



Supported by



## Accepted Article

**Title:** Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach

**Authors:** Anja Knorrscheidt, Pascal Püllmann, Eugen Schell, Dominik Homann, Erik Freier, and Martin J. Weissenborn

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). This work is currently citable by using the Digital Object Identifier (DOI) given below. The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

**To be cited as:** *ChemCatChem* 10.1002/cctc.202000618

**Link to VoR:** <https://doi.org/10.1002/cctc.202000618>

## FULL PAPER

# Identification of novel unspecific peroxygenase chimeras and unusual YfeX axial heme ligand by a versatile high-throughput GC-MS approach

Anja Knorrscheidt,<sup>[a]</sup> Pascal Püllmann,<sup>[a]</sup> Eugen Schell,<sup>[a]</sup> Dominik Homann,<sup>[a]</sup> Erik Freier,<sup>[b]</sup> and Martin J. Weissenborn\*<sup>[a, c]</sup>

<sup>[a]</sup> Anja Knorrscheidt, Pascal Püllmann, Eugen Schell, Dominik Homann, Jun.-Prof. Martin J. Weissenborn  
Bioorganic Chemistry  
Leibniz Institute of Plant Biochemistry  
Weinberg 3, 06120 Halle (Saale), Germany  
E-mail: martin.weissenborn@ipb-halle.de

<sup>[b]</sup> Dr. Erik Freier  
CARS Microscopy  
Leibniz-Institut für Analytische Wissenschaften -ISAS- e.V.  
Otto-Hahn-Str. 6b, 4227 Dortmund, Germany

<sup>[c]</sup> Jun.-Prof. Martin J. Weissenborn  
Institute of Chemistry  
Martin Luther University Halle-Wittenberg  
Kurt-Mothes-Str. 2, 06120 Halle (Saale), Germany

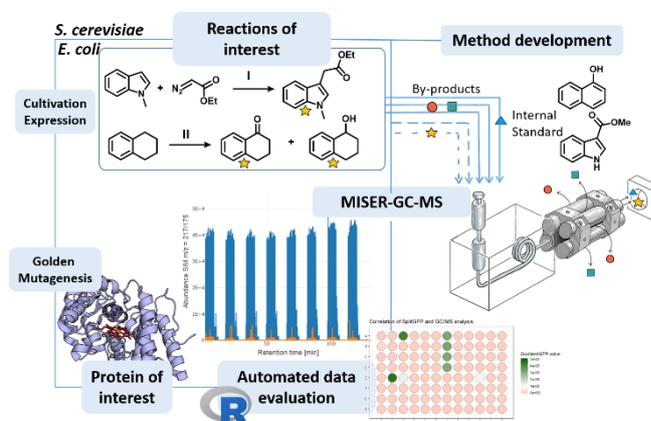
Supporting information for this article is given via a link at the end of the document.

**Abstract:** Catalyst discovery and development requires the screening of large reaction sets necessitating analytic methods with the potential for high-throughput screening. These techniques often suffer from substrate dependency or the requirement of expert knowledge. Chromatographic techniques (GC/LC) can overcome these limitations but are generally hampered by long analysis time or the need for special equipment. The herein developed multiple injections in a single experimental run (MISER) GC-MS technique allows a substrate independent 96-well microtiter plate analysis within 60 min. This method can be applied to any laboratory equipped with a standard GC-MS. With this concept novel, unspecific peroxygenase (UPO) chimeras, could be identified, consisting of subdomains from three different fungal UPO genes. The GC-technique was additionally applied to evaluate an YfeX library in an *E. coli* whole-cell system for the carbene-transfer reaction on indole, which revealed the thus far unknown axial heme ligand tryptophan.

## Introduction

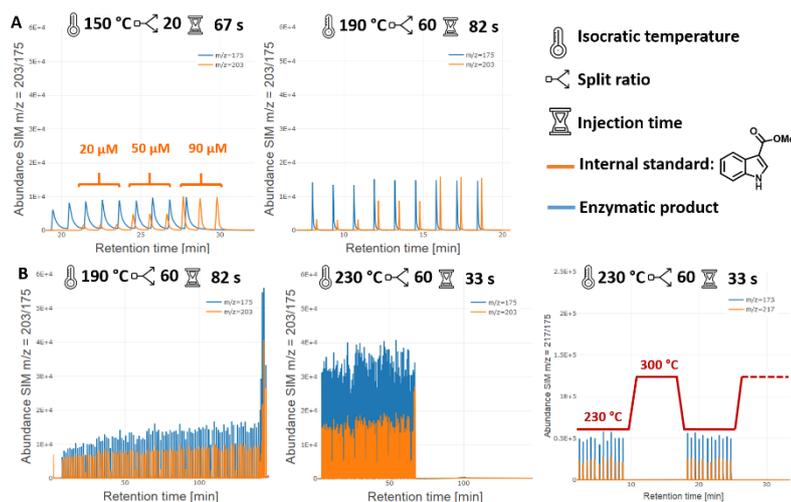
In the last decades, highly successful and environmentally benign catalytic methodologies have been developed in organic chemistry. The discovery and development of these catalysts have been primarily based on rational or intuitional approaches and serendipitous findings.<sup>[1]</sup> Improving the outcome of unexpected findings necessitates high-throughput experimentation that enables the analysis of several thousand reactions. Several smart strategies were developed like DNA templating,<sup>[2]</sup> sandwich immunoassay,<sup>[3]</sup> MALDI labelling,<sup>[4]</sup> fluorescence quenching and UV absorption.<sup>[5]</sup> These techniques have their distinct advantages but typically rely on expert knowledge, specialised laboratories or labelled substrates. For protein engineering, most setups rely on colourimetric or fluorescence based assays amongst them also microfluidic-based ultra-high throughput systems.<sup>[6]</sup> The ideal assay would i)

allow the screening with the exact substrate of interest, ii) be highly sensitive, iii) be exceedingly reproducible and iv) require minimal time-periods for analysis. To be applicable in standard chemical and biochemical laboratories, the necessary instrumentation should be sufficiently general and accessible. A flexible and sensitive analytical technique is provided by chromatographies such as liquid (LC) or gas chromatography (GC). While these techniques often provide the necessary sensitivity and are applicable to a wide range of substrates, they suffer from long analysis times preventing a high-throughput screening with several hundreds of samples a day. The time consuming parts are the separation, washing/heating and column equilibration. Higher throughput was enabled by method, column and gradient optimisation yielding total run times of less than four minutes.<sup>[7]</sup>



**Figure 1.** MISER-GC-MS and its implementation for the identification of novel enzymatic activities enabled the investigation of two case studies in different biological matrices: I) carbene-transfer reaction on 1-methyl indole (*E. coli*) and II) hydroxylation of 1,2,3,4-tetrahydronaphthalene (*S. cerevisiae*).

