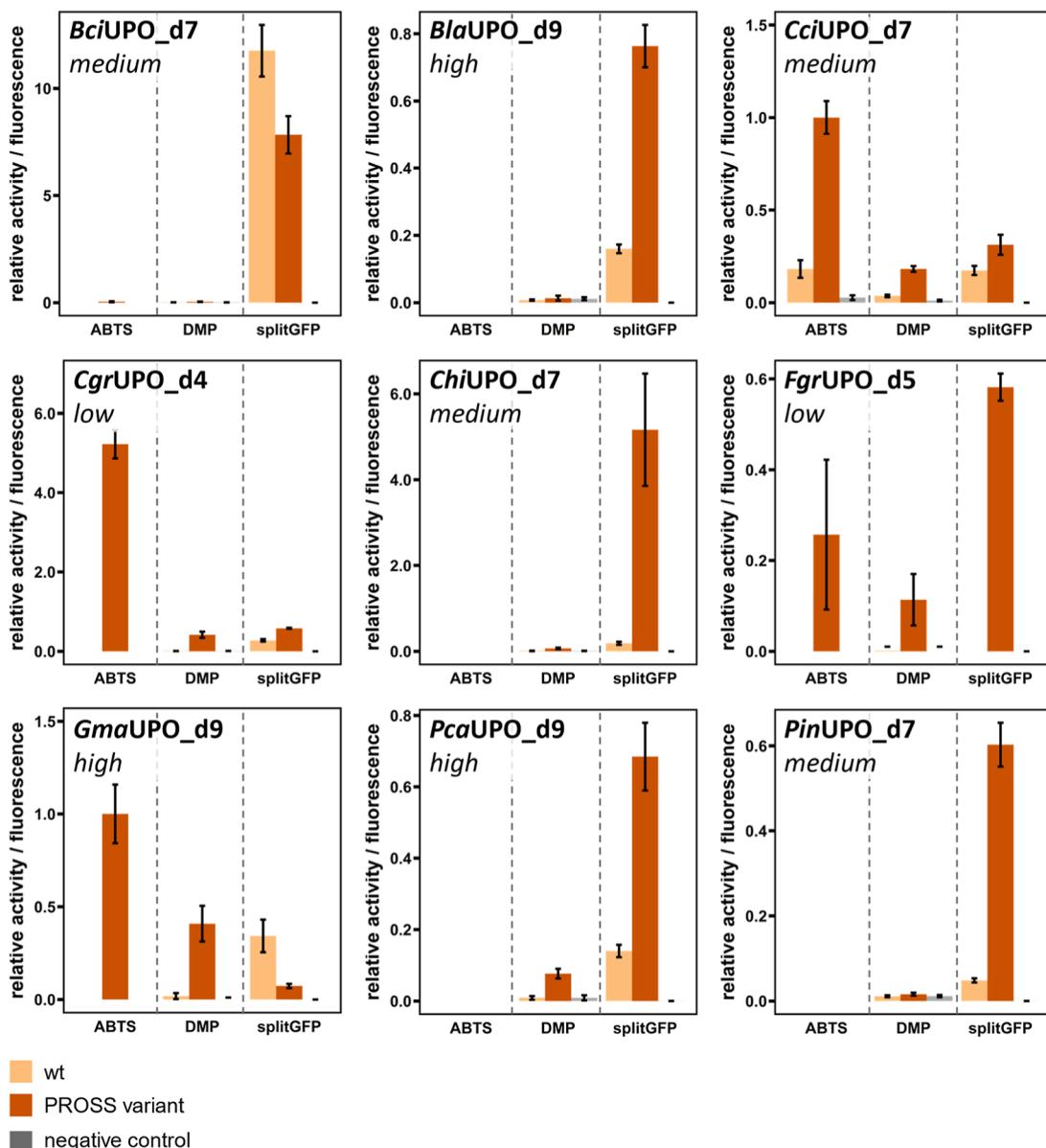


Figure 3. Comparison of the best PROSS design (orange) of each UPO with its corresponding wild type (yellow), expressed with the same signal peptide under screening conditions. Empty vector as the negative control (gray). Display of all three screening assays ABTS, DMP, and split-GFP. For each assay, the values are standardized to the positive control (*MthUPO*, not displayed). Only the ABTS assays of *CciUPO_d7* and *GmaUPO_d9* were standardized to the PROSS variant. The mutational load of the chosen PROSS variant (low, medium, or high) is given in *italics*. Data are mean \pm SD. Measurements were performed in general with eight replicates ($n = 8$), $n = 6$ in the case of *CgrUPO_d4*, and $n = 12$ in the case of *ChiUPO_d7*.

the highest secretion levels with the design carrying the highest number of mutations (Tables S1 and S2). Thus, as observed in previous studies, the number of mutations introduced during design is only weakly correlated with the observed improvement in production levels.⁵⁰

We also observed a huge variation in secreted enzyme levels, depending on the signal peptide. This variation can be demonstrated, for instance, by *ChiUPO* secretion relative to the positive control (*MthUPO*). The best *ChiUPO* PROSS variant (*ChiUPO_d7*, 19 mutations, medium mutational load) led to three different signal peptide-dependent secretion levels: (i) within a range of 100–300% of the fluorescence signal compared to the positive control, (ii) between 20 and 60% of the positive control signal, and (iii) below 10% of the positive control fluorescence signal (Figure S1). This observation

further emphasizes the enormous potential of signal peptide shuffling.^{18,19}

The best-performing signal peptide varied among the different UPOs. During the screening of *GmaUPO* and *CciUPO* variants, we confirmed that long-type UPO signal peptides were generally more prevalent for these long-type UPOs.^{18,19} *GmaUPO* showed the highest secretion level and activity with its natural signal peptide. In contrast, short-type UPOs did not show a preference for any signal peptide group, as previously observed.^{18,19} To reduce the screening effort of future studies of long-type UPOs, our results recommend using a reduced library containing only long-type UPO signal peptides. The diversity in signal peptides and differing mutational loads for the different PROSS variants indicates the strength of the combined approach, as neither the most

Table 2. Selection of Oxyfunctionalization Products Obtained with PROSS-Stabilized UPO Designs^a

