Electronic Supplementary Material

Chemogenetic engineering of nitrobindin toward an artificial epoxygenase

Daniel F. Sauer,[a†] Malte Wittwer,[a†] Ulrich Markel,[a†] Alexander Minges,[b] Markus Spiertz,[a] Johannes Schiffels,[a] Mehdi D. Davari,[a] Georg Groth,[b] Jun Okuda,*[c] and Ulrich Schwaneberg*[ad]

[a.] *Institute of Biotechnology, RWTH Aachen University, Worringerweg 3, 52074 Aachen, Germany.*
[b.] *Institute of Biochemical Plant Physiology, Heinrich Heine University Düsseldorf, Universitätsstraße 1, 40225 Düsseldorf, Germany.*
[c.] *Institute of Inorganic Chemistry, RWTH Aachen University, Landoltweg 1, 52074 Aachen, Germany.*
[d.] *DWI - Leibniz Institute for Interactive Materials, Forckenbeckstr. 50, D-52056, Aachen, Germany.*
† These authors contributed equally to this work.
Electronic Supplementary Information (ESI) available: [details of any supplementary information available should be included here]. See DOI: 10.1039/x0xx00000x

Table of contents

I. Experimental Procedures

1.1 General Remarks

All chemicals were obtained from commercial suppliers (*Sigma-Aldrich/Merck, TCI*), unless otherwise noted. MnPPIX and NiPPIX were synthesized according to the published procedure.[1] FePPIX (*hemin*) was obtained from Sigma-Aldrich/Merck, all other metalated PPIX cofactors were obtained from *Enzo Life Sciences GmbH*. Glucose-Oxidase (GOx) from *Aspergillus niger* was obtained from *Sigma Aldrich/Merck*. Enzymes for molecular cloning were obtained from *New England Biolabs* and *Thermo Fisher Scientific*. Lysozyme from chicken egg white was purchased from *Sigma-Aldrich*.

1.2 Cloning and site-saturation mutagenesis

NB4 was previously subcloned by Fukumoto *et al.*[2] The NB4 sequence was subcloned into the pET42b(+) plasmid (restriction enzymes *Nde*I and *Xho*I) and equipped with an *N*-terminal *Strep*-tag II. NB-wt was subcloned accordingly (all DNA and amino acid sequences are specified in section III). Site-saturation mutagenesis (SSM) libraries were generated using the 22c-trick method.[3] Primers were obtained from *Eurofins Genomics* (primer sequences are available upon request). PCR products were gel-purified (*NucleoSpin Gel and PCR Clean-up kit*, *Macherey-Nagel*), digested with *Dpn*I, repaired using the method of Gibson[4], and used for the transformation of *E. coli* BL21-Gold (DE3) cells by electroporation. SSM libraries were extensively sequenced (*Eurofins Genomics*) to confirm full coverage of all 20 amino acids. In case the library did not encode for all 20 amino acids, the missing amino acid exchanges were prepared by site-directed mutagenesis (SDM) using non-degenerated primers encoding for the missing amino acid. The cell surface-displayed version of NB4H (csdNB4H) was prepared by SDM using non-degenerated primers and the previously reported csdNB4 gene as template.[5] The used eGFP-adhesion-promoting peptide fusion proteins were cloned by Meurer *et al.*[6] NB4H-MH was generated by the method of Gibson as described above.

1.3 Protein purification

E. coli BL21-Gold (DE3) cells freshly transformed with the gene of a nitrobindin variant were grown overnight in 150 mL minimal medium[7] supplemented with kanamycin (MM-Kan, 50 µg/mL) at 37 °C (250 rpm). The overnight culture was used to inoculate 1 L of MM-Kan (50 µg/mL). For eGFP-adhesion-promotion peptides, LB medium was used for the preculture

and TB medium for the main culture. Cells were grown (37 °C, 200 rpm) for 2.5 h, chilled on ice for 5 min, and gene expression was induced with Isopropyl β-D-1-thiogalactopyranoside (IPTG; final conc.: 1 mM for NB4, NB4H, and NB4HA; 0.2 mM for eGFP-adhesion-promoting peptide fusion proteins; 0.1 mM for NB4H-MH). Expression was performed at 18 °C, 200 rpm for 16 h. Cells were harvested by centrifugation (2,820 g, 4 °C, 15 min) and cell pellets were stored at -20 °C. For the purification of NB4, NB4H, and NB4HA, frozen cell pellets were resuspended in buffer A (50 mM sodium phosphate (NaPi) pH 8, 300 mM NaCl; 8 mL buffer per gram cell wet weight) and lyzed by sonication. Crude cell lysates were centrifuged (14,000 g, 4 °C, 30 min) and the supernatant was filtered (0.45 µm syringe filter). The cleared lysate was applied twice on a gravity flow column filled with *Strep-Tactin Sepharose* resin (*IBA Lifescience*) and washed twice with 7.5 column bed volumes (BV) of buffer A. The target protein was eluted using buffer B (50 mM sodium phosphate (NaPi) pH 8, 300 mM NaCl, 2.5 mM desthiobiotin). Fractions containing nitrobindin were pooled and concentrated to a total volume of 1 mL by ultrafiltration (10 kDa molecular weight cut-off, *Amicon Ultra*). The buffer was exchanged to reaction buffer (100 mM NaPi, pH 7.0) using a HiTrap Desalting column (*GE Healthcare*) following the producer's manual. Purified protein was flash-frozen in liquid nitrogen and stored at -20 °C. Nitrobindin concentration was determined by measuring the absorbance at 280 nm (apo-nitrobindin variants: $\varepsilon = 26\ mM^{-1}cm^{-1}$).

For the purification of NB4H-MH, frozen cell pellets were resuspended in buffer C (50 mM HEPES pH 8.0, 250 mM KCl). A Ni^{2+}-NTA resin (2 mL BV; *Qiagen*, Venlo, NL) in a gravity flow column (20 mL *Maxi column*; *G-Biosciences*, St. Louis, USA) was equilibrated with 30 mL (15 BV) buffer C. 50 mL cleared lysate was added and the column was washed using 20 mL (10 BV) buffer C containing 50 mM imidazole followed by 10 mL (5 BV) buffer C containing 100 mM imidazole. Elution was performed using 6 mL (3 BV) buffer C containing 250 mM imidazole. For the purification of eGFP-MH, buffer D (50 mM Tris-HCl, pH 8.0) was used. Purification was performed as described above with the following modifications: column wash was performed using 15 mL (7.5 BV) buffer D containing 50 mM imidazole and 15 mL (7.5 BV) buffer D containing 100 mM imidazole. Elution was performed using 6 mL (3 BV) buffer D containing 500 mM imidazole. For the purification of all other eGFP-adhesion-promoting peptide fusion proteins, pellets were resuspended in buffer D. Purification was performed with an ÄKTA equipped with a HisTrap HP 5 mL column prepacked with *Ni Sepharose High* (*GE Healthcare*). Washing and elution were performed by using a linear imidazole gradient (3 mL/min, 100 mL

in total, up to 250 mM imidazole). For all purified proteins, imidazole was removed using PD 10 Desalting Columns (*GE Healthcare Life Sciences*, Amersham, UK). Concentrations of proteins were quantified by measuring the absorbance at 280 nm (eGFP-LCI: ε = 45.8 $mM^{-1}cm^{-1}$; eGFP-EBA: ε = 33.6 $mM^{-1}cm^{-1}$; eGFP-MH, eGFP-Stomoxyn, eGFP-hDermcidin, eGFP-Spinigerin, eGFP: ε = 23.4 $mM^{-1}cm^{-1}$).

Cofactor-conjugated protein was prepared by incubating nitrobindin variants with 1.5 equiv. MnPPIX cofactor (final conc.: 88 µM protein, 132 µM MnPPIX, 5% MeCN as cosolvent) for 5 min at 23 °C. Excess of cofactor was removed using a HiTrap Desalting column (*GE Healthcare Life Sciences*, Amersham, UK) according to the producer's manual. Cofactor-conjugated protein was flash-frozen in liquid nitrogen and stored at -20 °C. The concentration of MnPPIX-conjugated nitrobindin variants was determined by measuring the absorbance at 280 nm (MnPPIX@NB4: ε = 29.6 $mM^{-1}cm^{-1}$; MnPPIX@NB4H: ε = 33 $mM^{-1}cm^{-1}$; MnPPIX@NB4-HA: ε = 33 $mM^{-1}cm^{-1}$, MnPPIX@NB4H-MH: ε = 32.4 $mM^{-1}cm^{-1}$).

For protein crystallography, NB4H was supplemented with MnPPIX cofactor (final conc.: 25 µM protein, 50 µM MnPPIX, 10% (v/v) MeOH as cosolvent) and concentrated to a final volume of 3 mL by ultrafiltration (10 kDa molecular weight cut-off, Amicon Ultra). The sample was further purified using a HiLoadTM 16/600 Superdex 200 pg column (GE Healthcare *Life Sciences*, Amersham, UK) equilibrated with MES buffer (5 mM MES, pH 6.0, 50 mM NaCl, 0.3 mM TCEP). Eluted protein samples were concentrated to 1 mM by ultrafiltration (10 kDa molecular weight cut-off, Amicon Ultra).

1.4 Screening of site-saturation mutagenesis libraries

Single colonies of *E. coli* BL21-Gold (DE3) transformed with a site-saturation library were used to inoculate single wells of a 96 well plate containing LB soft agar supplemented with kanamycin (50 µg/mL). The soft agar plate was sequenced (*Eurofins Genomics*) to confirm full coverage of all 20 amino acids. In case the library did not encode for all 20 amino acids, the missing amino acid exchanges were prepared by site-directed mutagenesis (SDM) using non-degenerated primers encoding for the missing amino acid. After confirming full library coverage, a glycerol stock master plate containing all possible 20 nitrobindin variants of the respective mutagenesis round was used to inoculate wells of a 96 deep-well plate containing 400 µL/well of MM-Kan (50 µg/mL). As a control, *E. coli* BL21-Gold (DE3) cells transformed with pET42b(+) not containing the nitrobindin gene ("empty vector control") were used. A

fresh 96 deep-well plate containing 900 µL/well of MM-Kan was inoculated with 100 µL/well of the preculture. This inoculation step was performed in such a way that from each well of the pre-culture plate three wells of the expression plate were inoculated (*i.e.*, generating triplicate wells for each nitrobindin variant or control). Cultures were grown for 2.5 h (37 °C, 900 rpm), chilled on ice for 5 min, and supplemented with IPTG (1 mM final concentration) before further incubation (18 °C, 900 rpm, 16 h). Cells were pelleted by centrifugation (3,200 g, 10 min, 4 °C), the supernatants were discarded, and the cell pellets were stored overnight at -20 °C. The pellets were thawed for 30 min at room temperature, 175 µL/well buffer E (100 mM Tris-HCl, pH 8.0, 150 mM NaCl) supplemented with Trition X-100 (conc.: 0.5% (v/v) were added, and the cell pellets were briefly resuspended by vortexing. 175 µL/well buffer E containing lysozyme (final conc.: 2 mg/mL) were added and the MTP was incubated for 1 h (37 °C, 900 rpm). Crude cell lysates were transferred to a 96 deep-well V-bottom plate and cell debris was pelleted by centrifugation (3,200 g, 20 min, 4 °C). Quantitative SDS-PAGE revealed typical nitrobindin titers of ~20 µM in cleared lysates (Figure S9). For each variant 3 × 255 µL cleared lysate (obtained from the three triplicate wells; *vide supra*) were transferred to a 1.5 mL tube and supplemented with 90 µL of MnPPIX cofactor (stock of 500 µM in MeOH; corresponds to ~3 equiv. with respect to nitrobindin scaffold protein), incubated 5 min at 23 °C, and centrifuged (1 min, 10,000 g) to remove any precipitates. As an additional control ("buffer control"), the same amount of MnPPIX cofactor was added to buffer E (supplemented with the same amounts of Triton X-100 and lysozyme as specified above) instead of to the lysates. Into a *Chromafil-Multi* 96-well filter plate (*Macherey-Nagel GmbH*) mounted on a vacuum manifold 200 µL/well of a *Strep-Tactin Sepharose* suspension (50% w/v; *IBA Lifescience*) were added and washed twice using 800 µL buffer E. The supernatant of the centrifuged, cofactor conjugated lysate (*vide supra*) was added to the *Strep-Tactin Sepharose* beads and the 96-well filter plate was incubated on a shaker (5 min, 23 °C, 900 rpm) before removing the liquid by vacuum and washing the beads twice with 800 µL buffer E. Finally, the beads were resuspended in 213 µL/well reaction buffer (100 mM NaPi, pH = 7) and 25 µL/well *trans*-β-methylstyrene (100 mM stock solution in MeCN) were added. The reaction was started by addition of 12.5 µL AcOOH (100 mM stock solution). The upper end of the filter plate was sealed using an adhesive foil and the lower end was sealed with a layer of *Parafilm* before placing it into modeling clay. After incubation on a microplate shaker (900 rpm, 23 °C, 2.5 h), the seal was removed, and the liquid reaction mixture was collected in a 96-well plate.