## FULL PAPER



**Figure 2.** A) Method development with the analytes ethyl 3-indoleacetate ( $m/z$  203) and methyl indole-3-carboxylate ( $m/z$  175) with three different concentrations of ethyl 3-indoleacetate (20, 50 and 90  $\mu\text{M}$ ). B) Transferring the best conditions into 96-well format using a biological matrix (*E. coli* lysate). An alternative stacked method, including intermediate heating cycles, was developed with 3-(*N*-methyl-indole)acetate ( $m/z$  217).

Intriguing developments have brought substantially shortened analysis times for LC by ultrahigh performance liquid chromatography (UHPLC) and GC by flow field thermal gradient gas chromatography (FF-TG-GC).<sup>[8]</sup> However, these techniques are not available in every laboratory (UHPLC), only a few prototypes are existing (FF-TG-GC), or a higher throughput cannot be achieved with conventional approaches (GC/LC). The analysis of multiple, overlapping samples in one chromatogram, coined multiplexing, has been introduced in 1967 by Izawa for GC<sup>[9]</sup> and was extended by Trapp and co-workers for a high-throughput approach. This technique requires a specifically built injector system. A method to enhance throughput in standard LC measurements is the multiple injections in a single experimental run (MISER) approach.<sup>[10]</sup> This method does not require special equipment or expert knowledge. MISER relies on the injection of several samples under isocratic conditions into one chromatographical run requiring baseline separation of the peaks. This enabled the performance of long chromatographic separations in a high-throughput manner as usually signal and therefore information free areas in the chromatogram are filled by peaks of multiple injections.<sup>[10-11]</sup> The method of MISER-GC-MS has not yet been employed for large sample numbers,<sup>[12]</sup> but seems highly suitable for the application in a high-throughput screening based on several reasons:

- MISER allows a distinct throughput enhancement for GC without requiring special equipment,
- the quantification of target substances by GC-MS extracted from crude mixtures is far less prone to ion suppression by matrix effects than by LC-MS systems and,<sup>[13]</sup>
- MS-based detection is highly sensitive and allows quantification of low analyte concentrations in complex mixtures.

Herein, we report the development of a novel MISER-GC-MS strategy paving the way for a versatile, assay-independent platform for catalytic reactions. The developed system is applicable for any GC-MS equipped with an autosampler. Thereby, the MISER-GC-MS technique was broadened to measure several analytes combined with an internal standard to compensate deviations and enables the quantification of multiple molecules with overlapping peak areas eliminating the need for chromatographical separation. To facilitate the data evaluation, an R-script has been written, allowing rapid analysis and quality control. The developed technique enabled the screening of two libraries with different host organisms and reactions.

## Results and Discussion

### Development of a versatile MISER-GC-MS approach for highly reproducible 96-well analysis in biological matrices

A standard autosampler was modified for fast injection application by decoupling of the autosampler from the GC instrument (Supporting Information). The read-out-signal of the GC was suppressed, enabling an independent control of the autosampler. The method development commenced by altering various conditions at the autosampler. The setup optimisations like post-cleaning with isopropanol after each injection, as well as variations in filling speed and the number of filling strokes, were performed using the analyte ethyl 3-indoleacetate. This led to significantly decreased standard deviations from initially 51.6 to 6.5 % (Table S2) illustrating the importance of adjusting the settings of the autosampler for multiple injections set up. Since the MS-detection allows the simultaneous quantification of different  $m/z$  signals, the system was expanded to an internal standard (methyl indole-3-carboxylate). The correlation between the analyte and an internal standard enabled the error minimisation occurring due to solvent evaporation, extraction and sample injection and allows the comparability of microtiter plates. For MISER-GC-MS method verification, three different ethyl 3-indoleacetate concentrations (20, 50 and 90  $\mu\text{M}$ ) were employed with a constant concentration of internal standard. All MISER runs ought to be isocratic. The lowest oven temperature of 150 °C led to peak tailing and poor baseline separation (Figure 2A, left). Utilising a temperature of 190 °C resulted in an excellent baseline separation and improved peak shapes as well as a reduced standard deviation of 4.0 % using 20  $\mu\text{M}$  ethyl 3-indoleacetate (Table S3).

Further parameters, which were investigated, were the split ratio, the MS mode (SIM or Scan) and the injection interval (Figure 2, Figure S6). To avoid overlapping of the injection and analyte peaks, the injection interval was altered from 67 s to 82 s. This further lowered the standard deviation and allowed the increase of the split ratio up to 60, resulting in excellent standard deviations of 1.0 % for 20  $\mu\text{M}$  ethyl 3-indoleacetate. With these

## FULL PAPER

optimised analytical conditions in hand, we further challenged the system by using *E. coli* cell lysates spiked with ethyl 3-indoleacetate and analysed the resulting samples in 96-well experiments using methyl indole-3-carboxylate as an internal standard. As further quality control and for calibration purposes ethyl 3-indoleacetate (20  $\mu\text{M}$ ) extracted from a buffer system was injected after each microtiter plate row (12 samples) as well as standards at the end of the run. The system was assessed with the best conditions of the GC system (190  $^{\circ}\text{C}$ , split ratio 60, and 82 s injection interval) leading to a standard deviation of 4.0 % for 109 injections including controls (Figure 2B, left). This could be even further improved to a standard deviation of 2.5 % by increasing the oven temperature to 230  $^{\circ}\text{C}$ . Due to shorter retention times on the column, this also allowed a quicker injection interval of 33 s (Figure 2B, middle, Table S6).