A		B		C	
Enzyme	TON	Enzyme	TON	Enzyme	TON
	2		5		
<i>CciUPO_d7</i>	1793±353	<i>CciUPO_d7</i>	6±1	<i>CciUPO_d7</i>	24±9 (8)
	136±14		198±19		506±33 (10)
<i>GmaUPO_d9</i>	1452±85	<i>CgrUPO_d4</i>	38±1	<i>GmaUPO_d9</i>	430±27 (9)
	223±25		n.d.		56±1 (11)
		<i>GmaUPO_d9</i>	334±21	<i>GmaUPO_d9</i>	4±1 (8)
			2284±195		306±11 (10)
					331±43 (9)
					92±11 (11)
D		E		F	
Enzyme	TON	Enzyme	TON	Enzyme	TON ^[c]
			15		
<i>BciUPO_d7</i>	63±9	<i>BciUPO_d7</i>	39±9	<i>BciUPO_d7</i>	230±12
<i>BlaUPO_d9</i>	46±12	<i>BlaUPO_d9</i>	12±2	n.d.	
<i>CciUPO_d7</i>	3320±67	<i>CciUPO_d7</i>	67±7	<i>CciUPO_d7</i>	725±16
<i>CgrUPO_d4</i>	56±8	<i>CgrUPO_d4</i>	57±1	<i>CgrUPO_d4</i>	23±1
<i>GmaUPO_d9</i>	3747±378	<i>GmaUPO_d9</i>	98±13	<i>GmaUPO_d9</i>	699±14
<i>PcaUPO_d9</i>	179±13	<i>PcaUPO_d9</i>	n.d.	<i>PcaUPO_d9</i>	272± 6 ^[b]
<i>PinUPO_d7</i>	176±6			<i>PinUPO_d7</i>	270±7 ^[b]

^aTONs are based on GC-MS measurements if not stated otherwise and comparison with product standards. TON data are mean ± SD of measurements performed in triplicates. Reaction conditions are as follows: 4 mM substrate, 1.5 mM H₂O₂ (slow addition over the course of the reaction), 5% acetone (v/v), 250 nM enzyme, 0.1 mM potassium phosphate buffer (pH 7.0), 2 h reaction time, 30 °C. ^bReaction time overnight.

^cTONs are based on absorption measurement. TON data are mean ± SD of measurements made in triplicates.

suitable signal peptide nor the best PROSS design can be determined a priori; yet, the effort required to screen their combinations is more limited than a typical in vitro evolution campaign.

To assess the significance of the PROSS mutations on the enzymes, the best-performing signal peptide for each enzyme was combined with the corresponding wild-type variants. Most wild-type UPOs exhibited some degree of secretion, according to the split-GFP assay. Only wild-type *GmaUPO* and wild-type *BciUPO* showed increased secretion compared to their respective PROSS designs. All other wild-type UPOs displayed decreased secretion rates compared with the PROSS designs (Figure 3). Upscaling experiments revealed that secretion of wild-type *BciUPO* was not possible in shake flasks, while the best PROSS design of *BciUPO* (*BciUPO_d7*, medium mutational load) was actively secreted under these conditions, suggesting that the production conditions are another critical determinant of functional production.⁵⁰ The only wild-type enzyme that showed activity during this work was *CciUPO*, but its activity level was fivefold lower than the best PROSS design for this enzyme (Figure 3). From these experiments, we conclude that although some of the selected UPOs are producible in their wild-type form, the PROSS designs showed higher production levels, and the design process was essential for obtaining functional UPOs.

Diverse Enzyme Characteristics and Substrate Scope in Designs. For each enzyme, we chose the PROSS design and signal peptide combination that showed the highest

activity for DMP or ABTS (Figure 3 and Table S2) for subsequent shake flask expression and enzyme characterization. The occurrence and correct identity of all enzymes were confirmed via LC-MS (Figure S17 and Table S7). The UV absorption spectra showed the characteristic of heme-thiolate proteins with a Soret band maximum around 420 nm for *BciUPO_d7* (424 nm), *CciUPO_d7* (418 nm), *CgrUPO_d4* (420 nm), *ChiUPO_d7* (410 nm), *FgrUPO_d5* (415 nm), *GmaUPO_d9* (418 nm), and *PcaUPO_d9* (417 nm) (Figure S13). For *BlaUPO_d9* and *PinUPO_d7*, this maximum could not be detected due to challenges during purification and potential problems in heme incorporation (Figure S13).

We analyzed the substrate scope for each enzyme using concentrated supernatant, and TONs were determined with GC-MS. We chose test substrates that demand two-electron peroxygenase activity rather than one-electron peroxidase activity as these are the more challenging reactions (Table 2). Among the substrates were two with activated C–H bonds (NBD [5-nitro-1,3-benzodioxole] and ethylbenzene), styrene as an epoxidation substrate, the aromatic substrate naphthalene, and the two aliphatic substrates, cyclohexane and octane with nonactivated sp³-carbons. Very low enzyme concentrations (production titer <0.5 mg/L) for *ChiUPO_d7* and *FgrUPO_d5* after shake flask expression impeded determining their substrate scope.

The overall highest activities were demonstrated by the two long-type UPOs, *CciUPO_d7* and *GmaUPO_d9*. Styrene was

Table 3. Comparison of the Transition Temperature (T_m) and Onset of Aggregation (T_{agg}) for Selected PROSS-Optimized UPOs Expressed in Both *P. pastoris* and *S. cerevisiae*

enzyme	expression host <i>P. pastoris</i> C-terminal-tag TwinStrep (two additional tryptophan)			expression host <i>S. cerevisiae</i> C-terminal-tag His2 (no additional tryptophan)			$\Delta T_m = T_m (Pp) - T_m (Sc)$
	fluorescence shift	T_m	T_{agg}^a	fluorescence shift	T_m	T_{agg}^a	
<i>BciUPO_d7</i>	blue	44.3 ± 0.2 °C	33.9 ± 0.5 °C	red	59.4 ± 0.2 °C	n.d.	-15.1
<i>CciUPO_d7</i>	blue	58.7 ± 0.3 °C	63.9 ± 0.9 °C	red	46.6 ± 0.3 °C	56.2 ± 0.3 °C	+12.1
<i>CgrUPO_d4</i>	blue	42.8 ± 1.4 °C	37.7 ± 0.1 °C	red	54.3 ± 0.1 °C	n.d.	-11.5
<i>ChiUPO_d7</i>	red	62.7 ± 0.1 °C	n.d.	red	54.4 ± 0.1 °C	n.d.	+8.3
<i>GmaUPO_d9</i>	blue	58.5 ± 0.2 °C	49.9 ± 0.2 °C	red	46.7 ± 0.3 °C	46.0 ± 0.9 °C	+11.8
<i>PcaUPO_d9</i>	blue	50.0 ± 1.6 °C	41.1 ± 1.3 °C	red	52.9 ± 0.1 °C	n.d.	-2.9

^aOnset of aggregation; n.d.—not detected, data are mean ± SD, and measurements were performed in triplicates.