The reactions were extracted with 300 µL of dichloromethane (DCM) containing 1 mM mesitylene as internal standard. The samples were vortexed for 5 min and centrifuged (20,000 g; 5 min) and the organic phase was dried over magnesium sulfate prior to GC analysis. Background activity, as determined by the empty vector control, was subtracted from library members. The libraries were screened in two independent experiments, screening each variant in duplicates, respectively. Results of the screening on position 158 are shown in Figure S10.

1.5 Catalysis with purified protein

The purified ArM in reaction buffer (100 mM NaPi, pH = 7, total volume 500 µL, c(ArM) = 10 µM)) was incubated with a substrate (1 mM) and 30% acetonitrile. The reaction was started by addition of 2.5 equiv. of peracetic acid (2.5 mM). Alternatively, the reaction was started by addition of 40 equiv. H_2O_2 (buffer: 50 mM sodium borate, pH = 10.5). After 2 hours reaction time at room temperature, the excess peracetic acid was quenched by addition of sodium dithionite (5 mM). The reaction mixture was extracted with dichloromethane (175 µL) containing mesitylene as internal standard (1 mM). The extract was analyzed with GC-FID. The calibration curve for the product is depicted in Figure S11. Representative analysis of the catalysis runs with MnPPIX@NB4H and MnPPIX@NB4HA are shown in Figure S12 and Figure S13, respectively.

For the cascade reaction involving the GOx for H_2O_2 production, similar reaction conditions were applied. The purified ArM in aqueous buffer (50 mM sodium borate, pH = 10.5, total volume 500 µL, c(ArM) = 10 µM)) was incubated with *trans*-β-methylstyrene (1 mM) and 30% acetonitrile. The reaction was started by addition of GOx (1 U) and 2 equiv. of glucose. Work-up and analysis of the reaction was done as described above.

1.6 Determination of the kinetic parameters

To determine the kinetic parameters, purified ArM or MnPPIX in reaction buffer (100 mM NaPi, pH = 7, total volume 500 µL, c(ArM) = 10 µM)) was incubated with a substrate (0-2 mM) and 30% acetonitrile. The reaction was started by the addition of peracetic acid (2.5 mM). After exactly 5 minutes reaction time at room temperature, the reaction was quenched by the addition of sodium dithionite (5 mM). The reaction mixture was extracted with dichloromethane (175 µL) containing mesitylene as internal standard (1 mM). The extract was analyzed with GC-FID.

1.7 GC Analysis

GC analysis was carried out on a GC-FID system (*GC2010 Shimadzu*) using hydrogen as carrier gas. Reactions were analyzed on an *FS Lipodex-E* (*Macherey-Nagel*, Düren, Germany). Column oven temperature: 80 °C (12 min), 30 °C/min 210 °C (1 min). Signals for the products and starting materials appeared within the first 10 minutes.

1.8 UV/Vis spectroscopy

UV/Vis spectra were recorded on a *Varian Cary 50 Bio* spectrophotometer. UV-cuvette semi-micro (12.5 x 12.5 x 45 mm) made from polystyrene were used (Brand, Frankfurt, Germany). To ensure equal concentrations during the measurement, 15 µM MnPPIX were incubated with 20 µM of the corresponding nitrobindin variant. To measure the corresponding spectra after oxidation, the oxidant was added directly to the cuvette located inside the instrument and the measurement was started immediately after addition.

1.9 Crystallization and structure analysis

MnPPIX@NB4H for X-ray crystal structure determination was crystallized using the microbatch method at 285.15 K. Crystals were grown by microbatch under mineral oil (Sigma Aldrich, M3516) from a mixture that contained 1.0 µL of protein solution (1.0 mM protein in 5 mM MES buffer, pH 6.0, 50 mM NaCl, 0.3 mM TCEP) and 1.0 µL of precipitant (100 mM Tris-HCl pH 8.6, 18% (w/v) polyethylene glycol (PEG) 2000). Crystals formed within 48 h after setup.

Crystals were briefly soaked in a cryoprotectant solution (100 mM Tris-HCl pH 8.6, 18% (w/v) PEG 2000, 30% (v/v) ethylene glycol) and flash-cooled in liquid nitrogen. X-ray diffraction data were collected in-house and at the European Synchrotron Radiation Facility (ESRF) at beamline ID30B.

The data were integrated and scaled using the program XDS[8] and the autoPROC toolbox.[9] Further processing was performed using the CCP4[10] and PHENIX[11] software packages. Initial phases were obtained by molecular replacement using PHASER[12] with the structure of NB4 (PDB code 3WJB) serving as template. The resulting model was subjected to multiple rounds of iterative automatic and manual rebuilding with BUCCANEER[13] and COOT.[14] Each round of rebuilding was combined with crystallographic refinement using REFMAC5[15] or PHENIX.refine.[16] Geometry restraints for MnPPIX were generated using PHENIX.elbow.[17] Data collection and refinement statistics are summarized in Table S1. Figures that depict the

structures were prepared in PYMOL.[18] The atomic coordinates and structure factors were deposited in the Protein Data Bank (www.pdbe.org) under the accession code 7BBM.

Table S1. Data collection and refinement statistics

	MnPPIX@NB4H (PDB 7BBM)
Data collection	
Wavelength (Å)	0.9763
Resolution range (Å)	39.774 - 1.137 (1.157 - 1.137)
Space group	$P2_12_12$
Cell parameters	
a, b, c (Å)	59.9 79.547 36.56
α, β, γ (°)	90 90 90
Total reflections	789798 (34878)
Unique reflections	64928 (3190)
Multiplicity	12.2 (10.9)
Completeness (%)	100.0 (100.0)
Mean I/σ_I	14.5 (2.2)
Wilson B-factor (Å^2)	14.83
R_{merge}	0.075 (1.037)
R_{meas}	0.078 (1.088)
R_{pim}	0.022 (0.325)
$CC_{1/2}$	0.998 (0.811)
CC*	1 (0.956)
Model and refinement	
Reflections used in refinement	64903 (6413)
Reflections used for R_{free}	3158 (309)
R_{work}	0.1448 (0.2185)
R_{free}	0.1583 (0.2399)
CC_{work}	0.962 (0.925)
CC_{free}	0.934 (0.935)
Number of non-hydrogen atoms	1620
Protein residues	160
RMSD	
Bond lengths (Å)	0.012
Bond angles (°)	1.34
Ramachandran analysis	
Favored regions (%)	98.01
Allowed regions (%)	1.99
Outliers (%)	0.00
Rotamer outliers (%)	0.65
Clashscore	6.25
Average B-factor	21.64

Statistics for the highest-resolution shell are shown in parentheses

1.10 CD-Spectroscopy

Circular dichroism (CD) spectra were recorded on a JASCO J-1100 spectrometer equipped with a single position Peltier cell holder. If not stated otherwise, the temperature was set to

20 °C. The pathlength of the plate cuvette was 0.2 mm. The protein concentration was 20 µM.

For variable temperature CD, a cuvette with 1 mm pathlength was used. The temperature range was set from 2 °C to 92 °C and measured in 2 °C steps. The temperature was measured inside the cuvette.

Table S2: Melting temperatures of the variants determined by CD spectroscopy.

Entry	Variant	T_m [°C]
1	NB4	54
2	NB4H	56
3	MnPPIX@NB4H	74
4	MnPPIX@NB4	50 and 77

Nitrobindin variant NB4 has a T_m = 54 °C (Table 1, entry 1). This T_m is not significantly influenced by introducing the His158 (NB4H, Table 1, entry 2). Incorporating the MnPPIX cofactor into NB4H stabilizes the protein structure as suggested by a T_m increase of about 20 °C (Table 1, entry 3, Figure S4). The temperature-dependent CD spectra of NB4 indicate two subsequent denaturation events (Table 1, entry 4, Figure S5). At the first denaturation event (T_m = 50 °C), a ~50% loss in ellipticity was observed, followed by complete denaturation at T_m = 77 °C, which is close to that of MnPPIX@NB4H. We attribute the lower melting point to the absence of a proximal ligand in NB4, which results in an overall weaker MnPPIX binding. (Temporarily) unoccupied NB4 molecules are less stable and hence denature at lower temperatures. We, therefore conclude that the introduction of the proximal H158 is key to stabilize the ArM at elevated temperatures.

1.11 Binding of adhesion-promoting peptide fusion proteins to *E. coli* cells

Cell-binding was performed by mixing the components listed in Table S2. The mixtures were incubated for 2 minutes at 25°C in an incubation shaker (*ThermoMixer comfort*; *Eppendorf AG*, Hamburg, DE). Following incubation, cells were washed twice by sedimentation at 11,000 g for 90 s and resuspension in 300 µL HEPES (50 mM, pH 7.0).

Table S3: Components mixed for cell-binding experiments.

Component	Volume
E. coli cell suspension (OD 4.44) [a]	270 µL
Protein stock solution (20 µM) [b]	30 µL
Total volume	**300 µL**

[a] *E. coli* BL21-Gold (DE3) cells in the tested buffer. [b] eGFP-adhesion-promotion peptide fusion proteins were in 10 mM NaPi, pH 7.4. NB4H and NB4H-MH were in 50 mM HEPES + 250 mM KCl, pH 8.

1.12 Labeling of NB4H and NB4H-MH with Strep-Tactin Chromeo 546

To visualize NB4H-MH on *E. coli* cells, the fluorescent dye Strep-Tactin Chromeo 546 conjugate (*Chromeo 546*; *IBA Lifesciences*, Göttingen, DE) was used. To this end, NB4H-MH was bound to cells as described in section 1.11. Here, however, cells were resuspended in 300 µL HEPES (50 mM, pH 7.0) containing 500 nM Strep-Tactin Chromeo 546 after the first sedimentation step. Binding of Strep-Tactin Chromeo 546 to the *N*-terminal Strep-tag of NB4H-MH was performed for 15 minutes. Afterward, cells were washed twice by sedimentation at 11,000 g for 90 s and resuspension in 300 µL HEPES (50 mM, pH 7.0). The fluorescence was detected by fluorescence microscopy (λ_{ex} = 555 nm, λ_{em} = 580 nm). As control, NB4H (*i.e.*, the scaffold protein without MH) was treated as described above.

1.13 Coupling of ThioGlo 1 to NB4H-MH

For labeling of NB4H-MH with ThioGlo 1, 10 µM of NB4H-MH (in 50 mM HEPES, 250 mM KCl, pH 8) was incubated with 30 µM ThioGlo 1 (1.5 mM stock solution in acetonitrile; *Berry & Associates Inc*., Dexter, USA) for 30 min at 23 °C. The precipitate was removed by centrifugation (20,000 g at 23 °C for 3 min). To remove any excess of ThioGlo 1, the obtained supernatant was rebuffered in 50 mM HEPES, 250 mM KCl, pH 8 and concentrated. ThioGlo@NB4H-MH was bound to cells as described in section 1.11.

1.14 Characterization of the cell-binding of adhesion-promoting peptide fusion proteins

To determine the time required for cell attachment, cell-binding was performed as described in section 1.11 with the following modifications: cells were incubated with the tested protein

for various time points (0-30 min). After washing as described in section 1.11, the fluorescence intensity of resuspended cells was measured using a *Clariostar* plate reader (*BMG*). (eGFP: λ_{ex} = 470 nm, λ_{em} = 508 nm, ThioGlo 1: λ_{ex} = 379 nm, λ_{em} = 513 nm)

To determine the time eGFP-MH and NB4H-MH fusion proteins remain attached to *E. coli* cells, cell-binding was performed as described in section 1.11. Following washing, cells were incubated for various time points (1-48 hours) at 25°C in an incubation shaker. After the incubation, cells were washed again twice as described in section 1.11. The fluorescence intensity was measured using a *Clariostar* plate reader (*BMG*).