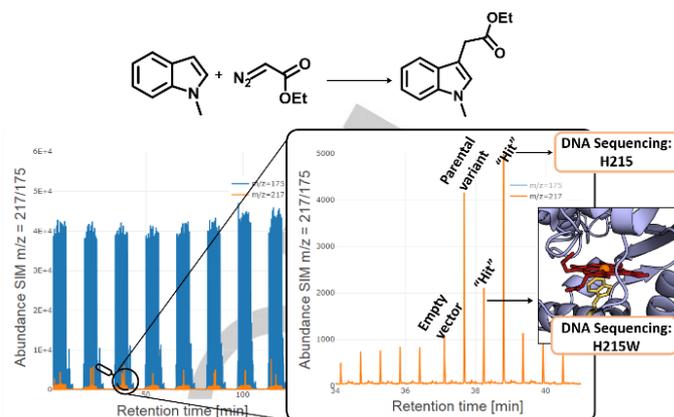
Since the developed methods shall be applicable for the screening of non-natural enzyme activities, which in general suffer from low turnovers and require high substrate loading, the contamination of the GC column could pose a substantial problem within 96 injections. Heating intervals can remove such contaminations.<sup>[12b]</sup> An additional method was therefore developed, which includes a heating cycle after every 12<sup>th</sup> sample (Figure 2B, right). With a split ratio of 60 a standard deviation of 3.0 % was reached. These different techniques and setups were applied to the screening of enzyme libraries (see below). Since the assessment of hundreds of samples results in large amounts of data points, an automated R-script was written to ensure high quality and correct peak integration in each microtiter plate (see Additional Material). The script assesses the data of the MISERgram for reproducibility using the internal standard. Based on the injection interval, the peaks are counted to verify whether every sample has been correctly injected (GC hardware) and integrated (GC software). The quotient of the internal standard and product peak is illustrated as a bar chart and colour-coded microtiter plate for fast data evaluation.

The new MISER-GC-MS system was now applied to two different enzyme systems to screen protein libraries in altering biological systems.

### Screening of a focussed enzyme library of the YfeX-catalysed carbene-transfer reaction revealed tryptophan as a novel axial heme ligand

The dye-decolourising peroxidase YfeX from *E. coli* was previously shown to perform non-natural carbene-transfer reactions<sup>[14]</sup> such as carbonyl olefination<sup>[15]</sup> and C—H functionalisation.<sup>[16]</sup> The starting activity for the latter reaction was previously improved by an alanine scan within the active site.<sup>[16]</sup>

We were subsequently interested in the influence of the axial ligand on the activity of YfeX regarding the C—H functionalisation reaction. The axial ligand complexes the heme iron and substantially influences its redox potential and electrophilicity and hence the overall activity of the occurring heme-carbenoid complex.<sup>[17]</sup> The starting point of the mutagenesis was a variant carrying the mutations D143V, S234C and F248V (parental variant). To study the influence of the axial ligand in YfeX, we performed saturation mutagenesis targeting the axial ligand residue histidine 215—using the recently developed Golden Mutagenesis protocol and its online



**Figure 3.** 96-well analysis of the carbene-transfer reaction by the previously developed stacked method for MISER-GC-MS. As controls, the empty vector was included in column 6 and the parental YfeX variant in column 7, which harbours the axial ligand H215.

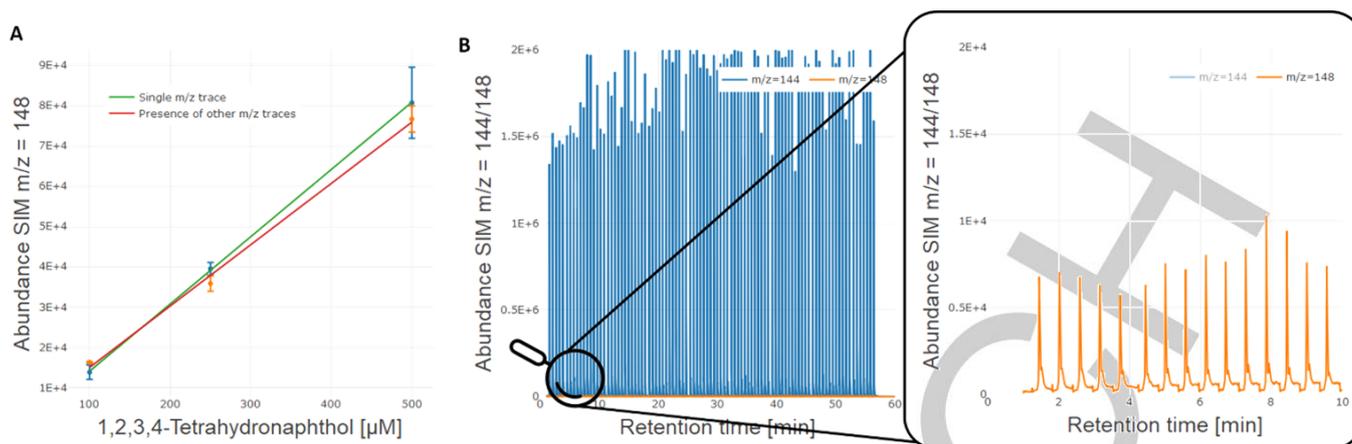
tool<sup>[18]</sup>—and screened the resulting library for the occurrence of other functional axial ligands in whole-cell reactions by using the interval heating method for MISER-GC-MS. To ensure the quality of the generated library, a quick quality control (QQC) was conducted, demonstrating the expected codon distribution (Figure S11). The reaction was performed using the carbene-transfer reaction on 1-methyl-indole with ethyl diazoacetate as carbene donor (Supporting Information).