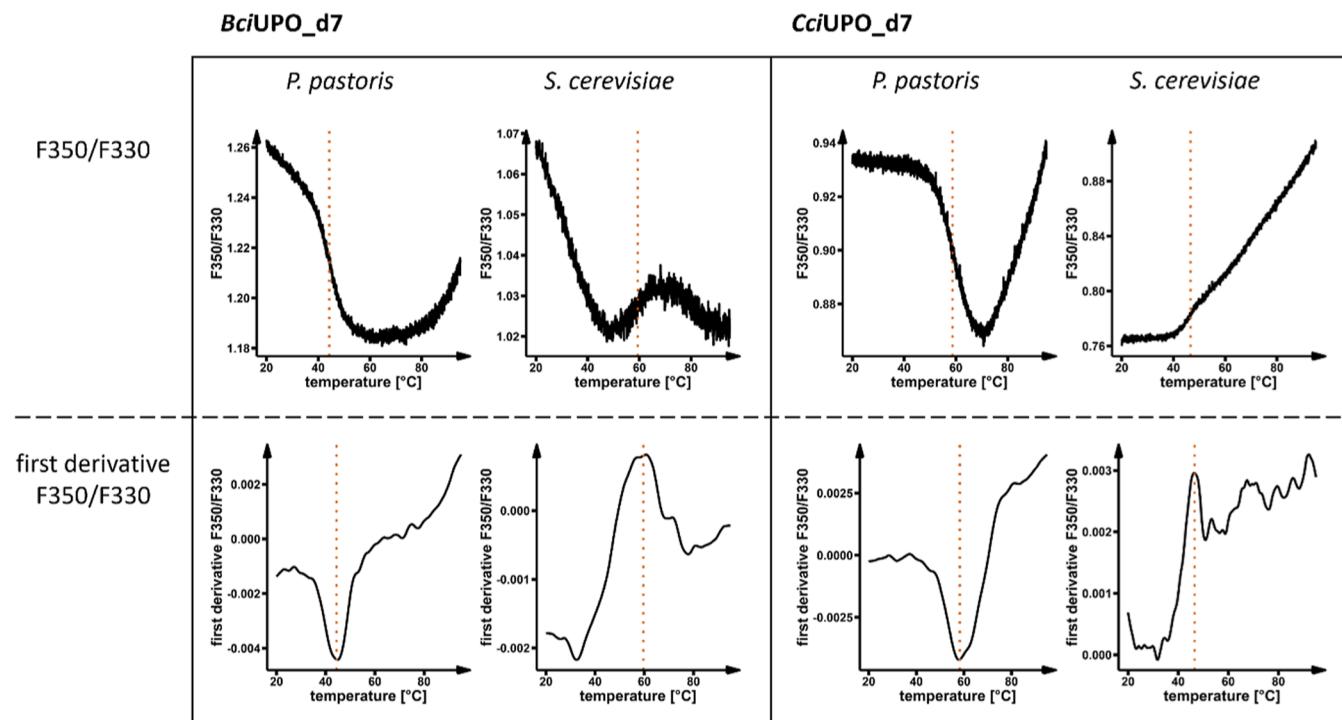


Figure 4. Comparison of fluorescence intensity change during temperature increase for *BciUPO_d7* and *CciUPO_d7*. Both enzymes were expressed in *P. pastoris* (with the C-terminal TwinStrep tag) and *S. cerevisiae* (with the C-terminal His2 tag). The ratio of fluorescence intensity at 350 nm/330 nm is shown in the upper row, and the first derivative of this fluorescence ratio is shown in the lower row. The inflection point, at which the transition temperature T_m in each case was determined, is marked (orange dotted line).

converted to styrene oxide with more than 3000 TONs by both enzymes, and ethylbenzene was mainly overoxidized to acetophenone with nearly 2500 TONs (Table 2). Both activities are within the range of other UPOs like PaDa-I (4000 TONs for ethylbenzene hydroxylation and 3000 TONs for styrene epoxidation) or *MthUPO* (500 TONs for ethylbenzene hydroxylation and 1100 TONs for styrene epoxidation).^{18,51} Both long-type UPOs also exhibit activities toward nonactivated C–H bonds, shown here by the conversion of cyclohexane to cyclohexanol with more than 1500 TONs and the oxyfunctionalization of octane with total TONs of 730 (*GmaUPO_d9*) and 1010 (*CciUPO_d7*) (Table 2), which is within the range of *MthUPO*,⁵¹ but significantly lower compared to 20,000 TONs for PaDa-I.⁵² Activity toward styrene epoxidation was also displayed by *BlaUPO_d9*, *BciUPO_d7*, *CgrUPO_d4*, *PcaUPO_d9*, and *PinUPO_d7*, the latter four also convert NBD, but all reactions lead only to low TONs between 20 and 270 (Table 2). These results

further demonstrate that *BlaUPO_d9* and *PinUPO_d7* are indeed functionally produced, even though they did not show activity toward ABTS or DMP under screening conditions, leading to a total of nine functionally produced UPOs. Aromatic hydroxylation was detected for *BciUPO_d7*, *BlaUPO_d9*, *CciUPO_d7*, *CgrUPO_d4*, *GmaUPO_d9*, and *PcaUPO_d9* with naphthalene as the substrate but with low TONs between 12 and 184.

The determined enantiomeric ratios (e.r.) for styrene oxide (13) and 1-phenylethanol (5) show only small differences between the investigated enzymes with a lowest e.r. of 59:41 (*CciUPO*, styrene oxide) and a highest e.r. of 85:15 (*GmaUPO*, 1-phenylethanol). All enzymes showed a preference for the *R* enantiomer in the case of both products (Table S8).

In addition to the substrate scope, we also determined the apparent transition temperature (T_m) for all best PROSS designs produced in both *P. pastoris* and *S. cerevisiae*. The