To determine the cell load (*i.e.*, the number of bound molecules per cell), cell-binding mixtures with varying concentrations of eGFP-MH were prepared as described in section 1.11 with the following modifications: A stock solution of 0-40 µM eGFP-MH was used. The fluorescence intensity of washed and resuspended cells was measured using a *Clariostar* plate reader (*BMG*). The cell load was calculated according to equation (eq. 1) using a calibration curve of eGFP-MH in presence of cells.

$$B = \frac{(F - n) \cdot N_A}{m \cdot c_{cells} \cdot OD_M} \qquad \text{(eq. 1)}$$

where

B	Cell load (the number of bound molecules per cell)	[-]
F	Measured fluorescence intensity	[RFU]
n	Ordinate intercept of the calibration curve	[RFU]
N_A	Avogadro constant $= 6.022 * (10^{17})$	[µmol^{-1}]
m	Slope of the calibration curve	[RFU µM^{-1}]
c_{cells}	*E. coli* cell number at $OD_{600} = 1$ $(= 8 * 10^{11})_{19}$	[L^{-1}]
OD_M	OD_{600} of the measured cell suspension ($= 4$)	[-]

The calculated cell load (B) was plotted against the applied protein concentration c_P. To estimate the maximum cell load (B_{max}), equation (2) was fitted to the data.

$$B = \frac{B_{max} \cdot c_P}{K_i + c_P} \qquad \text{(eq. 2)}$$

where

B	Cell load (the number of bound molecules per cell)	[-]
B_{max}	Maximal cell load	[-]
c_P	Applied protein concentration	[µM]
K_i	Concentration at which half of the maximum binding is reached	[µM]

1.15 Microscopy

For microscopy, the microscope BX51 (*Olympus*, Shinjuku, Japan) with a U-RFL-T burner and a BX-UCB control box was used. 0.5 µL cell suspension was transferred to a glass slide and covered with a cover-slip. A drop of immersion oil was pipetted onto the cover-slip and cells were inspected with the 100x magnification objective. For fluorescence microscopy, the exposure time and gain were set to 53.14 ms and 5 dB, respectively.

1.16 On-cell catalysis

MnPPIX@NB4H-MH was bound onto “carrier cells”. As carrier cells, *E. coli* BL21-Gold (DE3) cells were grown overnight in LB medium (32 °C, 250 rpm). On the following day, 50 mL of this preculture were transferred into 150 mL of fresh LB medium and grown for 5 h (32 °C, 250 rpm). The cells were harvested by centrifugation (2,820 g, 4 °C, 15 min) and the pellets were washed with HEPES buffer (50 mM; pH 7) before adjusting the optical density (OD_{600}) to OD_{600} = 56.

For on-cell catalysis experiments, 1.1 equiv. of MnPPIX (0.5 mM stock in MeOH) was added to 90 µM of the NB4H-MH fusion protein and incubated (23 °C, 15 min). The mixture was centrifuged (20,000 g, 3 min, 23 °C) to remove any precipitate and concentrated by ultrafiltration (10 kDa molecular weight cut-off, *Amicon Ultra*). The protein concentration of the solution was determined by measuring the absorbance at 280 nm (ε = 40.5 $mM^{-1}cm^{-1}$) and the concentration was adjusted to 300 µM. Next, 100 µL of MnPPIX@NB4H-MH (300 µM) was added to 900 µL of the carrier cells (OD_{600} = 56) and briefly mixed by pipetting (final OD_{600} =

50). As controls, MnPPIX ("Cells + MnPPIX" sample) or no metal cofactor ("Cells" sample) was added to the carrier cells instead of MnPPIX@NB4H-MH. The cells were immediately washed by centrifugation (11,000 g, 30 s, 23 °C) and resuspension: in a first washing step 1 mL of HEPES buffer (50 mM; pH 7) supplemented with 5% (v/v) MeCN was used; in a second washing step 1 mL of HEPES buffer (50 mM; pH 7) was used. Finally, the MnPPIX@NB4H-MH decorated cells (and the respective MnPPIX/no cofactor controls) were pelleted (11,000 g, 90 s, 23 °C) and resuspended in 780 µL HEPES buffer (50 mM; pH 7).

As a comparison to the adhesion-promoting peptide-based system, csdNB4H was used. Cells expressing csdNB4H were prepared as previously described,[5] washed once with HEPES buffer (50 mM; pH 7) and resuspended to OD_{600} = 180 using HEPES buffer (50 mM; pH 7). On average, 12,800 nitrobindin molecules were displayed per cell (corresponding to 3.2 µM csdNB4H).[5] 200 equiv. MnPPIX (50 mM stock in MeOH) was added (final volume: 1 mL) and the cells were immediately washed as described above before resuspending them in HEPES buffer (50 mM, pH 7).

To start on-cell epoxidation reactions, 20 µL copper(II) glycinate[20, 21] (100 mM stock), 100 µL *trans*-β-methylstyrene (50 mM stock in MeCN), and 100 µL AcOOH (50 mM stock) were added to the cells bearing either MnPPIX@NB4H-MH, or the MnPPIX-conjugated csdNB4, or a control ("Cells" or "Cells + MnPPIX"). Final concentrations were: 5 mM *trans*-β-methylstyrene, 5 mM AcOOH, 2 mM copper(II) glycinate, cells at OD_{600} = 50, HEPES buffer (50 mM, pH 7); 10% (v/v) MeCN as cosolvent. The reactions were incubated for 16 h (23 °C, 900 rpm) before extraction with 350 µL of dichloromethane (DCM) containing 1 mM mesitylene as internal standard. The samples were vortexed for 5 min and centrifuged (20,000 g; 5 min) and the organic phase was dried over magnesium sulfate prior to GC analysis.

1.17 Whole-Cell Screening

Cells containing the NB4-MH construct with the NRT library were grown and harvested as described in Section 1.4. The harvested cell pellets were stored overnight at -20 °C. The pellets were thawed for 30 min at room temperature, 125 µL/well buffer F (50 mM NaPi, pH 8.0, 250 mM KCl) supplemented with Trition X-100 (conc.: 0.3% (v/v) were added, and the cell pellets were briefly resuspended by vortexing. 125 µL/well buffer F containing lysozyme (final conc.: 2 mg/mL) and DNase I (final conc.: 0.15 mg/mL) were added and the MTP was incubated for 30 min (37 °C, 900 rpm). Crude cell lysates were transferred to a 96 deep-well V-bottom plate

and cell debris was pelleted by centrifugation (3,200 g, 20 min, 4 °C). Into the cell-free lysates, MnPPIX was added (4.5 mM stock in methanol, final concentration: 20 µM). 200 µL of the cell-free lysate containing MnPPIX were transferred to a tube containing *E. coli* carrier cells (1750 µL, HEPES-Buffer (50 mM, pH 7.0), OD_{600} = 7.5). To avoid damaging the membrane integrity of the carrier cells, the cells were washed without further incubation time twice by sedimentation at 11,000 g for 2 min and resuspended in 1 mL HEPES-buffer (50 mM, pH = 7). Finally, the washed cell-pallet was resuspended in 500 µL HEPES-buffer (50 mM, pH = 7). *trans*-β-methylstyrene (1 mM) and 10 (v/v)% acetonitrile were added. The reaction was started by addition of 2.5 equiv. of peracetic acid (2.5 mM). After 2 hours reaction time at room temperature, the excess peracetic acid was quenched by addition of sodium dithionite (5 mM). The reaction mixture was extracted with dichloromethane (175 µL) containing mesitylene as internal standard (1 mM). The extract was analyzed with GC-FID.

For the screening of different catalysts, 150 µL of purified NB4H-MH (15 µM in 50 mM HEPES buffer, pH = 7) were incubated with an access of the corresponding metal-PPIX cofactors (fincal concentrations 20 µM taken from 3 mM stocks (in methanol)). The mixture was transferred onto *E. coli* carrier cells (900 µL, OD_{600} = 15) and immeadiately washed twice by sedimentation at 11,000 g for 2 min and resuspended in 1 mL HEPES-buffer (50 mM, pH = 7). Finally, the washed cell-pallet was resuspended in 500 µL NaPi buffer (100 mM, pH = 8), the cells collected by centrifugation and supernated used for catalysis. *trans*-β-methylstyrene (1 mM) and 10 (v/v)% acetonitrile were added. The reaction was started by addition of 2.5 equiv. of peracetic acid (2.5 mM). After 2 hours reaction time at room temperature, the excess peracetic acid was quenched by addition of sodium dithionite (5 mM). The reaction mixture was extracted with dichloromethane (175 µL) containing mesitylene as internal standard (1 mM). The extract was analyzed with GC-FID.

1.18 Computational methods

Structural models of MnPPIX@NB4H and MnPPIX@NB4HA were built with YASARA Structure version 17.4.17[22] using the swap function based on wildtype nitrobindin (PDB-code: 3emm[23]). The initial structures were optimized using the SCWRL rotamer library search[24] for the designed substitutions. The Fe atom was replaced by a Mn atom. The models were neutralized and solvated in a periodic box containing TIP3P[25] water. All energy minimizations and

molecular dynamics (MD) simulations were carried out using the AMBER14[26] and GAFF force field.[27] Bond length to hydrogen atoms and bond angles in water were constrained to increase the speed for the simulations.[25] Atomic partial charges were derived using the AM1/BCC procedure implemented in YASARA. Electrostatic interactions were calculated using a cut-off of 7.86 Å and long-range interactions were calculated by using the particle-mesh Ewald (PME) integration. Initial energy minimizations by steepest descent minimization were followed by simulated annealing until reaching convergence in potential energy (<0.02 kJ mol^{-1} per atom during 200 steps). MD simulations were performed for 100 ns at 298 K by rescaling the time-averaged atom velocities using a Berendsen thermostat and a solvent density of 0.997 kg m^{-3}. The protein structures were visualized using Pymol.[18]

A structural model of the macaque histatin (MH) was constructed using the I-TASSER server[28] based on the multiple threading method. The structure validation was performed at the SAVES server (https://servicesn.mbi.ucla.edu/SAVES/) by PROCHECK. The protonation states of titratable residues of MH were assigned on the basis of pKa values calculated using the PROPKA 3.1 program[29] and visual inspection at pH 7 (net charge of +4). To evaluate the model and force field, we performed a systematic study directed toward the secondary structure propensity with different molecular dynamics (MD) force-field variants (using the AMBER03,[26] AMBER99SB-ILDN,[30] and GROMOS54A7 force field[31]) in explicit solvent. As can be seen in Figure S13, GROMOS54A7 force field and AMBER99SB-ILDN retained the secondary structure propensity of MH. Therefore, to serve the purpose of force field compatibility with the *E. coli* outer membrane model developed by Piggot *et al.*,[32] the GROMOS54A7 force field was further used for simulations of peptide binding to the membrane. The initial coordinates and force field parameters of the *E. coli* outer membrane were taken from an atomistic model developed by Piggot *et al.*[32] In this model, the inner leaflet of the *E. coli* outer membrane contains 160 phospholipids. 144 of those phospholipids are 1-palmitoyl, 2-*cis*-vaccenyl phosphatidylethanolamines (PVPE), 8 are 1-palmitoyl, 2-*cis*-vaccenyl phosphatidylglycerols (PVPG), and 8 are diphosphatidylglycerols (DPG) with the fatty acyl chain composition 1-palmitoyl, 2-*cis*-vaccenyl, 3-pamitoly, 4-*cis*-vaccenyl (PVPV). The outer leaflet consists of 64 Rd1 LPS containing the core type R1. The GROMOS54A7 force field[31] was used for MH and the membrane model. To neutralize the system charge, solvent water molecules were replaced by $Na^{+}Cl^{-}$ ions. Because the main aim of this study was to understand the MH-membrane interaction, we started from configurations in which the MH model was placed in four

different orientations approximately 7 Å above the *E. coli* outer membrane. Three independent MD simulation runs were conducted for each configuration of the MH peptide on *E. coli*. The systems contained ~ 86,000 atoms in total, including ~ 57800 SPC/E[33] water molecules. Initially, the solvent and the ions and subsequently the whole system were subjected to minimization using 10,000 steps of steepest descent minimization followed by 3,000 steps of conjugate-gradient minimization. The system was then slowly heated from 0 to 300 K for 50 ps. In all MD simulations, constant pressure periodic boundary conditions using the Particle Mesh Ewald (PME)[34] method were employed. To calculate the electrostatic interactions a cutoff of 10 Å was used. After heating, the systems were equilibrated for 1000 ps at 300 K. Finally, three independent production runs were performed for 100 ns for each starting conformation resulting in a total of 1.2 μs simulation time. All classical MD simulations were performed with GROAMCS.[35] VMD[36] and GROAMCS tools were used for molecular visualizations and analysis of molecular dynamics trajectories.