As a control, cells harbouring the empty plasmid (pAGM22082) and the parental YfeX were included within the 96 well plate, which could be clearly distinguished by MISER-GC-MS (Figure 3). To our delight, a new variant was identified, which carried a highly unusual tryptophan residue as axial heme ligand. To prove that the MISERgram revealed a “true positive” result, the corresponding plasmid was freshly transformed, expressed, and the corresponding protein purified by metal affinity chromatography. These results confirmed the YfeX-H215W variant, showing only slightly reduced activities compared to the parental variant (Figure S12). This discovery could broaden the spectrum of canonical amino acids as axial ligands for non-natural reactions<sup>[17c, 19]</sup> as well as represent an interesting target structure for the further development of non-canonical amino acids for heme complexing.<sup>[17b, 20]</sup>

### Screening of a fungal unspecific peroxygenase (UPO) chimera library in *S. cerevisiae* with the substrate tetralin revealing six novel peroxygenase constructs

To demonstrate that the MISER-GC-MS method can be readily applied to other reactions and environments, the screening was applied to fungal unspecific peroxygenases<sup>[21]</sup> (UPOs) for its hydroxylation of tetralin. The previous method development demonstrated that simultaneous quantification of product and an internal standard is possible. We here wanted to expand the analysis to three different molecules: tetrahydronaphthol (main reaction product),  $\alpha$ -tetralone (side-product) and 1-naphthol (internal standard). Method development was based on the results, as stated above. The two products and the internal standard were injected in three different concentrations, and the MS response was compared to the obtained values when injecting only one analyte at a time and all three simultaneously. The results revealed minimal deviation comparing single and

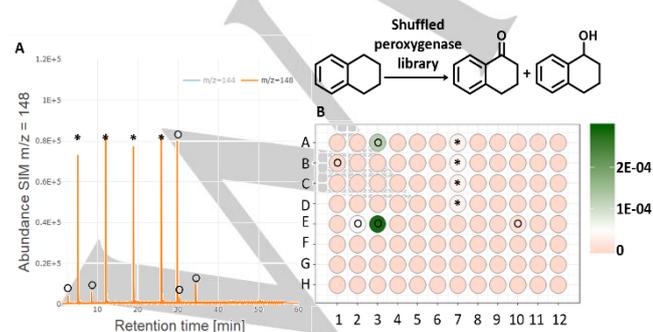
## FULL PAPER



**Figure 4.** A) Extraction and mass spectrometric analysis of 1,2,3,4-tetrahydronaphthol in presence and absence of the other two compounds within a MISER experiment, B) MISERgram of the hydroxylation of 1,2,3,4-tetrahydronaphthalene with 96 biological replicates in microtiter plate format with a standard deviation of 9.7 %.

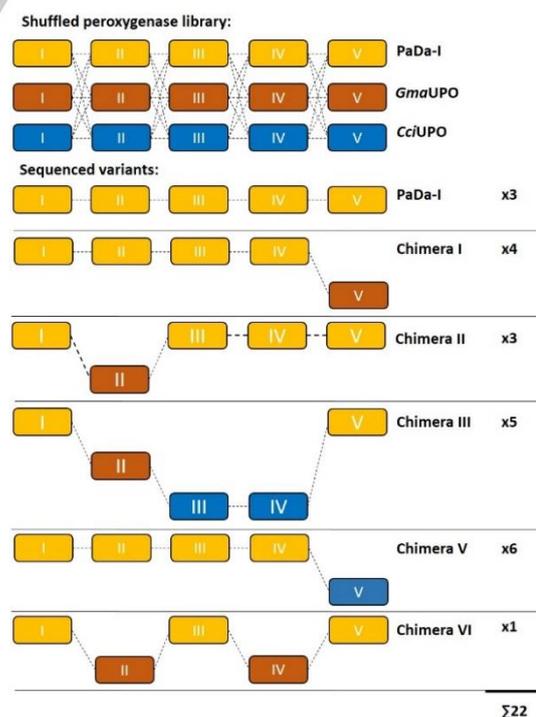
multiple  $m/z$  trace analysis (Figure 4A, Figure S4). To confirm the accuracy of the refined MISER-GC-MS method, we selected one possible functional variant (chimera I, Figure 4B), which was previously identified by a standard colourimetric assay (unpublished results). The aim was to validate a MISER-GC-MS technology enabling the analysis of an entire microtiter plate with individually expressed variants with an overall standard deviation of less than 10 %. We were delighted to see that the measurement of the entire 96 well plate within an analysis time of 60 minutes showed a standard deviation of only 9.7 % for the formation of tetrahydronaphthol (Figure 4B).

UPOs represent a class of highly promising enzymes, which show remarkable activities as well as stabilities and solely consume hydrogen peroxide as co-substrate. A major bottleneck, however, is the heterologous enzyme expression. Even though several thousands of putative peroxygenases have been assigned, only very few could be produced heterologously thus far.<sup>[21b, 22]</sup> To create structural diversity three (putative) UPO genes from different fungal origins bearing high sequence similarity were selected for the construction of a shuffled peroxygenase library: the yeast secretion variant PaDa-I originating from *Agrocybe aegerita*, *GmaUPO* from *Galerina marginata* (72 % identity) and *CciUPO* from *Coprinopsis cinerea* (62 % identity).<sup>[21b, 22]</sup> The wildtype enzymes *GmaUPO* and *CciUPO* showed no activity and no expression in *S. cerevisiae*, respectively. The secondary structural units were grouped, and sequence subunits were created by loop cuts yielding five subunits for each gene (Figure 6). The structural assignment was done based on the crystal structure of PaDa-I (pdb: 5OXU).<sup>[23]</sup>



**Figure 5.** Screening of a fungal unspecific peroxygenase library. A) MISERgram of microtiter plate 1, B) automated R-script evaluation of plate 1 (○=identified hit, \* = PaDa-I as control).