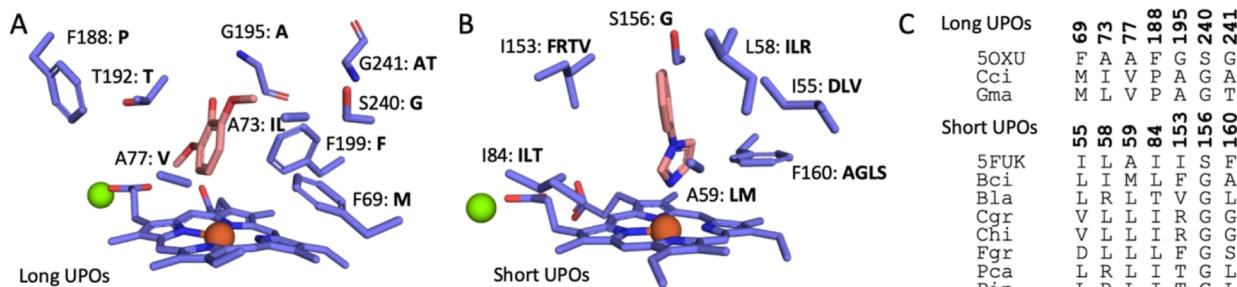


Figure 5. Active-site sequence diversity among successfully designed UPOs. (A) Long UPO active site. PaDa-I (pdb:5OXU) heme and active-site residues are presented in purple sticks (unless glycine, only side chain atoms are presented) and Mg²⁺ is presented as a green sphere. DMP was grafted from pdb:SOY2 and is presented in pink sticks. PaDa-I positions and identities are written next to each presented residue (left side). For each position, the identities of *Cci*UPO and *Gma*UPO in the same position are written on the right side, respectively. (B) Short UPO active site. *Mro*UPO (pdb:5FUK) heme and active-site residues are presented in purple sticks (only side chain atoms are presented) and Mg²⁺ is presented as a green sphere. 1-Phenylimidazole was grafted from pdb:7O1X and is presented in pink sticks. *Mro*UPO positions and identities are written next to each presented residue (except for position 58 that is assigned as leucine, as 5FUK possesses T58L mutation; left side). For each position, the identities of *Bci*UPO, *Bla*UPO, *Cgr*UPO, *Chi*UPO, *Fgr*UPO, *Pca*UPO, and *Pin*UPO in the same position are written on the right side. (C) Identities of each UPO in the presented positions (only positions that are not completely conserved are shown). Long UPOs: PaDa-I, *Cci*UPO, and *Gma*UPO are 63–72% identical to one another. Short UPOs: *Cgr*UPO and *Chi*UPO are 80% identical, *Pca*UPO and *Pin*UPO are 87% identical, *Bla*UPO is 66% identical to both *Pca*UPO and *Pin*UPO, and all other UPO pairs are 25–50% identical to one another.

enzymes expressed in *P. pastoris* contained a C-terminal TwinStrep tag for purification while the counterparts from *S. cerevisiae* carried a C-terminal His2 tag. Both sets of UPOs were purified via respective affinity chromatography. To investigate the thermal stability, we applied nano differential scanning fluorimetry (nanoDSF). Parallel to the fluorescence measurements, we detected light scattering to determine the potential onset of aggregation. Interestingly, the resulting temperature values for both sets were different depending on the production host and C-terminal tag. For *P. pastoris* as the production host, *ChiUPO_d7* showed the highest apparent transition temperature with 62.7 °C. Both long-type UPOs *CciUPO_d7* and *GmaUPO_d9* showed similar transition temperatures of about 58.6 °C. The lowest transition temperature was measured for *CgrUPO_d7* with 42.8 °C (Table 3, Figures 4 and S15A). Comparing the transition temperatures of the recombinant enzymes secreted from *S. cerevisiae*, *BciUPO_d7* had the highest T_m with 59.4 °C, whereas the lowest value was found for *CciUPO_d4* with 46.6 °C (Table 3, Figures 4 and S15B). The transition and onset temperatures of *BlaUPO*, *FgrUPO*, and *PinUPO* could not be determined with satisfying reliability due to very low enzyme concentrations after purification. The data we obtained revealed strong differences in T_m for the individual PROSS designs, e.g., ΔT_m of 15.1 °C for *BciUPO_d7* with a higher thermal stability when expressed in *S. cerevisiae* and ΔT_m of 12.2 °C for *CciUPO_d7* with a higher thermal stability when expressed in *P. pastoris*. We further noticed a reverse trajectory of the fluorescence intensity change (ratio at 350 nm/330 nm) for *BciUPO_d7*, *CciUPO_d7*, *CgrUPO_d4*, and *GmaUPO_d9*. With increasing temperature, an increase in fluorescence is observed for enzymes expressed in *S. cerevisiae* with the His2-GFP11 tag, whereas a decrease is mainly observed for variants expressed in *P. pastoris* with the TwinStrep-GFP11 tag (Table 3, Figures 4 and S15). Considering our approach, we introduced two additional surface-exposed Trp residues which are present in the TwinStrep-GFP11 tag⁵³ in contrast to the His2-GFP11 tag. If the majority of Trp residues beforehand are buried in the enzyme, as is the case, e.g., for *BciUPO_d7*, adding two surface-exposed Trp residues can lead to a different

fluorescence behavior as observed in our case. Regarding the positioning of Trp in *CciUPO_d7*, which has rather surface-exposed Trp residues (Figure S14), the blue shift trajectory, as detected for the construct from *P. pastoris* as the host, can be expected. Hence, the red-shifted progress of the enzyme from *S. cerevisiae* is rather surprising. We assume that the glycosylation pattern (both O- and N-glycosylation) might play an important role in this context as well as in the differences in protein stability. It is well known that glycosylation can strongly affect protein folding and stability^{54,55} as well as intrinsic protein dynamics⁵⁶ and therefore activity.^{57,58} Previous research also revealed the influence of glycans on the fluorescence behavior of intrinsic Trp fluorescence itself.^{59,60} Considering the differences published for glycosylation patterns from *P. pastoris* and *S. cerevisiae*,⁶¹ we therefore hypothesize that these differences affect both enzyme stability and the fluorescence behavior of our PROSS designs. Glycosylation has also been shown to have an influence on the aggregation tendency of proteins.^{62–64} The light scattering measurement to detect the temperature for the onset of aggregation supports this finding also for the enzymes investigated in the present study (Table 3 and Figure S16). While only *CciUPO_d7* and *GmaUP_d9* expressed in *S. cerevisiae* show typical aggregation behavior, the aggregation tendency for the UPOs expressed in *P. pastoris* is more pronounced. These data could be due to hyperglycosylation in *S. cerevisiae*.⁶⁵ Overall, it is crucial to acknowledge that both the host organism, along with the resulting glycosylation pattern, and the chosen tag can influence enzyme stability and folding behavior.

Structural Underpinnings of Functional Diversity.

Finally, we analyzed the AF2 structure models of the nine UPOs that were functional following the design (Figure 5). The models exhibit only minor backbone differences from the reference crystallographic structure (Figure S18). The enzyme active sites were not allowed to be designed, and any differences in amino acids in them are strictly due to the natural diversity of UPOs. Whereas the long-type UPOs exhibit few amino acid changes in the active site (Figure 5A,C), the short-chain ones exhibit high diversity (Figure 5B,C). For instance, positions 55, 58, and 153 (pdb 5FUK

numbering) exhibit hydrophobic identities in most of the short UPOs and charged identities in some of the variants. Even among the hydrophobic-to-hydrophobic exchanges, some are predicted to change the sterics of the active site dramatically, as in Val/Phe at position 153 and Gly/Leu at position 160. These large changes in electrostatics and sterics are likely to change the positioning of the substrate relative to that of the heme, leading to the observed functional differences.