II. Supplementary Figures

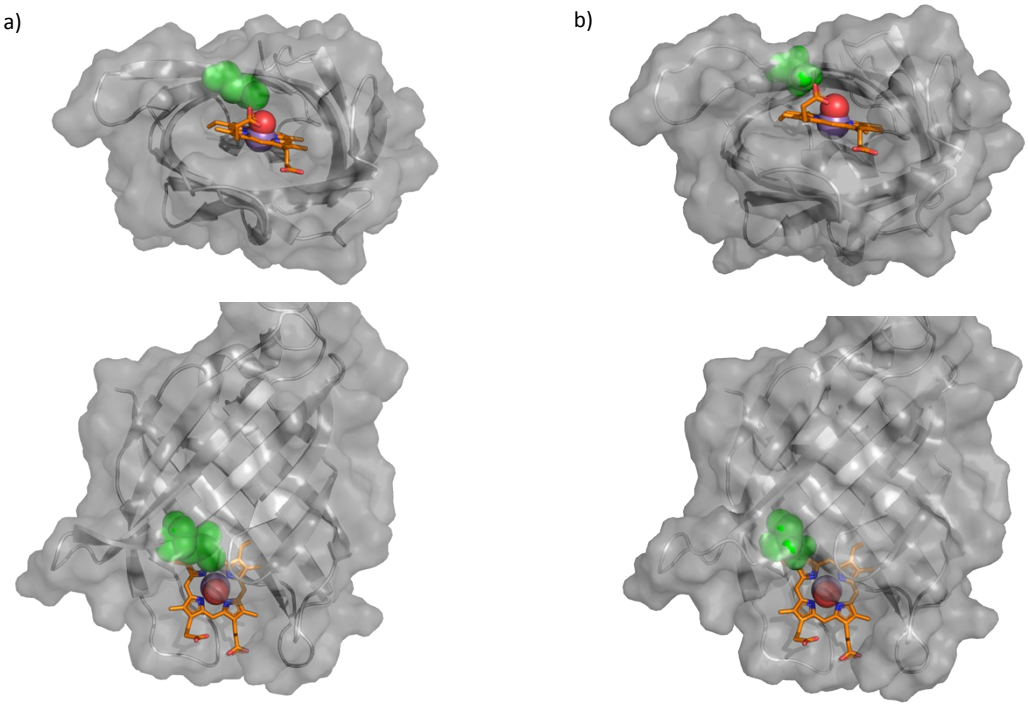

Figure S1: Molecular model of a) MnPPIX@NB4H and b) MnPPIX@NB4HA. The protein is shown as gray cartoon with its light-gray surface. The PPIX framework is shown as orange sticks, the Mn atom as purple sphere, and the oxygen bound to the Mn atom as red sphere. The L76 (NB4H) and the A76 (NB4HA) are shown in green spheres. In the top row, the front view is shown, in the bottom row the view from the top is shown. The distance between the oxygen atom to L76 is closer (a) than the distance to A76 (b), due to the shorter side chain of the alanine.

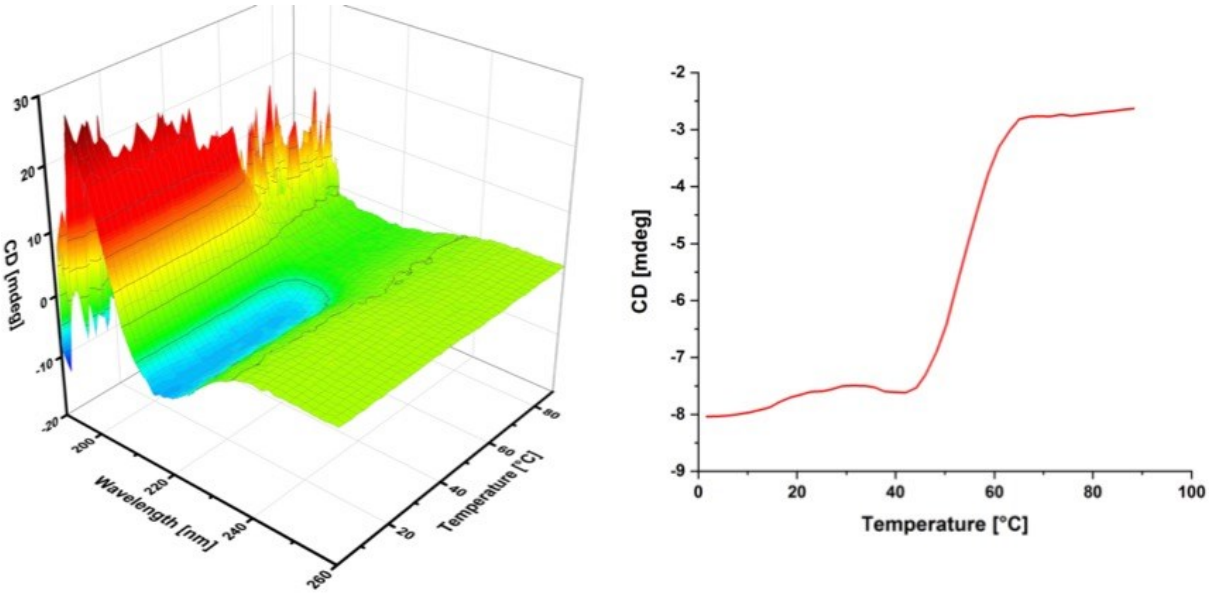

Figure S2: VT-CD spectrum of apo NB4 (left) and the CD-signal at 218 nm over the temperature range from 2 °C to 92 °C (right). The protein concentation for the measurement was c(NB4) = 11 µM. The melting temperature was T_m = 54 °C, which is in good agreement with the previously reported value.[37]

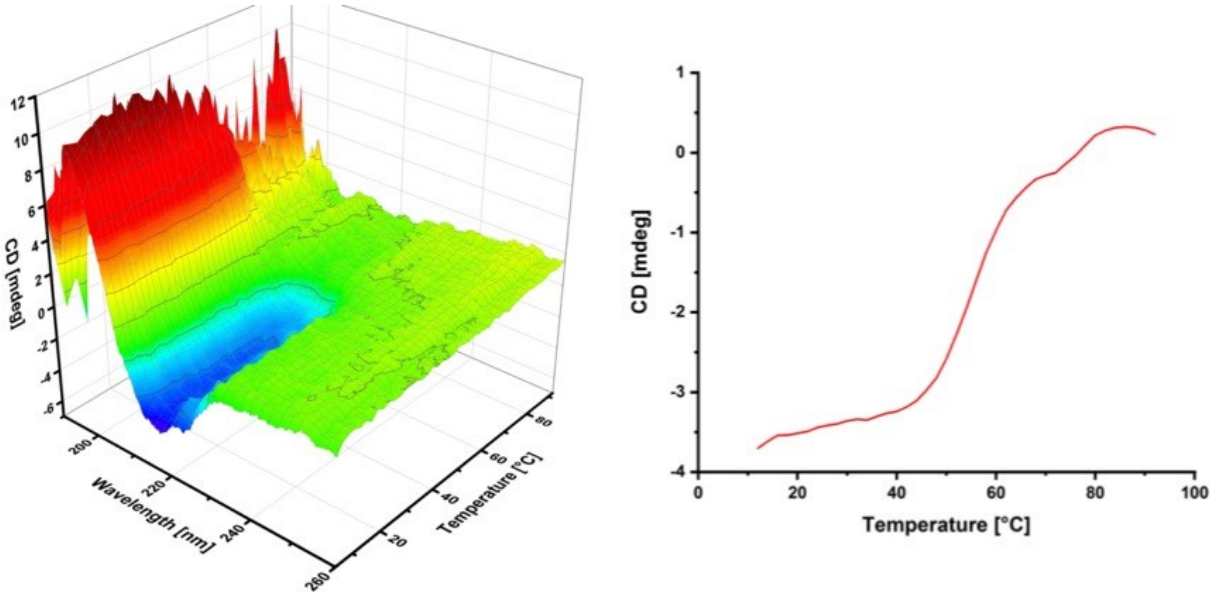

Figure S3: VT-CD spectrum of apo NB4H (left) and the CD-signal at 218 nm over the temperature range from 10 °C to 92 °C (right). The protein concentation for the measurement was c(NB4H) = 7 µM. The melting temperature was T_m = 56 °C.

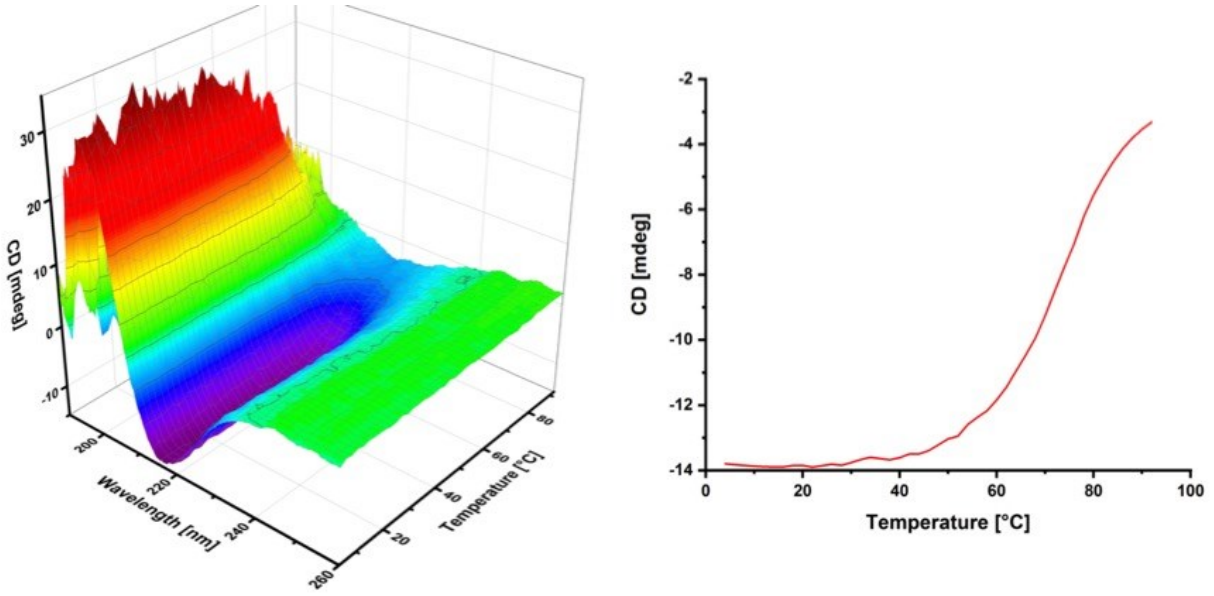

Figure S4: VT-CD spectrum of MnPPIX@NB4H (left) and the CD-signal at 218 nm over the temperature range from 2 °C to 92 °C (right). The protein concentation for the measurement was c(MnPPIX@NB4H) = 10 µM. The melting temperature was T_m = 74 °C.

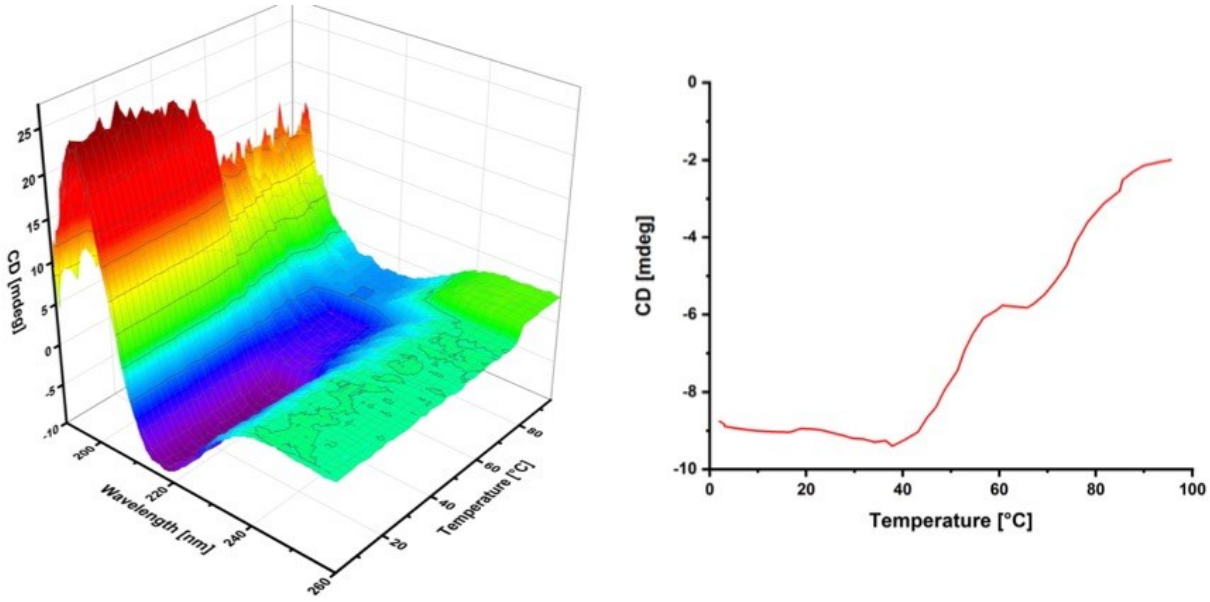

Figure S5: VT-CD spectrum of MnPPIX@NB4 (left) and the CD-signal at 218 nm over the temperature range from 2 °C to 92 °C (right). The protein concentation for the measurement was c(MnPPIX@NB4H) = 10 µM. Two melting temperature are observed at T_m = 50 °C and T_m = 77 °C.