The secondary structure consists of 13 helices (42 % of overall sequence) and 15 beta sheets (6 %). The units were then grouped based on not disrupting pivotal catalytic motives (PCP; EGD; E196) and secondary structure elements such as alpha helices and beta sheets (see Figure S15 for details). These subunits were randomly shuffled, leading to  $3^5 = 243$  possible combinations. Cultivation, as well as biotransformation, were implemented in a high-throughput manner using 96 well microtiter plates. Besides the analytics of the product formation, the peroxygenases were equipped with a C-terminal split-GFP-tag thus allowing the determination of the protein concentration.<sup>[24]</sup> The screening of 672 transformants by MISER-GC-MS was done in seven hours and revealed 34 hits. Figure 5 depicts a resulting MISERgram of one microtiter plate. The best performing variants from this library were reproduced in four individual biological replicates in a microtiter plate, confirming the accuracy of this method (Figure S14).



**Figure 6.** Subunits of the shuffled peroxygenase library and repetitive identified hits which were analysed by MISER-GC-MS and subsequently sequenced.

## FULL PAPER

From the previously identified chimeras I-V (unpublished results), only the chimera I, II, III and V showed activity towards tetralin and could be identified during the MISER-GC-MS screening. However, a novel construct Chimera VI could be identified, and the parental variant PaDa-I was rediscovered three times. Based on the GFP signal, the chimera VI showed a 3.8 fold enhanced secretion compared to PaDa-I (Supporting Figure S14B) and demonstrated a good starting point for a heterologously expressed UPO, which can be screened on additional substrates.

## Conclusion

The development of MISER-GC-MS methods was successfully utilised for its identification of novel enzyme activities. Two methods were developed: one method for the screening of natural enzymatic reactions, which generally have only a few side-products and therefore an injection interval of 33 s was applied for 96 well microtiter plate analysis within 60 min. The other developed technique was the stacked method with particular relevance to low activities and large amounts of side-products, including a heating step after every 12<sup>th</sup> injection to improve the quality of the acquired data.

Both systems were employed for the screening of two different enzyme classes and reactivities within two biological systems. For the YfeX catalysed carbene-transfer reaction, the screening of a focused library led to a highly unusual axial heme ligand (H215W). By screening a chimera library of three shuffled unspecific peroxygenase genes, MISER-GC-MS aided the detection of five peroxygenase chimeras. These chimeras were active and heterologous producible catalysts and hence provide access to new peroxygenase scaffolds.

To put the MISER-GC-MS in perspective to the colourimetric assay based on 4-nitrocatechol formation: the 96-well analysis of the colourimetric product 4-nitrocatechol takes 2 minutes. It has a sensitivity in the range of 10-20  $\mu\text{M}$ , which amounts to 2000-4000 pmol (200  $\mu\text{l}$  volume in one microtiter plate well). MISER-GC-MS revealed an analysis time of 60 minutes with a detection limit of 5-10  $\mu\text{M}$  corresponding to 0.08-0.17 pmol (1  $\mu\text{L}$  sample injection and a split ratio of 1:60). The high sensitivity of MISER-GC-MS could enable the engineering of minute initial activities for new unnatural reactions in the future.

The herein demonstrated MISER-GC-MS technology can be implemented into any laboratory with a GC-MS equipped with an autosampler and has proven potential as a versatile, specific and cost-effective high-throughput approach.

## Experimental Section

### Biological procedures

**Lysate preparation for extraction experiments.** The YfeX WT gene (cloned into pAGM22082) was transformed into chemically competent *E. coli* BL21 (DE3) pLysS cells (Merck Millipore, Darmstadt, DE) by heat shock. Freshly-plated transformants were grown overnight in 5 mL TB medium containing kanamycin and chloramphenicol (50  $\mu\text{g}/\text{mL}$  each). 2 mL of the preculture was then used to inoculate 400 mL of main culture consisting of TB autoinduction medium containing kanamycin and chloramphenicol (50  $\mu\text{g}/\text{mL}$  each). Cells were incubated at 37 °C and 120 rpm. After 4 h of initial cultivation, aqueous solutions of  $\text{FeCl}_3/5$ -aminolevulinic acid (final concentration: 100  $\mu\text{M}$ ) were added, the temperature was reduced to 30 °C, and the cells were incubated for further 16.5 h. Cells were harvested by centrifugation (3000 x g, 20 min, 4 °C). The supernatant was discarded, and the pellet was resuspended in binding buffer (50 mM KPi, pH 7.0, 200 mM NaCl). Cells were lysed by sonication (Bandelin Sonoplus HD3100: 6 x 30 s, 70 % amplitude, pulse mode). The resulting lysate was stored at -20 °C until further utilisation for lysate extraction experiments.

**Site-saturation mutagenesis (SSM) and *E. coli* cultivation in 96-deep-well plates.** Mutagenesis was performed using the Golden Mutagenesis technique<sup>[19]</sup> combined with the “22c-trick”<sup>[25]</sup> for randomisation. The YfeX gene from *E. coli* was chosen as a template, targeting amino acid residue position Histidine 215. The created library was then transformed into chemically competent *E. coli* BL21 (DE3) pLysS cells. A preculture of the transformants was grown in 350  $\mu\text{L}$  TB medium with added kanamycin and chloramphenicol (50  $\mu\text{g}/\text{mL}$  each) in an EnzyScreen plate at 37 °C and 300 rpm overnight. For the main culture, 730  $\mu\text{L}$  of TB autoinduction medium per well (plus 50  $\mu\text{g}/\text{mL}$  kanamycin and chloramphenicol and 100  $\mu\text{M}$  of  $\text{FeCl}_3/5$ -aminolevulinic acid) was inoculated with 20  $\mu\text{L}$  of the respective preculture. In a first phase the cells were cultivated at 37 °C and 300 rpm for 4 h. Afterwards the temperature was decreased to 25 °C and protein expression was continued overnight. The cells were harvested by centrifugation (30 min, 3000 g, 4 °C) and the supernatant discarded.