CONCLUSIONS

Due to their challenging heterologous production, only a few UPOs are available for engineering campaigns, and their respective crystal structures are rarely determined. The improvement of ab initio structure prediction tools like AF2 provides access to numerous—if not all—UPO structures and thus opens the way to structure-based stability design. We combined AF2 for structure prediction, PROSS for introducing stabilizing mutations, and signal peptide shuffling to increase enzyme secretion. These three methods had not been previously used in conjunction and enabled the production of a highly challenging class of UPOs. Our success rate in functional expression reached 90%, as only *Fox*UPO did not show any activity or secretion. This rate is significantly higher compared to our previous workflow where we combined only the legacy trRosetta and PROSS without a signal peptide shuffling, which resulted in the successful expression of only three of 11 target VPs.³¹ PROSS proved to be very beneficial for gaining functionally produced enzymes as (i) secretion levels are increased substantially in most cases for the PROSS design compared to the wild-type enzyme and (ii) eight out of nine wild-type enzymes did not show any activity for the tested screening substrates (ABTS and DMP), while seven PROSS UPOs displayed activity on at least one of the screening substrates (ABTS or DMP). All nine produced PROSS UPOs were active on at least one tested substrate, if not a screening substrate, then during substrate scope analysis with six additional substrates. This high success rate demonstrates a clear path to protein engineering of even challenging enzymes using modern modeling and design software. In fact, our success rate surpasses that of other PROSS applications^{31,50} suggesting that either the combination with signal peptide shuffling has a beneficial effect or the AF2-predicted structures offer an advantage over crystal structures. For the latter possibility, we speculate that the high success rate may be due to restricting designs in positions that exhibit low modeling confidence (AF2 pLDDT scores <90%) and their vicinity (this functionality is automatically enabled in the PROSS web server when AF2 models are used; <https://PROSS.weizmann.ac.il>). Crystallographic structure analysis does not provide a comparable way to assign confidence to structures.

Nevertheless, PROSS on its own is insufficient to gain functionally active, secreted UPOs. In our work, we combined 30 PROSS designs of 10 UPOs with 17 different signal peptides and screened them with three different assays to determine the secretion level and activity for all combinations. Our work shows again, in accordance with the previous work,^{18,19} that no single most suitable signal peptide was found for all UPOs, and secretion rates differed greatly between different samples during the signal-peptide-shuffling screening.

Successful secretion of UPOs in a fast-growing microbial host is the first essential step on the path to the facile access of customized enzymes for industrial application. Seven of the newly characterized enzymes derive from phytopathogenic

organisms. It may be interesting to study their effect on plant material to gain further insights into the natural functions of UPOs and their possible interaction with the plant defense system. Three of the UPOs are derived from oomycetes, extending the range of available UPOs, for the first time, to nonfungal organisms.

It is important to note that design calculations were not applied to the active site. The sequence variations within the active site (Figure 5) are therefore entirely derived from natural UPO diversity. In most cases, the variants introduce multiple simultaneous changes relative to one another. Combinations of mutations within the active site are likely to be epistatic with one another or with mutations and backbone differences outside of the active site.^{66–68} Such epistatic relationships are known to slow functional innovation in natural and laboratory evolution.⁶⁹ Thus, our structural analysis highlights the major strength of the design process: instead of painstakingly re-engineering the active site and the enzyme backbone, it exploits the natural structural and sequence diversity to expose new substrate specificities and reactivities. In the case of UPOs, thousands of natural sequences are known, only a few dozen of which have been successfully characterized until now. We envision that the computational design will dramatically accelerate the discovery of oxyfunctionalization reactions in this family.

Finally, the workflow we employed can be readily adopted by other laboratories, as AF2 and PROSS are available by web servers, and all required signal peptides and vectors for yeast production and secretion have been previously deposited at Addgene (Yeast Secrete and Detect Kit #1000000166).^{18,19} In this study, we could readily identify the heme-binding site by a comparison to two of the known experimental structures of UPOs. If a homologous experimental structure is not known, however, the active site may be identified by using other means. Conservation analysis or prior experimental studies may provide guidance in such cases. Additionally, recent advances enable ab initio structure prediction of the protein with known ligands.⁷⁰ These new methods may extend the pipeline that we have demonstrated here even to cases in which no homologous structure has been experimentally determined. Thus, the combination of AI-based structure prediction and atomistic design can provide enzymologists and protein engineers with access to a vast array of functions from previously untappable natural enzymes.

ASSOCIATED CONTENT

SI Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acscatal.4c00883>.

Material and methods; gene and protein sequence of all enzymes investigated in this study; initial screening data; GC-MS temperature programs; utilized linker and tags; initial screening data and overview over best PROSS design and signal peptide combination; relative abundance of signal peptides during the initial screening; transition temperature profiles of all purified enzymes; calibration curve for GC-MS measurements; and absorption spectra (PDF)

AUTHOR INFORMATION

Corresponding Authors

Sarel J. Fleishman — Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7600001, Israel; [ORCID iD](https://orcid.org/0000-0003-3177-7560); Email: sarel@weizmann.ac.il

Martin J. Weissenborn — Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany; [ORCID iD](https://orcid.org/0000-0002-1200-4485); Email: martin.weissenborn@chemie.uni-halle.de

Authors

Judith Münch — Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany

Niklas Dietz — Institute of Chemistry, Martin Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany; Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany

Shiran Barber-Zucker — Department of Biomolecular Sciences, Weizmann Institute of Science, Rehovot 7600001, Israel; Present Address: Scala Biodesign LTD, 50 Dizengoff St., Tel Aviv 6433222, Israel

Franziska Seifert — Department of Pharmaceutical Technology and Biopharmacy, Institute of Pharmacy, Martin-Luther-University Halle-Wittenberg, Halle (Saale) 06120, Germany

Susanne Matschi — Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany

Pascal Püllmann — Leibniz Institute of Plant Biochemistry, Halle (Saale) 06120, Germany; Present Address: Molecular Design and Engineering, Bayer AG, Aprather Weg 18A, 42113 Wuppertal, Germany

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acscatal.4c00883>

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript. Both J.M. and N.D. contributed equally and have the right to list their name first in their CV.

Funding

Friedrich-Naumann-Stiftung für die Freiheit Bundesministerium für Bildung und Forschung (Maßgeschneiderte Inhaltsstoffe 2, 031B0834A), German Research Foundation (DFG, project ID 436494874, TP A05, RTG 2670), German Research Foundation (DFG, project ID 505185500), Israel Science Foundation (ISF, #1844).

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

J.M. thanks the Friedrich-Naumann-Stiftung für die Freiheit for a PhD scholarship. M.J.W. and J.M. thank the Bundesministerium für Bildung und Forschung (Maßgeschneiderte Inhaltsstoffe 2, 031B0834A) and German Research Foundation (DFG, project ID 436494874, TP A05, RTG 2670) for their generous funding. N.D. thanks the German Research Foundation (DFG, project ID 505185500) for their support. We would like to thank Domenika Thieme (Leibniz-Institute for Plant Biochemistry) for performing

protein identification measurements at the LC-MS after trypsin digestion.