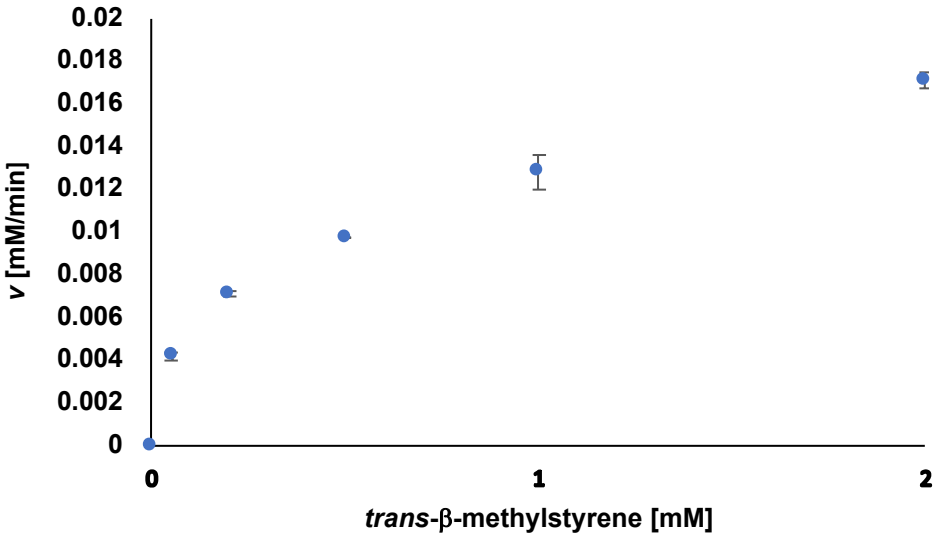

Figure S6: Kinetic characterization of the artificial epoxygenase MnPPIX@NB4H with *trans*-β-methylstyrene as the substrate and AcOOH as the oxidant. Experimental details are given in section 1.6.

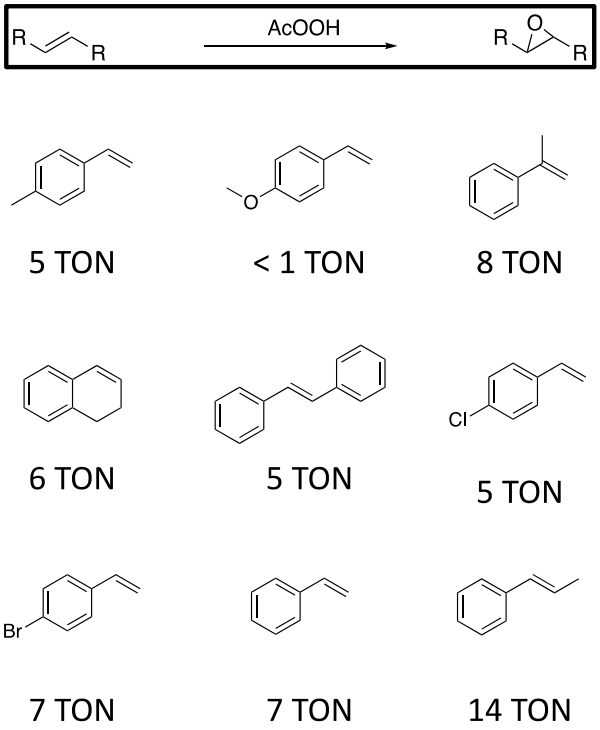

Figure S7: Substrate scope of the artificial epoxygenase MnPPIX@NB4H. Experimental procedures are given in section 1.5.

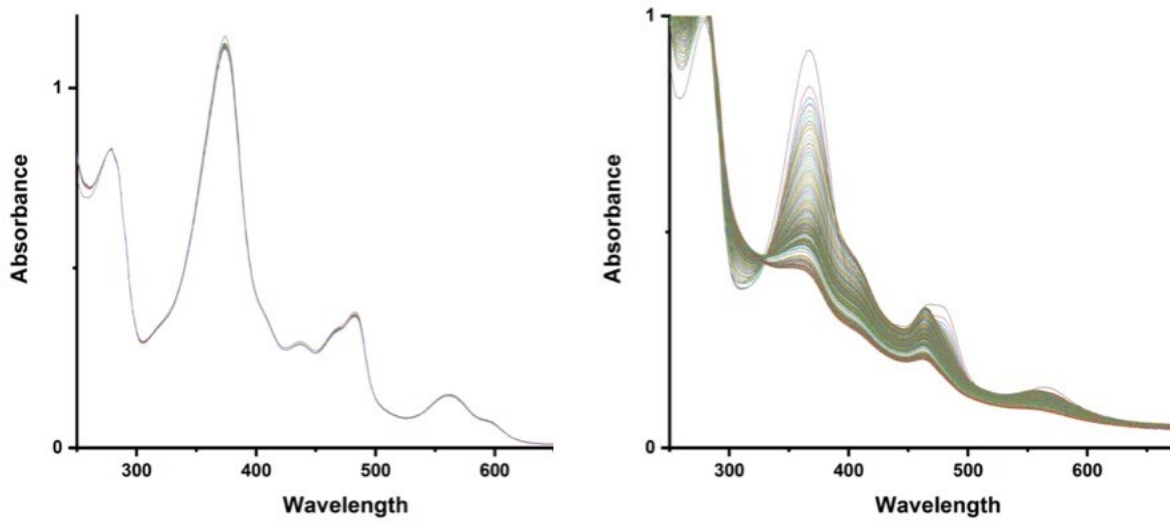

Figure S8: UV/Vis spectra of MnPPIX@NB4 in the presence of 100 equiv H_2O_2 at pH = 7 (left) and pH = 10.5 (right) over a 15 min. time period.

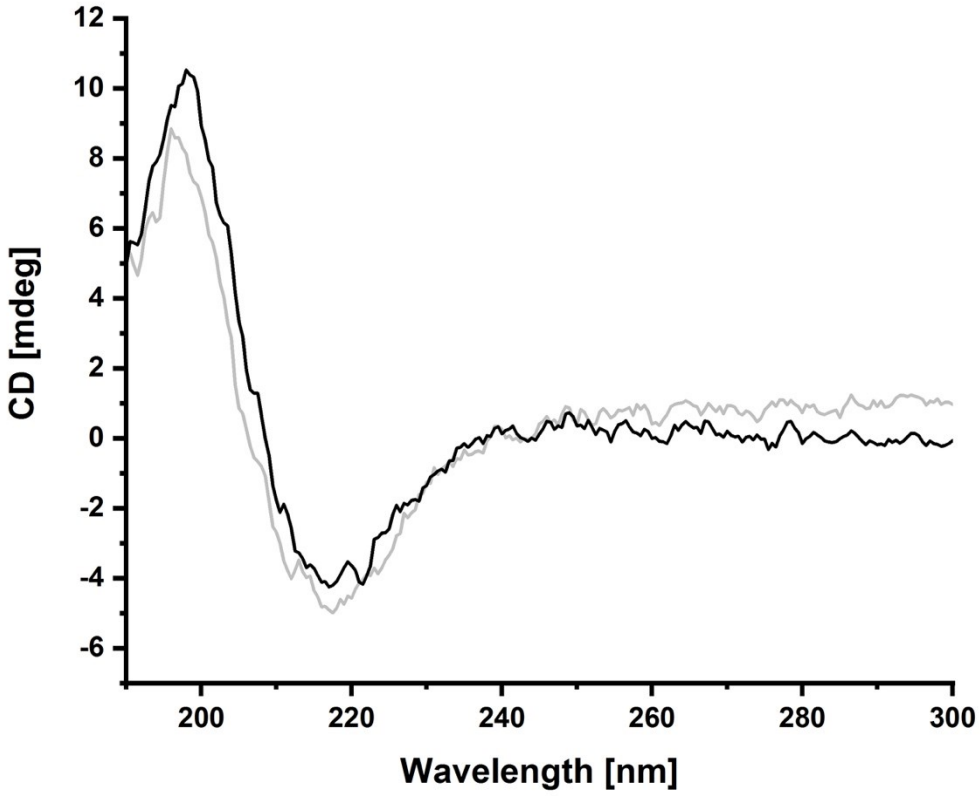

Figure S9: CD-spectra of MnPPIX@NB4H in the range from 190-300 nm without MeCN (black) and with 30 (v/v)% MeCN (gray).

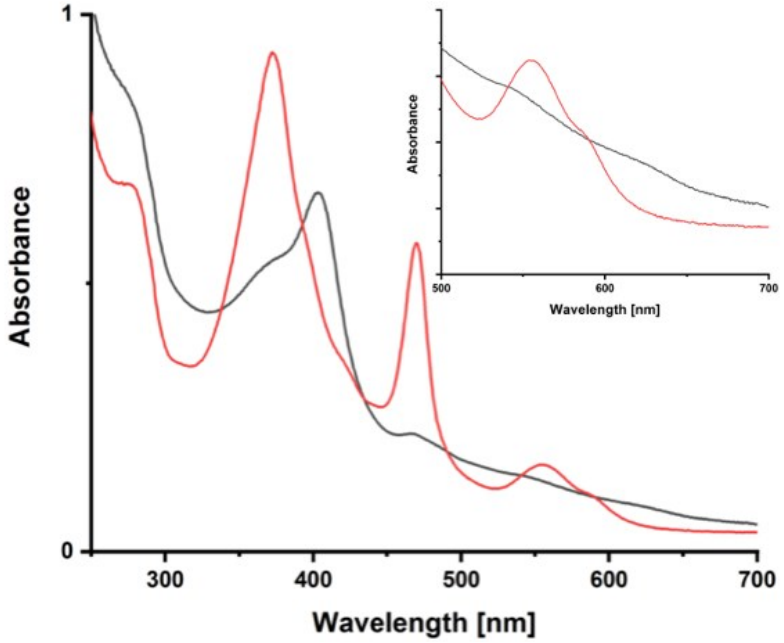

Figure S10: Activation of the oxygen source in the presence of 30 (v/v)% MeCN analyzed by UV/Vis spectroscopy. MnPPIX@NB4H before addition (red) and upon addition of 100 equiv. AcOOH (black). The inset shows the spectra in the range of 500-700 nm.

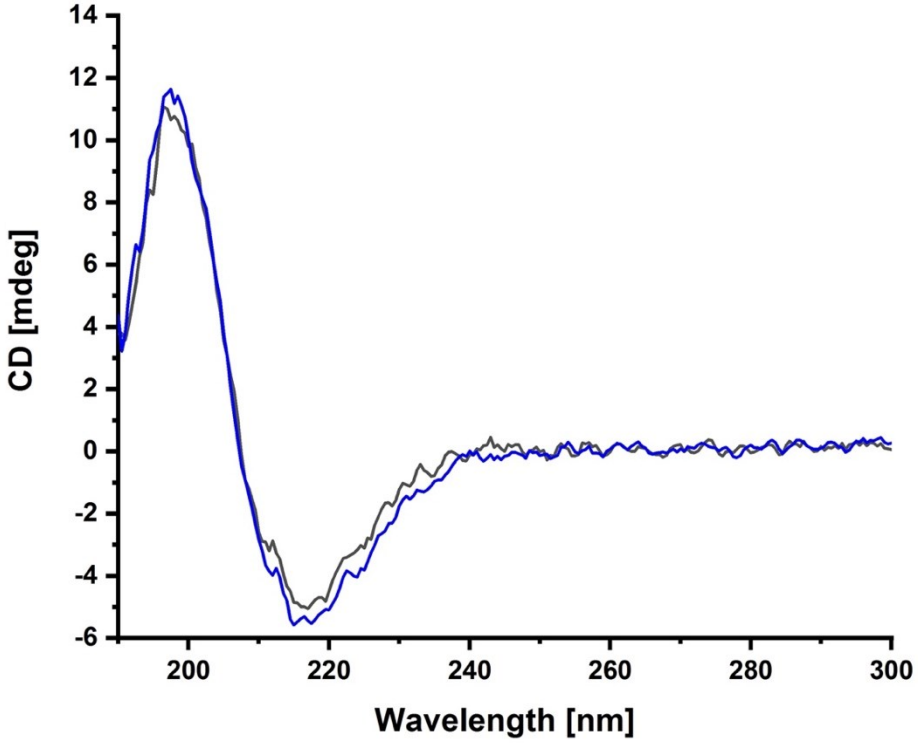

Figure S11: CD-spectra of apo NB4H (black) and MnPPIX@NB4H (blue) at pH = 10.5 in the range from 190-300 nm.