**Whole-cell biotransformation with YfeX.** The 96 deep well plate harbouring the cell pellets was transferred into a glove box ( $\text{N}_2$  atmosphere) and incubated on ice for 1 h to remove residual oxygen from the gas phase. 200  $\mu\text{L}$  of degassed 50 mM KPi buffer (pH 7.0, 200 mM NaCl, 2 mM  $\text{MgCl}_2$ ) was added to each well, and the cells were resuspended by vortexing for 1 min. 200  $\mu\text{L}$  of a reaction master mix (stock solutions: 50 mM 1-methylindole, 50 mM ethyl diazoacetate, 100 mM sodium dithionite) were added to each well (final concentration: 2.5 mM 1-methylindole, 2.5 mM ethyl diazoacetate, 10 mM sodium dithionite and 20 % ethanol as co-solvent) and the plate was closed tightly with a cover. The reaction was performed at 30 °C and 300 rpm for 1 h. The samples were extracted by addition of 1 mL ethyl acetate (EtOAc) containing 100  $\mu\text{M}$  methyl indole-3-carboxylate as internal standard and shaking at 300 rpm for 20 min at 25 °C. After centrifugation (10 min, 3000 g, 10 °C) 400  $\mu\text{L}$  of the organic phase were transferred into a glass-coated 96-well plate for GC-MS analysis.

## FULL PAPER

**Expression and purification of YfeX variants for hit verification.** The YfeX gene and its corresponding mutants (plasmid backbone: pAGM22082) were transformed into chemically competent *E. coli* BL21(DE3) pLysS cells by heat shock procedure. Freshly-plated transformants were grown overnight (160 rpm) as preculture in 5 ml TB medium containing chloramphenicol and kanamycin (50 µg/mL each). 2 ml of the precultures were used to respectively inoculate 400 ml TB autoinduction medium (+ kanamycin and chloramphenicol). Cells were incubated at 37 °C (120 rpm shaking). After 4 hours of cultivation, aqueous solutions of FeCl<sub>3</sub> and 5-aminolevulinic acid (final concentration: 100 µM) were added and the temperature reduced to 30 °C. The cells were incubated for further 16.5 h. Cells were finally harvested by centrifugation (3000 x g, 20 min, 4 °C). The cultivation supernatant was discarded, and the pellets were resuspended in binding buffer (50 mM KPi; pH = 7.4, 200 mM NaCl, 1 mg/ml lysozyme, 100 µg/ml DNase I). Cells were lysed by sonication (Bandelin Sonoplus HD3100: 6x30 s, 70% amplitude, pulse mode). The cell debris was removed by centrifugation for 45 min at 4 °C and 6000 x g. The proteins exhibiting an N-terminal attached hexahistidine-Tag were purified by IMAC (immobilised metal ion affinity chromatography) using 1 ml His GraviTrap TALON columns (GE Healthcare Europe GmbH, Freiburg, DE). After applying the cleared supernatant to the column, the column was washed with 10 column volumes (10 ml) of washing buffer (50 mM KPi; pH = 7.4, 200 mM NaCl, 5 mM imidazole). YfeX variants were finally eluted by the addition of elution buffer (50 mM KPi; pH = 7.4, 200 mM NaCl and 250 mM imidazole). For subsequent buffer exchange of the pooled elution fractions PD-10 desalting columns (GE Healthcare Europe GmbH, Freiburg, DE) were used according to the manufacturer's protocol. YfeX samples were eluted using 50 mM KPi (pH = 7.0) with addition of 10 % glycerol (v/v) as storage buffer. Afterwards samples were flash frozen in liquid nitrogen and stored at -20 °C. Protein concentrations were determined in duplicates using the Pierce TM BCA Protein Assay Kit (Thermo Scientific, Rockford, US).

**Purified enzyme biotransformation.** For the purified enzyme biotransformation the whole cell biotransformation conditions were adapted and the final enzyme concentration was set to 15 µM. The reaction was performed at 30 °C and 600 rpm for 1 h. The samples were extracted by addition of 1 mL EtOAc containing 100 µM methyl indole-3-carboxylate as internal standard and vortexing. After centrifugation (5 min, 20.000 g) 600 µL of the organic phase were transferred into glass vials for subsequent GC-MS analysis.

**Amino acid sequence of the YfeX parental variant.** Plasmid derived N-terminal hexahistidine-tag + T7 tag indicated in *italic*. The C-terminal GFP11 detection tag is underlined. Mutations in comparison to the YfeX wild type protein sequence (UniProt ID: P76536) are indicated as **bold** letters.

*MGSSHHHHHSSGLVPRGSHMASMTGGQQMGRDGMSSQV  
QSGILPEHCRAAIWIEANVKGEVDALRAASKTFADKLATFEAK  
FPDAHLGAVVAFGNNTWRALSGGVGAELKDFPGYKGLAP  
TTQFDVLIHILSLRHVDVNFSAQAAMEAFGDCIEVKKEIHGFR  
VWEERDLSGFVVGTEHPAGEETRREVAVIKDGVDAGGSYVF  
VQRWEHNLKQLNRMSVHDQEMVIGRTKEANEIDGDERPET  
SHLTRVDLKEDGKGLKIVRQCCLPYGTASGTHGLYVCAYCARL*

HNIEQQLLSMFGD TDGKRDAMLRFTKPV TGGYYFAPSLDKL  
MALSDGSGGGSTSRDHMVLHEYVNAAGIT

**Yeast supernatant preparation for extraction experiments.**

The empty plasmid (lacking UPO gene) for yeast expression was transformed into chemically competent yeast cells (INVSc1 strain) by polyethylene glycol/lithium acetate transformation. For the preculture 100 ml SC Drop selection Media (lacking Uracil as supplement; containing 2 % raffinose as carbon source and 25 µg/ml chloramphenicol) were inoculated with a single yeast colony from a previously grown SC Drop out plate (lacking Uracil) at 30 °C and 130 rpm for 2 days. For expression rich expression media (Yeast extract Peptone Galactose) containing 2 % of galactose as inducer was utilised. The final main culture OD was adjusted by addition of the preculture to 0.3. Expression was performed for further 72 h (120 rpm; 30 °C). After 72 h cultivation time cells were separated from the supernatant by centrifugation (3400 rpm; 45 min; 4 °C). The supernatant was stored at 4 °C until further utilisation for supernatant extraction experiments.