ABBREVIATIONS

AF2, AlphaFold2; ABTS, 2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid); DMP, 2,6-dimethoxyphenol; GC-MS, gas chromatography-mass spectrometry; NBD, 5-nitro-1,3-benzodioxole; *P. pastoris*, *Pichia pastoris* (*Komagataella phaffii*); PROSS, Protein Repair One-Stop Shop; T_{agg} , onset of aggregation; T_m , transition temperature; TON, turnover number; UPO, unspecific peroxygenase

REFERENCES

- (1) Beltrán-Nogal, A.; Sánchez-Moreno, I.; Méndez-Sánchez, D.; Gómez de Santos, P.; Hollmann, F.; Alcalde, M. Surfing the wave of oxyfunctionalization chemistry by engineering fungal unspecific peroxygenases. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102342.
- (2) Monterrey, D. T.; Menés-Rubio, A.; Keser, M.; Gonzalez-Perez, D.; Alcalde, M. Unspecific peroxygenases: The pot of gold at the end of the oxyfunctionalization rainbow? *Curr. Opin. Green Sustain. Chem.* **2023**, *41*, 100786.
- (3) Kinner, A.; Rosenthal, K.; Lütz, S. Identification and expression of new unspecific peroxygenases—Recent advances, challenges and opportunities. *Front. Bioeng. Biotechnol.* **2021**, *9*, 705630.
- (4) Hobisch, M.; Holtmann, D.; Gomez de Santos, P.; Alcalde, M.; Hollmann, F.; Kara, S. Recent developments in the use of peroxygenases—Exploring their high potential in selective oxyfunctionalizations. *Biotechnol. Adv.* **2021**, *51*, 107615.
- (5) Li, Z.; Jiang, Y.; Guengerich, F. P.; Ma, L.; Li, S.; Zhang, W. Engineering cytochrome P450 enzyme systems for biomedical and biotechnological applications. *J. Biol. Chem.* **2020**, *295* (3), 833–849.
- (6) Guengerich, F. P.; Waterman, M. R.; Egli, M. Recent structural insights into cytochrome P450 function. *TIPS* **2016**, *37* (8), 625–640.
- (7) Manikandan, P.; Nagini, S. Cytochrome P450 structure, function and clinical significance: a review. *Curr. Drug Targets* **2018**, *19* (1), 38–54.
- (8) Münch, J.; Püllmann, P.; Zhang, W.; Weissenborn, M. J. Enzymatic hydroxylations of sp³-carbons. *ACS Catal.* **2021**, *11* (15), 9168–9203.
- (9) Hobisch, M.; De Santis, P.; Serban, S.; Basso, A.; Byström, E.; Kara, S. Peroxygenase-Driven Ethylbenzene Hydroxylation in a Rotating Bed Reactor. *Org. Process Res. Dev.* **2022**, *26* (9), 2761–2765.
- (10) Aranda, C.; Municoy, M.; Guallar, V.; Kiebist, J.; Scheibner, K.; Ullrich, R.; del Río, J. C.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Selective synthesis of 4-hydroxisophorone and 4-ketoisophorone by fungal peroxygenases. *Catal. Sci. Technol.* **2019**, *9* (6), 1398–1405.
- (11) Babot, E. D.; del Río, J. C.; Kalum, L.; Martínez, A. T.; Gutiérrez, A. Oxyfunctionalization of aliphatic compounds by a recombinant peroxygenase from *Coprinopsis cinerea*. *Biotechnol. Bioeng.* **2013**, *110* (9), 2323–2332.
- (12) Bormann, S.; Kellner, H.; Hermes, J.; Herzog, R.; Ullrich, R.; Liers, C.; Ulber, R.; Hofrichter, M.; Holtmann, D. Broadening the biocatalytic toolbox—Screening and expression of new unspecific peroxygenases. *Antioxidants* **2022**, *11* (2), 223.
- (13) Carro, J.; González-Benjumea, A.; Fernández-Fueyo, E.; Aranda, C.; Guallar, V.; Gutiérrez, A.; Martínez, A. T. Modulating fatty acid epoxidation vs hydroxylation in a fungal peroxygenase. *ACS Catal.* **2019**, *9* (7), 6234–6242.
- (14) Ma, Y.; Liang, H.; Zhao, Z.; Wu, B.; Lan, D.; Hollmann, F.; Wang, Y. A Novel Unspecific Peroxygenase from *Galatian Marginata* for Biocatalytic Oxyfunctionalization Reactions. *J. Mol. Catal.* **2022**, *531*, 112707.
- (15) Molina-Espeja, P.; García-Ruiz, E.; Gonzalez-Perez, D.; Ullrich, R.; Hofrichter, M.; Alcalde, M. Directed evolution of unspecific