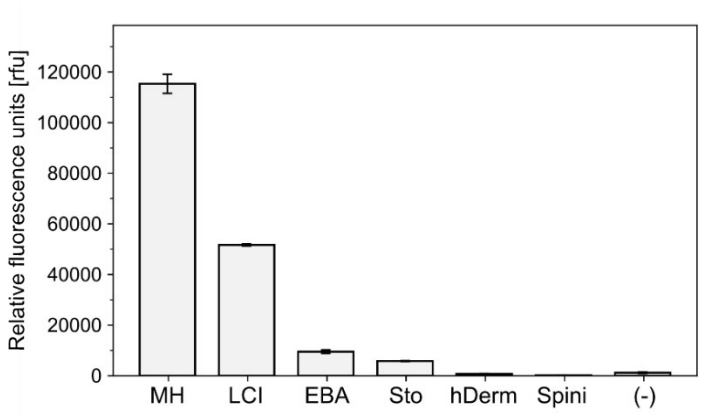

Figure S12: Binding of eGFP-adhesion-promoting peptide fusion proteins to *E. coli* cells. *E. coli* cells in 50 mM HEPES, pH 7 were incubated with the fusion proteins for 2 min, washed, and used for fluorescence spectroscopy. MH; Macaque histatin, LCI; liquid chromatography peak I, EBA; European bumblebee abaecin, Sto; Stomoxyn, hDerm; hDermcidin, Spini; Spinigerin, (-); eGFP without any adhesion-promoting peptide.

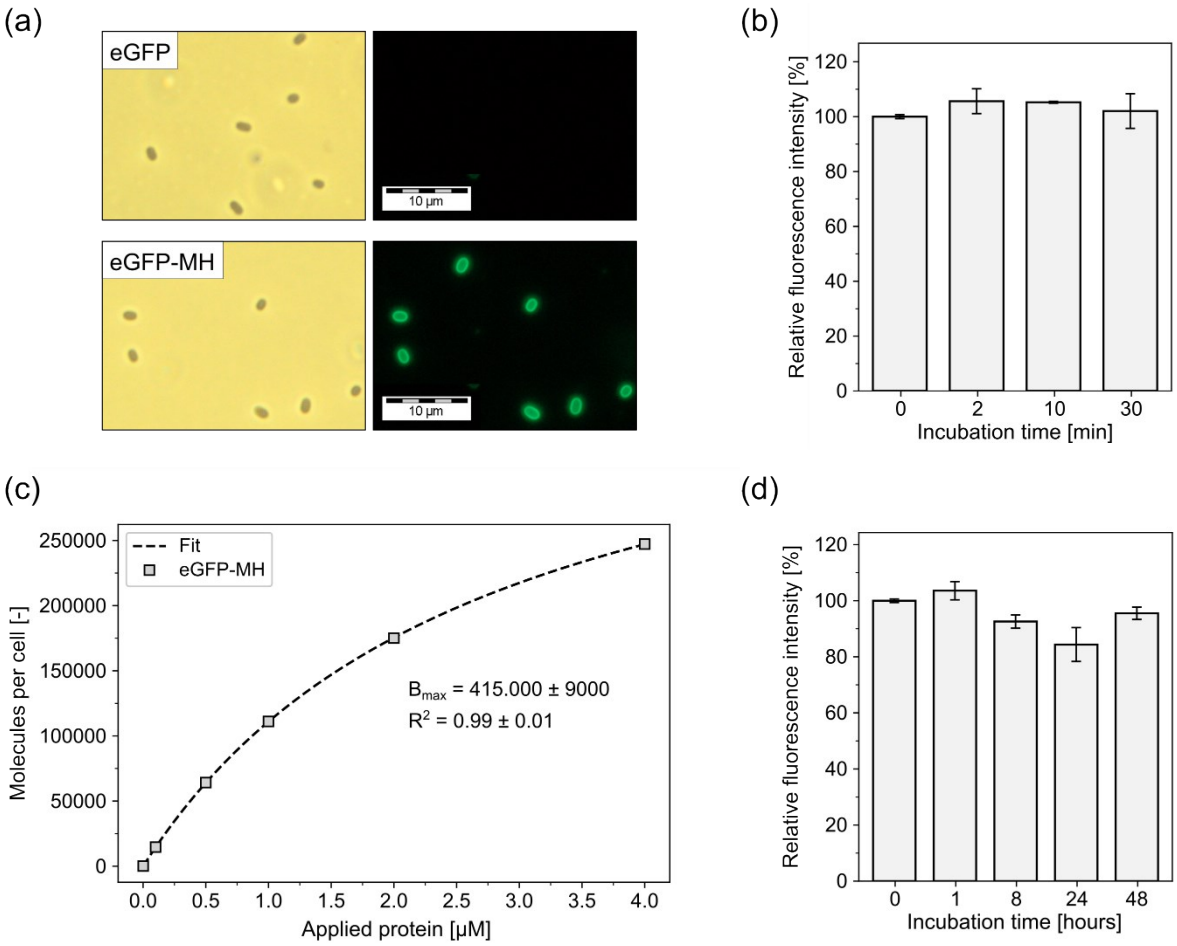

Figure S13: Characterization of the binding of eGFP-MH to *E. coli* cells. (a) Fluorescence microscopy pictures of *E. coli* cells decorated with eGFP-MH. *E. coli* cells (in 50 mM HEPES; pH = 7) were incubated with eGFP or eGFP-MH, washed, and used for fluorescence microscopy. (b) Time required for eGFP-MH to attach to *E. coli* cells. *E. coli* cells (in 50 mM HEPES ; pH = 7) were incubated with eGFP-MH for various time points, washed, and used for fluorescence spectroscopy. The fluorescence intensity for no incubation time (t = 0 min) was set to 100%. (c) Estimation of the maximum cell load (B_{max}). *E. coli* cells (in 50 mM HEPES ; pH = 7) were incubated with various amounts of eGFP-MH, washed, and used for fluorescence spectroscopy. To estimate B_{max}, curve-fitting was performed. (d) Determination of the amount of eGFP-MH that remains bound to *E. coli* cells over time. *E. coli* cells (in 50 mM HEPES ; pH = 7) were incubated with eGFP-MH, washed, and again incubated for various time points in fresh buffer. Following incubation, cells were washed again and used for fluorescence spectroscopy. The fluorescence intensity for no incubation time (t = 0 min) was set to 100%.

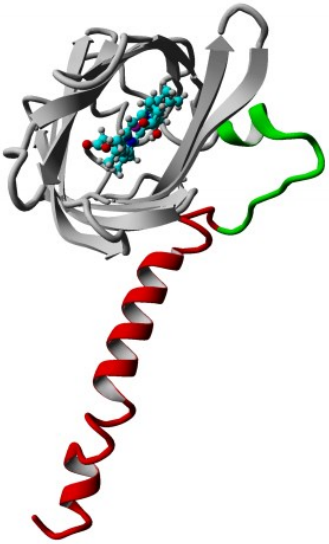

Figure S14: Illustration of MnPPIX@NB4H-MH. Homology modeling was performed by YASARA Structure version 17.4.17.[22] Colour coding: NB4 (gray), MnPPIX (in ball-stick), linker structure (including a 17 amino acid helix spacer and a TEV cleavage site; green), macaque histatin (red).

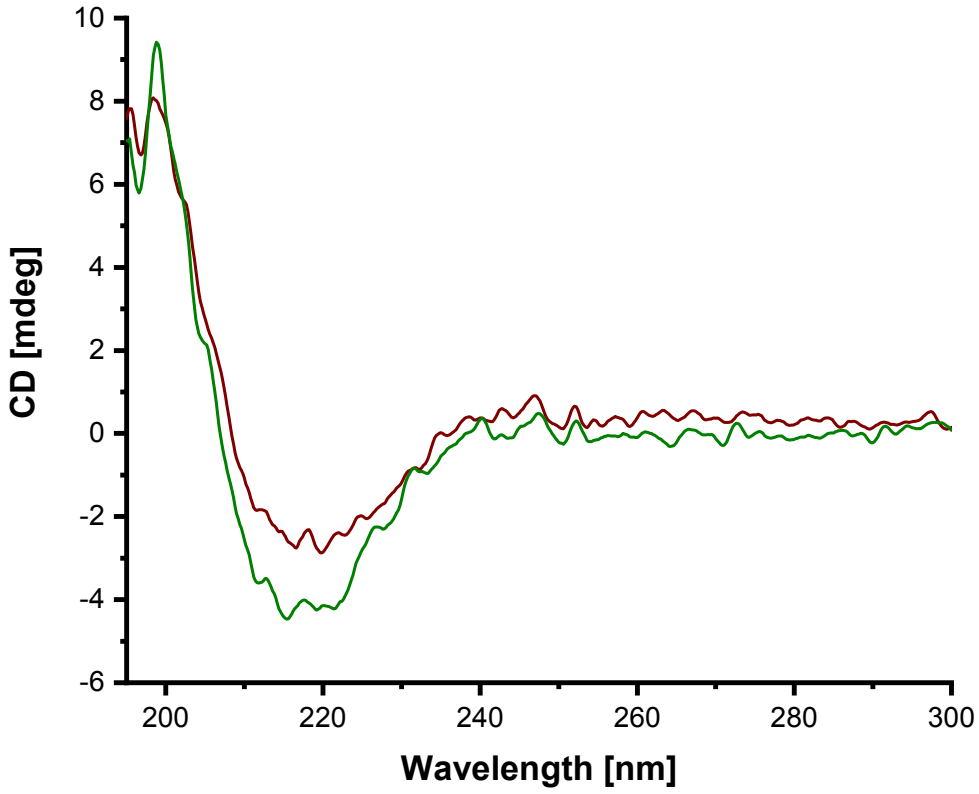

Figure S15: CD-spectra of apo NB4H-MH (wine red) and MnPPIX@NB4H-MH (green) in the range from 190-260 nm. The CD spectra indicate structural integrity of the protein scaffold upon protein conjugation.

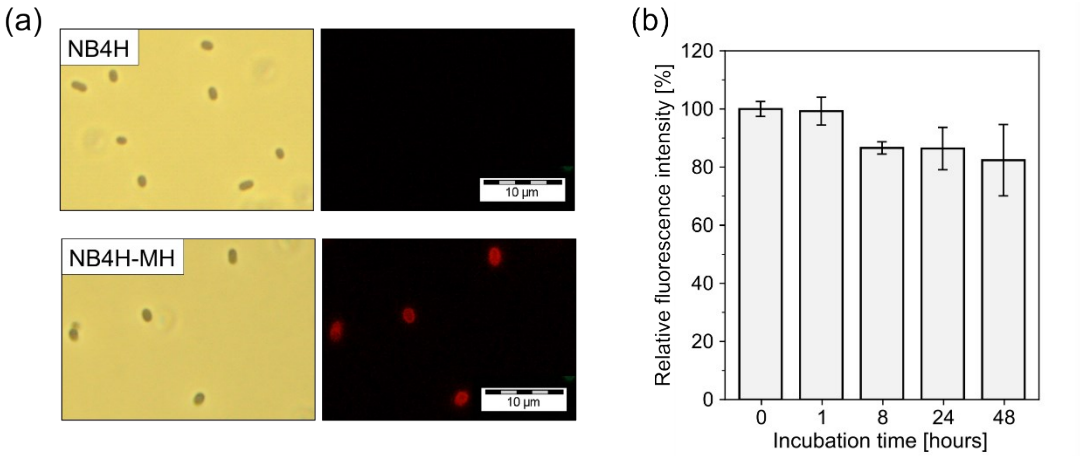

Figure S16: (a) Fluorescence microscopy pictures of *E. coli* cells decorated with NB4H-MH. *E. coli* cells were incubated with NB4H-MH or NB4H, washed, and used for fluorescence microscopy. NB4H(-MH) was fluorescently labeled with the Strep-Tactin Chromeo 546 dye. (b) Determination of the amount of NB4H-MH that remains bound to *E. coli* cells over time. *E. coli* cells were incubated with NB4H-MH, washed, and again incubated for various time points in fresh buffer. Following incubation, cells were washed again and used for fluorescence spectroscopy. The fluorescence intensity for no incubation time (t = 0 min) was set to 100%. NB4H-MH was fluorescently labeled with the Strep-Tactin Chromeo 546 dye.