**Generation of a shuffled peroxygenase gene library and Yeast cultivation in 96-well plates.**

The gene of the peroxygenase yeast secretion mutant PaDa-I,<sup>[21b]</sup> an annotated peroxygenase gene from the fungus *Galerina marginata* and previously described peroxygenase from *Coprinopsis cinerea*<sup>[22]</sup> were divided into 5 structural subunits and randomly shuffled together (243 possible combinations). Full length fragments were then reassembled into a yeast expression plasmid (Galactose inducible promoter) with an N-terminal signal peptide and a C-terminal GFP 11 detection tag within a single Golden Gate cloning reaction (unpublished results). Corresponding plasmid mixtures were transformed into yeast cells (INVSc1 strain) by polyethylene glycol/lithium acetate transformation. Yeast cells were cultivated in liquid culture as described by Molina Espeja *et al.*<sup>[21b]</sup> with slight modifications. 220 µl of minimal expression medium per well (containing 2 % (w/v) Galactose final concentration as carbon source and inducer) were inoculated with a single yeast colony from a previously grown SC Drop out plate (lacking Uracil). For cultivation EnzyScreen half-deepwell plates were utilised. Expression was performed for 72 h (230 rpm; 30 °C). After 72 h cultivation time cells were separated from the peroxygenase containing supernatant by centrifugation (3400 rpm; 45 min; 4 °C). For subsequent screening 20 µL (for split GFP Assay) and 100 µL (for biotransformation) were transferred to a respective plate using a multichannel pipet.

**Supernatant biotransformation in *S. cerevisiae* for peroxygenases.**

A volume of 100 µL of the peroxygenase containing yeast supernatant was transferred to a 96-well EnzyScreen plate (CR1496), followed by the addition of 240 µl of 100 mM citrate buffer (pH 6), 40 µl of 1,2,3,4-tetrahydronaphthalene stock solution (10 mM in acetonitrile, final concentration: 1 mM) and 20 µl H<sub>2</sub>O<sub>2</sub> stock solution (6 mM in H<sub>2</sub>O, final concentration: 300 µM). After a centrifugation step (1 min, 100 x g, 10 °C), the reaction was shaken for 16 h at 230 rpm and 30 °C. The reaction was stopped by addition of 400 µl freshly prepared 500 µM internal standard extraction solution (1-naphthol in EtOAc). For extraction the aqueous and organic

## FULL PAPER

solution were shaken for additional 20 min at 300 rpm. After centrifugation (3000 x g, 5 min, 10 °C), 300 µl of the organic layer was transferred to a glass coated plate for GC analysis.

**Split-GFP assay.** Protein normalisation was performed employing the principle of a split GFP normalisation assay as described by Santos-Aberturas *et al.*<sup>[24b]</sup> with slight modifications. The complementation fragment sfGFP 1-10 was cloned into the Golden Mutagenesis plasmid pAGM22082\_cRed<sup>2</sup> for T7 promoter controlled expression in *E. coli*. For measurement 20 µl of yeast expression supernatant was transferred to a previously BSA blocked 96 well Nunc MaxiSorp Fluorescence plate (ThermoFisherScientific, Waltham, US) and 180 µl of sfGFP 1-10 inclusion body preparation added. Immediate fluorescence values (GFP fluorophore: excitation wavelength: 485 nm; emission wavelength: 535 nm; top read mode) was measured using a 96 well plate fluorescence reader (TECAN, Grödig, AT). After storage over two nights (at 4 °C) final fluorescence values were measured. Protein quantities were then normalised based on the relative fluorescence increase (differential values) and in comparison to the empty plasmid backbone.

## Acknowledgements

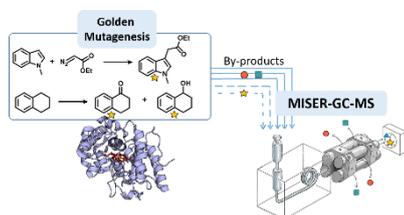
E.F. gratefully acknowledges the financial support by the Ministerium für Innovation, Wissenschaft und Forschung des Landes Nordrhein-Westfalen and M.J.W, A.K., E.S. and E.F thank the Bundesministerium für Bildung und Forschung („Biotechnologie 2020+ Strukturvorhaben: Leibniz Research Cluster“, 031A360B and 031A360E) for generous funding. P.P. thanks the Landesgraduiertenförderung Sachsen-Anhalt for a PhD scholarship.

The authors thank Miguel Alcalde (CSIC Madrid) and Dirk Holtmann (DFI Frankfurt) for providing the peroxygenase genes and Benjamin Jones for substantial support on Figure 1.