- peroxyxygenase from Agrocybe aegerita. *Appl. Environ. Microbiol.* **2014**, *80* (11), 3496–3507.
- (16) Pecyna, M. J.; Schnorr, K. M.; Ullrich, R.; Scheibner, K.; Kluge, M. G.; Hofrichter, M. Fungal peroxyxygenases and methods of application. WO 2008/119780 A2, 2008.
- (17) Rotilio, L.; Swoboda, A.; Ebner, K.; Rinnofner, C.; Glieder, A.; Kroutil, W.; Mattevi, A. Structural and biochemical studies enlighten the unspecific peroxyxygenase from Hypoxylon sp. EC38 as an efficient oxidative biocatalyst. *ACS Catal.* **2021**, *11* (18), 11511–11525.
- (18) Püllmann, P.; Knorrseidt, A.; Münch, J.; Palme, P. R.; Hoehnwarter, W.; Marillonnet, S.; Alcalde, M.; Westermann, B.; Weissenborn, M. J. A modular two yeast species secretion system for the production and preparative application of unspecific peroxyxygenases. *Commun. Biol.* **2021**, *4* (1), S62.
- (19) Püllmann, P.; Weissenborn, M. J. Improving the heterologous production of fungal peroxyxygenases through an episomal *Pichia pastoris* promoter and signal peptide shuffling system. *ACS Synth. Biol.* **2021**, *10* (6), 1360–1372.
- (20) Ebner, K.; Pfeifenberger, L. J.; Rinnofner, C.; Schusterbauer, V.; Glieder, A.; Winkler, M. Discovery and Heterologous Expression of Unspecific Peroxyxygenases. *Catalysts* **2023**, *13* (1), 206.
- (21) König, R.; Kiebist, J.; Kalmbach, J.; Herzog, R.; Schmidtke, K.-U.; Kellner, H.; Ullrich, R.; Jehmlich, N.; Hofrichter, M.; Scheibner, K. Novel unspecific peroxyxygenase from *Truncatella angustata* catalyzes the synthesis of bioactive lipid mediators. *Microorganisms* **2022**, *10* (7), 1267.
- (22) Magliery, T. J. Protein stability: computation, sequence statistics, and new experimental methods. *Curr. Opin. Struct. Biol.* **2015**, *33*, 161–168.
- (23) Goldenzweig, A.; Goldsmith, M.; Hill, S. E.; Gertman, O.; Laurino, P.; Ashani, Y.; Dym, O.; Unger, T.; Albeck, S.; Prilusky, J.; et al. Automated structure-and sequence-based design of proteins for high bacterial expression and stability. *Mol. Cell* **2016**, *63* (2), 337–346.
- (24) Akbulut, N.; Tuzlakoğlu Öztürk, M.; Pijning, T.; İşsевер Öztürk, S.; Gümuşel, F. Improved activity and thermostability of *Bacillus pumilus* lipase by directed evolution. *J. Biotechnol.* **2013**, *164* (1), 123–129.
- (25) Socha, R. D.; Tokuriki, N. Modulating protein stability-directed evolution strategies for improved protein function. *FEBS J.* **2013**, *280* (22), 5582–5595.
- (26) Modarres, H. P.; Mofrad, M.; Sanati-Nezhad, A. Protein thermostability engineering. *RCS Adv.* **2016**, *6* (116), 115252–115270.
- (27) Kulshreshtha, S.; Chaudhary, V.; Goswami, G. K.; Mathur, N. Computational approaches for predicting mutant protein stability. *J. Comput. Aided Mol. Des.* **2016**, *30*, 401–412.
- (28) Goldenzweig, A.; Fleishman, S. J. Principles of protein stability and their application in computational design. *Annu. Rev. Biochem.* **2018**, *87*, 105–129.
- (29) Wijma, H. J.; Fürst, M. J. L. J.; Janssen, D. B. A computational library design protocol for rapid improvement of protein stability: FRESCO. *Protein Eng.* **2018**, *1685*, 69–85.
- (30) Weinstein, J. J.; Goldenzweig, A.; Hoch, S.; Fleishman, S. J. PROSS 2: a new server for the design of stable and highly expressed protein variants. *Bioinformatics* **2021**, *37* (1), 123–125.
- (31) Barber-Zucker, S.; Mindel, V.; Garcia-Ruiz, E.; Weinstein, J. J.; Alcalde, M.; Fleishman, S. J. Stable and functionally diverse versatile peroxidases designed directly from sequences. *J. Am. Chem. Soc.* **2022**, *144* (8), 3564–3571.
- (32) Barber-Zucker, S.; Mateljak, I.; Goldsmith, M.; Kupervaser, M.; Alcalde, M.; Fleishman, S. J. Designed high-redox potential laccases exhibit high functional diversity. *ACS Catal.* **2022**, *12* (21), 13164–13173.
- (33) Ramirez-Escudero, M.; Molina-Espeja, P.; Gomez de Santos, P.; Hofrichter, M.; Sanz-Aparicio, J.; Alcalde, M. Structural insights into the substrate promiscuity of a laboratory-evolved peroxyxygenase. *ACS Chem. Biol.* **2018**, *13* (12), 3259–3268.
- (34) Piontek, K.; Strittmatter, E.; Ullrich, R.; Gröbe, G.; Pecyna, M. J.; Kluge, M.; Scheibner, K.; Hofrichter, M.; Plattner, D. A. Structural basis of substrate conversion in a new aromatic peroxyxygenase: cytochrome P450 functionality with benefits. *J. Biol. Chem.* **2013**, *288* (48), 34767–34776.
- (35) Linde, D.; Santillana, E.; Fernández-Fueyo, E.; González-Benjumea, A.; Carro, J.; Gutiérrez, A.; Martínez, A. T.; Romero, A. Structural characterization of two short unspecific peroxyxygenases: two different dimeric arrangements. *Antioxidants* **2022**, *11* (5), 891.
- (36) Krysztafovy, A.; Schwede, T.; Topf, M.; Fidelis, K.; Moult, J. Critical assessment of methods of protein structure prediction (CASP)—Round XIV. *Proteins* **2021**, *89* (12), 1607–1617.
- (37) Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. *Nature* **2021**, *596* (7873), 583–589.
- (38) Karich, A.; Ullrich, R.; Scheibner, K.; Hofrichter, M. Fungal unspecific peroxyxygenases oxidize the majority of organic EPA priority pollutants. *Front. Microbiol.* **2017**, *8*, 1463.
- (39) Westrick, N. M.; Smith, D. L.; Kabbage, M. Disarming the host: detoxification of plant defense compounds during fungal necrotrophy. *Front. Plant Sci.* **2021**, *12*, 651716.
- (40) Beakes, G. W.; Glockling, S. L.; Sekimoto, S. The evolutionary phylogeny of the oomycete “fungi”. *Protoplasma* **2012**, *249*, 3–19.
- (41) Cavalier-Smith, T.; Chao, E. E. Phylogeny and megasystematics of phagotrophic heterokonts (kingdom Chromista). *J. Mol. Evol.* **2006**, *62*, 388–420.
- (42) Aranda, C.; Ullrich, R.; Kiebist, J.; Scheibner, K.; del Río, J. C.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Selective synthesis of the resveratrol analogue 4, 4'-dihydroxy-trans-stilbene and stilbenoids modification by fungal peroxyxygenases. *Catal. Sci. Technol.* **2018**, *8* (9), 2394–2401.
- (43) Karich, A.; Scheibner, K.; Ullrich, R.; Hofrichter, M. Exploring the catalase activity of unspecific peroxyxygenases and the mechanism of peroxide-dependent heme destruction. *J. Mol. Catal. Chem.* **2016**, *134*, 238–246.
- (44) Aranda, C.; Olmedo, A.; Kiebist, J.; Scheibner, K.; del Río, J. C.; Martínez, A. T.; Gutiérrez, A. Selective epoxidation of fatty acids and fatty acid methyl esters by fungal peroxyxygenases. *ChemCatChem* **2018**, *10* (18), 3964–3968.
- (45) Knorrseidt, A.; Püllmann, P.; Schell, E.; Homann, D.; Freier, E.; Weissenborn, M. J. Identification of novel unspecific peroxyxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach. *ChemCatChem* **2020**, *12* (19), 4788–4795.
- (46) Olmedo, A.; Ullrich, R.; Hofrichter, M.; del Río, J. C.; Martínez, Á. T.; Gutiérrez, A. Novel Fatty Acid Chain-Shortening by Fungal Peroxyxygenases Yielding 2C-Shorter Dicarboxylic Acids. *Antioxidants* **2022**, *11* (4), 744.
- (47) Babot, E. D.; Aranda, C.; Kiebist, J.; Scheibner, K.; Ullrich, R.; Hofrichter, M.; Martínez, A. T.; Gutiérrez, A. Enzymatic epoxidation of long-chain terminal alkenes by fungal peroxyxygenases. *Antioxidants* **2022**, *11* (3), 522.
- (48) Ullrich, R.; Poraj-Kobielska, M.; Scholze, S.; Halbout, C.; Sandvoss, M.; Pecyna, M. J.; Scheibner, K.; Hofrichter, M. Side chain removal from corticosteroids by unspecific peroxyxygenase. *J. Inorg. Biochem.* **2018**, *183*, 84–93.
- (49) Santos-Aberturas, J.; Dörr, M.; Waldo, G. S.; Bornscheuer, U. T. In-depth high-throughput screening of protein engineering libraries by split-GFP direct crude cell extract data normalization. *Chem. Biol.* **2015**, *22* (10), 1406–1414.
- (50) Peleg, Y.; Vincentelli, R.; Collins, B. M.; Chen, K.-E.; Livingstone, E. K.; Weeratunga, S.; Leneva, N.; Guo, Q.; Remans, K.; Perez, K.; et al. Community-wide experimental evaluation of the PROSS stability-design method. *J. Mol. Biol.* **2021**, *433* (13), 166964.
- (51) Knorrseidt, A.; Soler, J.; Hünecke, N.; Püllmann, P.; García-Borràs, M.; Weissenborn, M. J. Simultaneous screening of multiple substrates with an unspecific peroxyxygenase enabled modified alkane