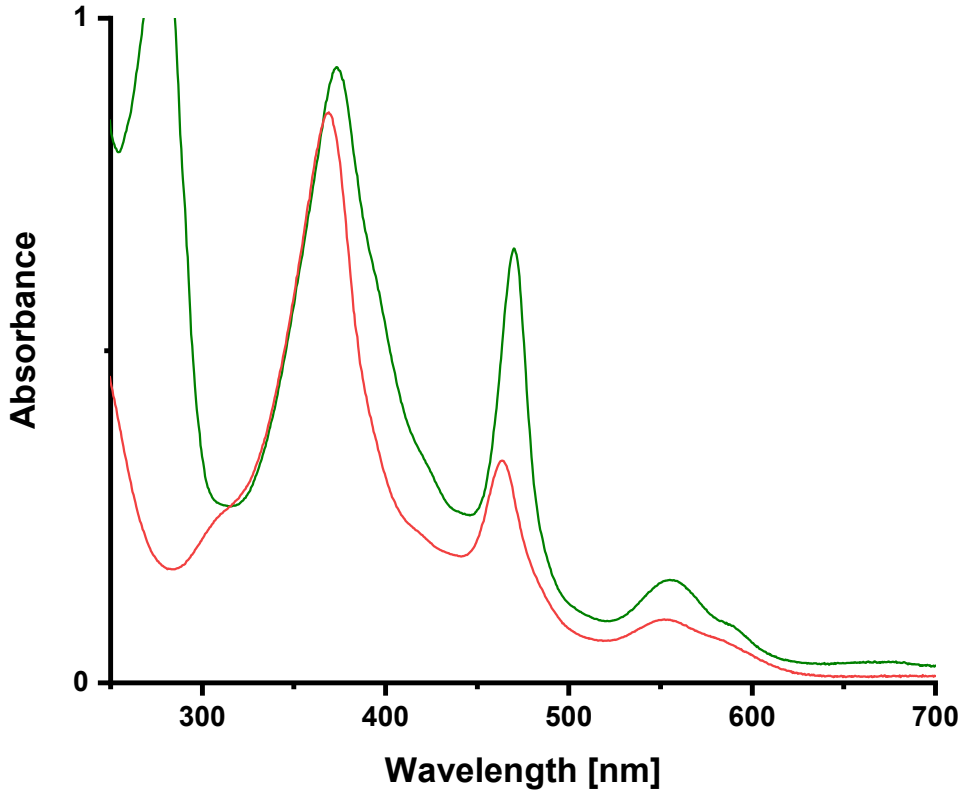

Figure S17: UV/Vis spectra of MnPPIX@NB4H-MH (green) and MnPPIX (red). The shift of the Soret bands from 368 nm to 374 nm and 464 nm to 471 nm as well as the increase in absorbance of the Soret band at 471 nm indicate the interaction of MnPPIX with the protein scaffold NB4H-MH.

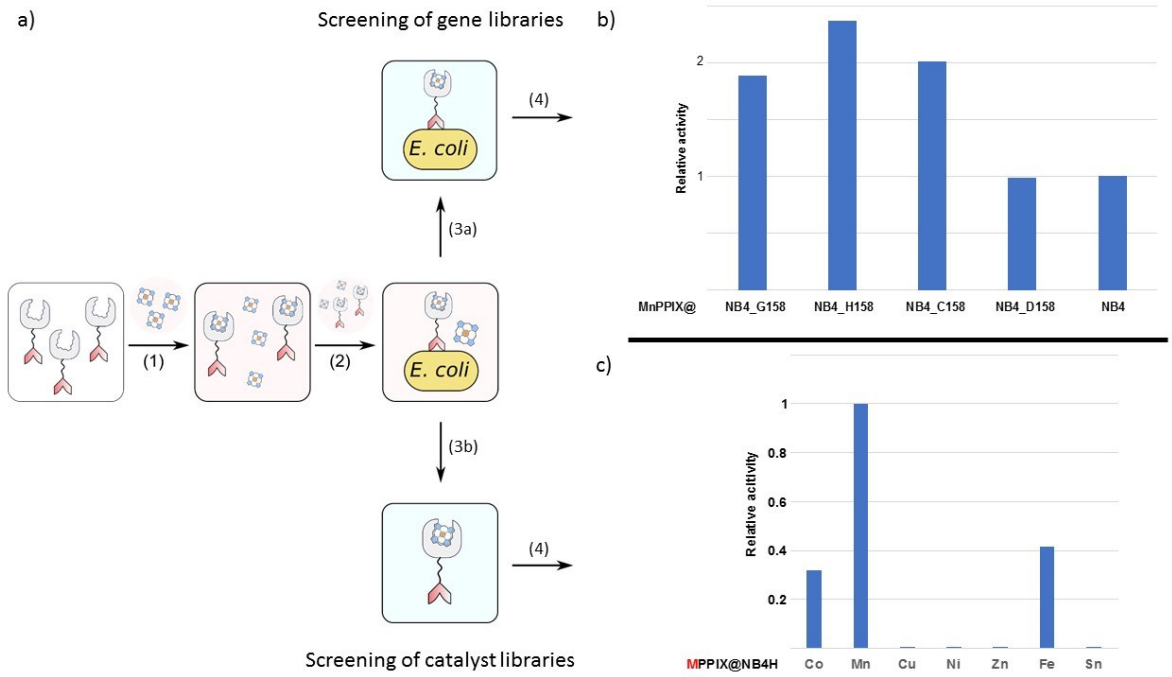

Figure S18: Screening of ArM with the adhesion-promoting peptide technology. (a) Schematic depiction of the screening procedure. The ArM scaffold is shown in gray. The adhesion-promoting peptide is shown in red. For the screening of protein variants, the starting point is cell-free lysates of cells producing the ArM scaffold-adhesion-promoting peptide fusion protein. The starting point for metal catalyst screening is the purified ArM scaffold-adhesion-promoting peptide fusion protein. In the first step (1) the cofactor is added to assemble the ArM. Following incubation, the solution is added to *E. coli* "carrier cells" (2). The assembled ArM attaches to cells via the adhesion-promoting peptide. Next, cells are washed by sedimentation and resuspension to remove unbound cofactor. Following washing, the cell/ArM solution can be used for screening (3a). Alternatively, cells can be washed and resuspended in a buffer capable of detaching the ArM (*e.g.,* 100 mM NaPi buffer pH = 8). Following pelleting of the "carrier cells" by centrifugation, the supernatant (containing the detached ArM) can be used for screening (3b). (b) Rescreening results after screening the NRT library with the whole-cell MH technology. MnPPIX@NB4H was the best variant, confirming the results of the screening on Strep-Tactin beads. (c) Screening results of the different Metal-PPIX cofactors attached to NB4H-MH. MnPPIX remains the best cofactor for the epoxidation reaction.

While an efficient screening approach accelerates the optimization of the scaffold protein, it might as well be used for the selection of the most suitable metal cofactor during the chemogenetic optimization of an ArM. In a further proof-of-principle experiment, seven different metal-containing protoporphyrin cofactors (MnPPIX, FePPIX, CoPPIX, NiPPIX, CuPPIX, ZnPPIX, and SnPPIX) were screened using the NB4H-MH scaffold and *E. coli* as carrier cells. The cofactor was added to the purified proteins, followed by addition to carrier cells. Cells were washed, resuspended in 100 mM NaPi-buffer (pH = 8) causing the detachment of the ArM, and then tested (Figures S18a and S19). Here, MnPPIX was identified as be the best metal cofactor for the epoxidation reaction, followed by FePPIX and CoPPIX (Figure S18c). In the future, gene libraries and cofactor libraries could be screened combinatorially.

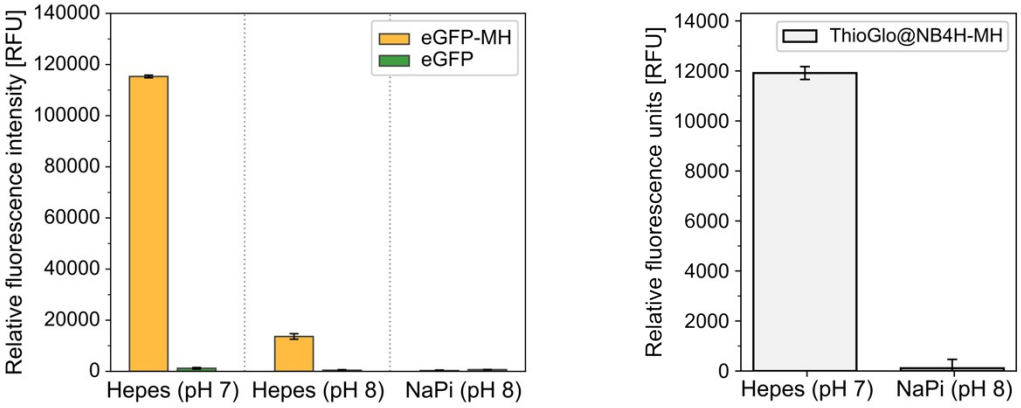

Figure S19: pH-dependent attachment of MH to *E. coli* cells. *E. coli* cells were incubated with eGFP, eGFP-MH or ThioGlo@NB4H-MH in the buffer denoted below the chart, washed, and used for fluorescence spectroscopy.

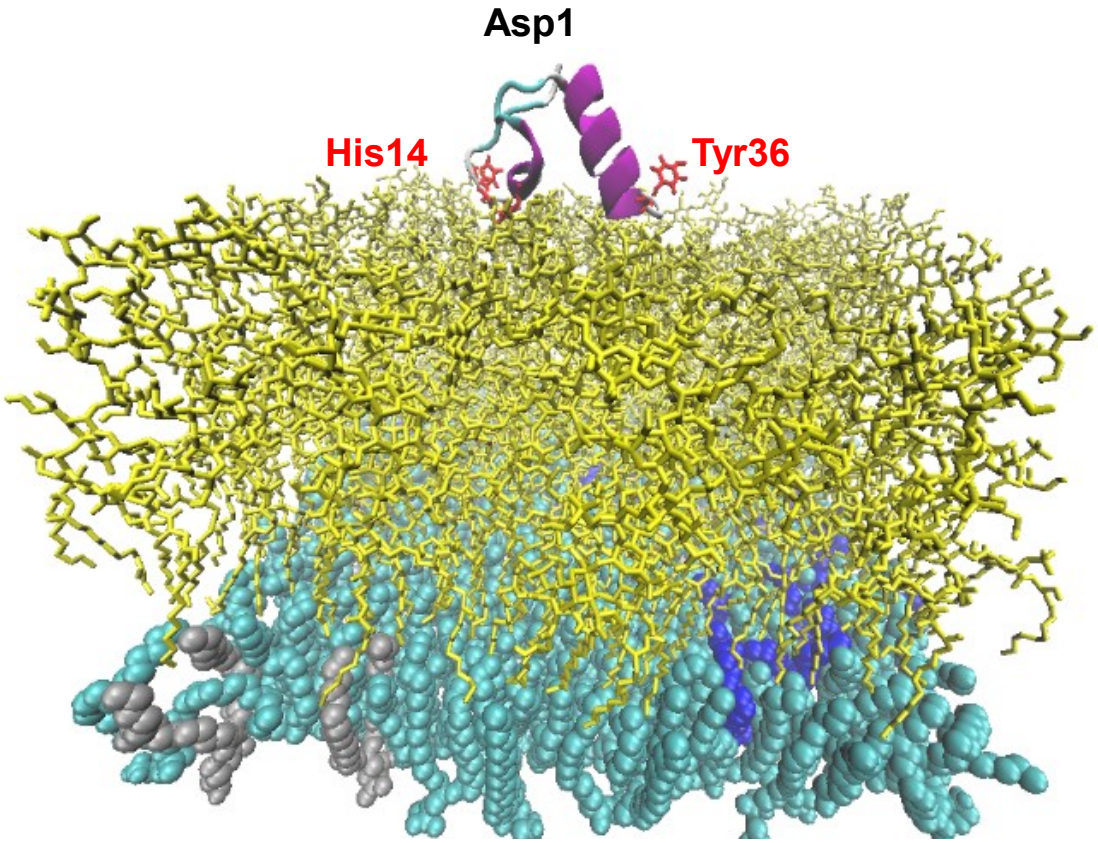

Figure S20: Binding orientation of MH on the E. coli outer membrane. Representative snapshot from the MD simulations of MH on the E. coli outer membrane. Surface binding of the MH molecule is initiated via anchoring of histidine residues (His8 and His14; depicted as red sticks). Miscellaneous lipids of the membrane are shown as cyan, silver, and blue balls; LPS sugars are depicted as yellow sticks (details of the membrane composition are given in the Supporting Information).