**Keywords:** High-throughput analytics • Biocatalysis • Carbene transfer • Unspecific peroxygenase • Hydroxylation

- [1] K. D. Collins, T. Gensch and F. Glorius, *Nat. Chem.* **2014**, *6*, 859-871.
- [2] Z. J. Gartner and D. R. Liu, *J. Am. Chem. Soc.* **2001**, *123*, 6961-6963.
- [3] E. Engvall and P. Perlmann, *Immunochemistry* **1971**, *8*, 871-874.
- [4] J. R. Cabrera-Pardo, D. I. Chai, S. Liu, M. Mrksich and S. A. Kozmin, *Nat. Chem.* **2013**, *5*, 423-427.
- [5] a) K. H. Shaughnessy, P. Kim and J. F. Hartwig, *J. Am. Chem. Soc.* **1999**, *121*, 2123-2132; b) A. C. Cooper, L. H. McAlexander, D.-H. Lee, M. T. Torres and R. H. Crabtree, *J. Am. Chem. Soc.* **1998**, *120*, 9971-9972.
- [6] a) H. A. Bunzel, X. Garrabou, M. Pott and D. Hilvert, *Curr. Opin. Struct. Biol.* **2018**, *48*, 149-156; b) P. Mair, F. Gielen and F. Hollfelder, *Curr. Opin. Chem. Biol.* **2017**, *37*, 137-144.
- [7] a) B. Sun, G. Miller, W. Y. Lee, K. Ho, M. A. Crowe and L. Partridge, *J. Chromatogr. A* **2013**, *1271*, 163-169; b) S. Kille, F. E. Zilly, J. P. Acevedo and M. T. Reetz, *Nat. Chem.* **2011**, *3*, 738-743.
- [8] P. Boeker and J. Leppert, *Anal. Chem.* **2015**, *87*, 9033-9041.
- [9] K. Izawa, K. Furuta, T. Fujiwara and N. Suyama, *Ind. Chim. Belge* **1967**, *32*, 223-226.
- [10] C. J. Welch, X. Y. Gong, W. Schafer, E. C. Pratt, T. Brkovic, Z. Pirzada, J. F. Cuff and B. Kosjek, *Tetrahedron Asymmetry* **2010**, *21*, 1674-1681.
- [11] a) K. Zawatzky, M. Biba, E. L. Regalado and C. J. Welch, *J. Chromatogr. A* **2016**, *1429*, 374-379; b) C. J. Welch, E. L. Regalado, E. C. Welch, I. M. K. Eckert and C. Kraml, *Anal. Methods* **2014**, *6*, 857-862; c) C. Welch, E. Regalado, C. Kraml, E. Welch, M. Welch, H. Semmelhack, D. Almstead, A. Kress, N. Hidalgo and M. Kress, *LC GC N. Am.* **2015**, *33*, 262-269.
- [12] a) J. P. Vistuba, M. Piovezan, M. G. Pizzolatti, A. M. Rebelo, M. S. Azevedo, L. Vitali, A. C. Costa and G. Amadeu Micke, *J. Chromatogr. A* **2013**, *1274*, 159-164; b) W. Schafer, H. Wang and C. J. Welch, *J. Sep. Sci.* **2016**, *39*, 2978-2985.
- [13] L. Silvestro, I. Tarcomnicu and S. R. Savu in *Matrix effects in mass spectrometry combined with separation methods—comparison HPLC, GC and discussion on methods to control these effects*, IntechOpen, **2013**.
- [14] M. J. Weissenborn and R. M. Koenigs, *ChemCatChem* **2020**, *12*, 2171-2179.
- [15] M. J. Weissenborn, S. A. Low, N. Borlinghaus, M. Kuhn, S. Kummer, F. Rami, B. Plietker and B. Hauer, *ChemCatChem* **2016**, *8*, 1636-1640.
- [16] K. J. Hock, A. Knorrscheidt, R. Hommelsheim, J. M. Ho, M. J. Weissenborn and R. M. Koenigs, *Angew. Chem. Int. Ed.* **2019**, *58*, 3630-3634.
- [17] a) K. Chen, S. Q. Zhang, O. F. Brandenburg, X. Hong and F. H. Arnold, *J. Am. Chem. Soc.* **2018**, *140*, 16402-16407; b) T. Hayashi, M. Tinzl, T. Mori, U. Krengel, J. Proppe, J. Soetbeer, D. Kloose, G. Jeschke, M. Reiher and D. Hilvert, *Nat. Catal.* **2018**, *1*, 578-584; c) T. K. Hyster and F. H. Arnold, *Isr. J. Chem.* **2015**, *55*, 14-20.
- [18] P. Püllmann, C. Ulpinnis, S. Marillonnet, R. Gruetzner, S. Neumann and M. J. Weissenborn, *Sci. Rep.* **2019**, *9*, 10932.
- [19] Z. J. Wang, H. Renata, N. E. Peck, C. C. Farwell, P. S. Coelho and F. H. Arnold, *Angew. Chem. Int. Ed.* **2014**, *53*, 6810-6813.
- [20] E. J. Moore and R. Fasan, *Tetrahedron Asymmetry* **2019**, *75*, 2357-2363.
- [21] a) R. Ullrich, J. Nuske, K. Scheibner, J. Spantzel and M. Hofrichter, *Appl. Environ. Microbiol.* **2004**, *70*, 4575-4581; b) P. Molina-Espeja, E. Garcia-Ruiz, D. Gonzalez-Perez, R. Ullrich, M. Hofrichter and M. Alcalde, *Appl. Environ. Microbiol.* **2014**, *80*, 3496-3507; c) M. Faiza, S. F. Huang, D. M. Lan and Y. H. Wang, *BMS Evol. Biol.* **2019**, *19*, 19; d) S. J. Willot, E. Fernandez-Fueyo, F. Tieves, M. Pesic, M. Alcalde, I. Arends, C. B. Park and F. Hollmann, *ACS Catal.* **2019**, *9*, 890-894.
- [22] E. D. Babot, J. C. del Rio, L. Kalum, A. T. Martinez and A. Gutierrez, *Biotechnol. Bioeng.* **2013**, *110*, 2323-2332.
- [23] M. Ramirez-Escudero, P. Molina-Espeja, P. Gomez de Santos, M. Hofrichter, J. Sanz-Aparicio and M. Alcalde, *ACS Chem. Biol.* **2018**, *13*, 3259-3268.
- [24] a) S. Cabantous and G. S. Waldo, *Nat Methods* **2006**, *3*, 845-854; b) J. Santos-Aberturas, M. Dorr, G. S. Waldo and U. T. Bornscheuer, *Chem. Biol.* **2015**, *22*, 1406-1414.
- [25] S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z. G. Zhang, D. J. Opperman, M. T. Reetz and J. P. Acevedo, *ACS Synth. Biol.* **2013**, *2*, 83-92.

## Entry for the Table of Contents



The developed multiple injections in a single experimental run (MISER) GC-MS technique allows a substrate independent assay for the 96-well plate analysis. Utilisation of this concept yielded novel unspecific peroxygenase chimeras, produced in *S. cerevisiae*, consisting of subdomains from three different fungal UPO genes for the hydroxylation of tetralin. The GC-technique was additionally applied to evaluate a YfeX library in an *E. coli* whole-cell system for the carbene-transfer reaction on indole, which revealed an thus far unknown axial heme ligand: tryptophan.

Institute and/or researcher Twitter usernames: Martin Weissenborn (@Weissenborn\_Lab)