- and alkene oxyfunctionalisations. *Catal. Sci. Technol.* **2021**, *11* (18), 6058–6064.
- (52) Hilberath, T.; van Oosten, R.; Victoria, J.; Brasselet, H.; Alcalde, M.; Woodley, J. M.; Hollmann, F. Toward Kilogram-Scale Peroxygenase-Catalyzed Oxyfunctionalization of Cyclohexane. *Org. Process Res. Dev.* **2023**, *27*, 1384–1389.
- (53) Schmidt, T. G.; Batz, L.; Bonet, L.; Carl, U.; Holzapfel, G.; Kiem, K.; Matulewicz, K.; Niermeier, D.; Schuchardt, I.; Stanar, K. Development of the Twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.* **2013**, *92* (1), 54–61.
- (54) Shental-Bechor, D.; Levy, Y. Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proc. Natl. Acad. Sci. U.S.A.* **2008**, *105* (24), 8256–8261.
- (55) Powell, L. M.; Pain, R. H. Effects of glycosylation on the folding and stability of human, recombinant and cleaved α 1-antitrypsin. *J. Mol. Biol.* **1992**, *224* (1), 241–252.
- (56) Lee, H. S.; Qi, Y.; Im, W. Effects of N-glycosylation on protein conformation and dynamics: Protein Data Bank analysis and molecular dynamics simulation study. *Sci. Rep.* **2015**, *5* (1), 8926.
- (57) Chung, D.; Sarai, N. S.; Knott, B. C.; Hengge, N.; Russell, J. F.; Yarbrough, J. M.; Brunecky, R.; Young, J.; Supekar, N.; Vander Wall, T.; et al. Glycosylation is vital for industrial performance of hyperactive cellulases. *ACS Sustain. Chem. Eng.* **2019**, *7* (5), 4792–4800.
- (58) Pol-Fachin, L.; Siebert, M.; Verli, H.; Saraiva-Pereira, M. L. Glycosylation is crucial for a proper catalytic site organization in human glucocerebrosidase. *Glycoconj. J.* **2016**, *33* (2), 237–244.
- (59) Sun, F.; Zong, W.; Liu, R.; Chai, J.; Liu, Y. Micro-environmental influences on the fluorescence of tryptophan. *SAA* **2010**, *76* (2), 142–145.
- (60) Zhu, B. C.; Laine, R. A.; Barkley, M. D. Intrinsic tryptophan fluorescence measurements suggest that poly lactosaminy glycosylation affects the protein conformation of the gelatin-binding domain from human placental fibronectin. *Eur. J. Biochem.* **1990**, *189* (3), 509–516.
- (61) Tran, A.-M.; Nguyen, T.-T.; Nguyen, C.-T.; Huynh-Thi, X.-M.; Nguyen, C.-T.; Trinh, M.-T.; Tran, L.-T.; Cartwright, S. P.; Bill, R. M.; Tran-Van, H. *Pichia pastoris* versus *Saccharomyces cerevisiae*: a case study on the recombinant production of human granulocyte-macrophage colony-stimulating factor. *BMC Res. Notes* **2017**, *10*, 148.
- (62) Kwon, K.-S.; Yu, M.-H. Effect of glycosylation on the stability of α 1-antitrypsin toward urea denaturation and thermal deactivation. *Biochim. Biophys. Acta* **1997**, *1335* (3), 265–272.
- (63) Høiberg-Nielsen, R.; Westh, P.; Skov, L.; Arleth, L. Interrelationship of steric stabilization and self-crowding of a glycosylated protein. *Biophys. J.* **2009**, *97* (5), 1445–1453.
- (64) Kayser, V.; Chennamsetty, N.; Voynov, V.; Forrer, K.; Helk, B.; Trout, B. L. Glycosylation influences on the aggregation propensity of therapeutic monoclonal antibodies. *Biotechnol. J.* **2011**, *6* (1), 38–44.
- (65) Vieira Gomes, A. M.; Souza Carmo, T.; Silva Carvalho, L.; Mendonça Bahia, F.; Parachin, N. S. Comparison of yeasts as hosts for recombinant protein production. *Microorganisms* **2018**, *6* (2), 38.
- (66) Miton, C. M.; Tokuriki, N. How mutational epistasis impairs predictability in protein evolution and design. *Protein Sci.* **2016**, *25* (7), 1260–1272.
- (67) Lipsh-Sokolik, R.; Khersonsky, O.; Schröder, S. P.; de Boer, C.; Hoch, S.-Y.; Davies, G. J.; Overkleft, H. S.; Fleishman, S. J. Combinatorial assembly and design of enzymes. *Science* **2023**, *379* (6628), 195–201.
- (68) Weinstein, J. Y.; Martí-Gómez, C.; Lipsh-Sokolik, R.; Hoch, S. Y.; Liebermann, D.; Nevo, R.; Weissman, H.; Petrovich-Kopitman, E.; Margulies, D.; Ivankov, D.; et al. Designed active-site library reveals thousands of functional GFP variants. *Nat. Commun.* **2023**, *14* (1), 2890.
- (69) Breen, M. S.; Kemeny, C.; Vlasov, P. K.; Notredame, C.; Kondrashov, F. A. Epistasis as the primary factor in molecular evolution. *Nature* **2012**, *490* (7421), 535–538.
- (70) Krishna, R.; Wang, J.; Ahern, W.; Sturmels, P.; Venkatesh, P.; Kalvet, I.; Lee, G. R.; Morey-Burrows, F. S.; Anishchenko, I.; Humphreys, I. R.; et al. Generalized biomolecular modeling and design with RoseTTAFold All-Atom. *bioRxiv* **2023**, 2023.10.09.561603.



CAS BIOFINDER DISCOVERY PLATFORM™

CAS BIOFINDER HELPS YOU FIND YOUR NEXT BREAKTHROUGH FASTER

Navigate pathways, targets, and diseases with precision

Explore CAS BioFinder