To gain insights into the remarkable binding properties of the MH-peptide with the *E. coli* cell surface and to understand the binding mechanism to the membrane, all-atom molecular dynamics (MD) simulations with an explicit solvent model were performed. The validated atomistic *E. coli* membrane developed by Piggot *et al.*[31] was used for the simulation of interactions between the peptide and the

membrane. In this model, the outer leaflet of the E. coli outer membrane bilayer is composed of the Rd1 lipopolysaccharide (LPS) molecules, which comprises lipid A and the inner core of the LPS (see computational methods in SI). The MD trajectories analysis shows that MH interacts mainly with the E. coli outer membrane through interaction between aromatic residues (His and Tyr) and LPS molecules (Figure S20). The surface-binding of the MH molecule is initiated *via* anchoring of histidine residues (His8 and His14). A complete membrane-bound state is found to be stable due to the interaction of three regions of MH including contact residues His7, His8, His14, and Tyr36. The *N*-terminus (Asp1, Figure 20) is facing away from the membrane toward the bulk solution. The fusion partners (*i.e.,* eGFP and NB4H) are connected to that *N*-terminus with spacing units, spatially separating the proteins from the MH moiety. We therefore conclude that the binding behaviour of MH remained unchanged with a fusion partner attached to the *N*-terminus.

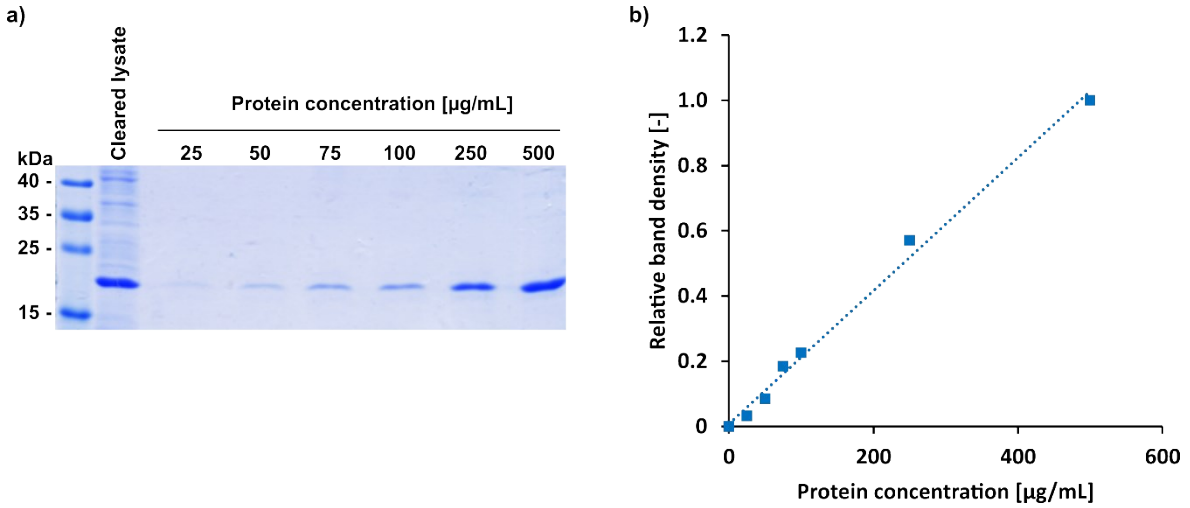

Figure S21: Quantification of NB4 in cleared lysates. a) SDS-PAGE analysis of a representative sample of cleared lysate containing NB4 (19.5 kDa). As standards, purified NB4 was used at different concentrations. b) Densitometric analysis revealed typical scaffold protein titers of ~20 µM in cleared lysates for NB4 and its variants. Densitometric analysis was done using *ImageJ* (version 1.52a).

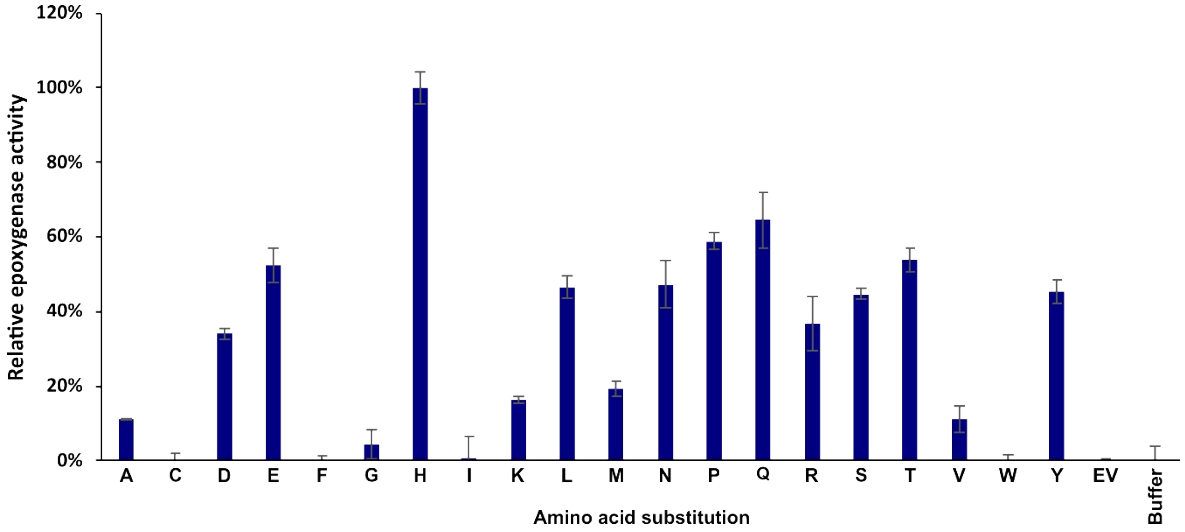

Figure S22: Results of the first round of directed evolution (SSM on position 158). The parent variant of this round (NB4) carries a leucine (L) at position 158. In addition to the library members, the "empty vector control" (EV) and the "buffer control" (Buffer) are depicted.

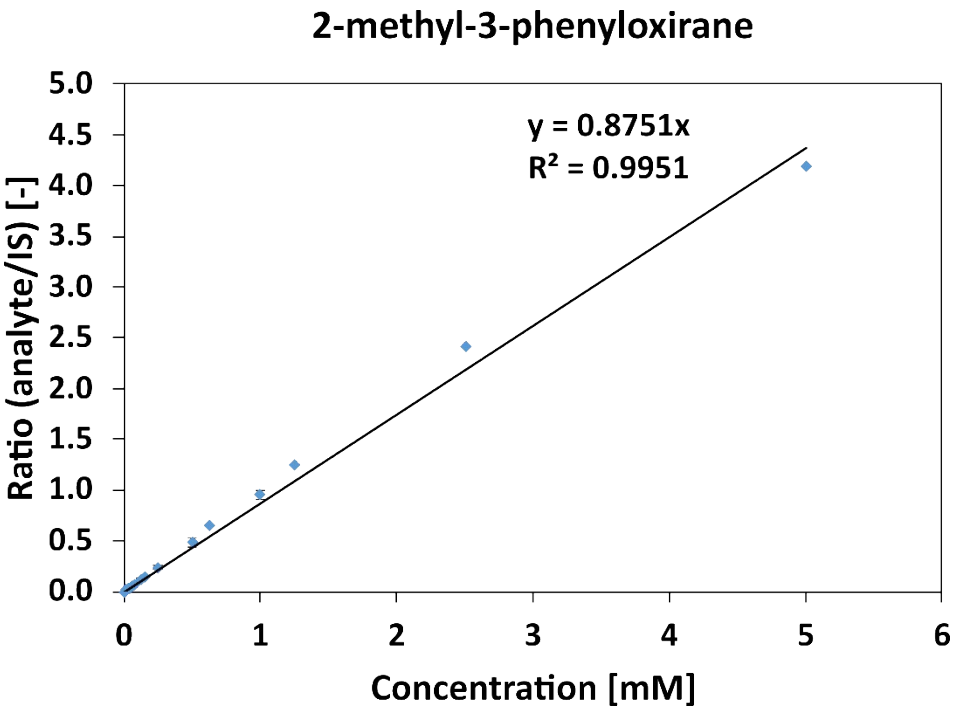

Figure S23: Calibration curve for 2-methyl-3-phenyloxirane.

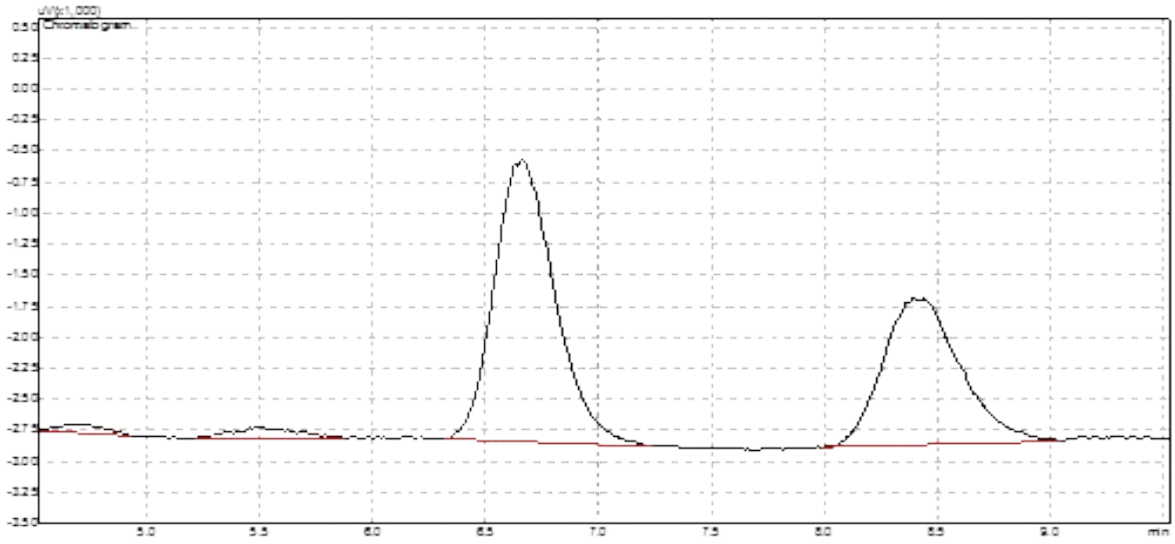

Figure S24: Representative GC-FID chromatogram for the epoxidation of *trans*-β-methylstyrene catalyzed by MnPPIX@NB4HA. The ratio of the signals at 6.66 and 8.44 minutes is 1.5/1. This corresponds to an *ee* of 20% for the (*S,S*) enantiomer.

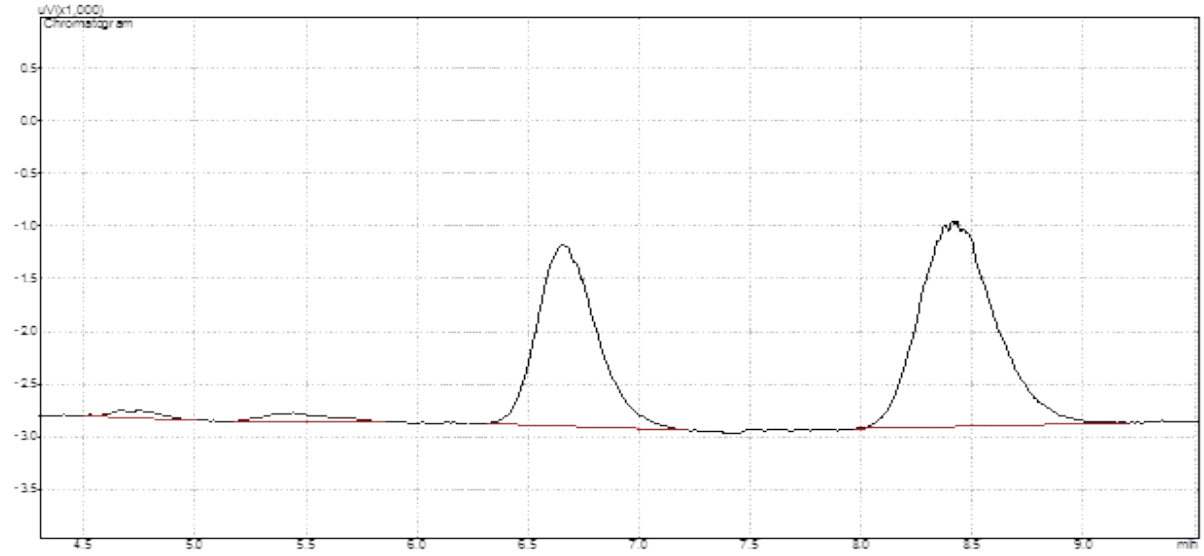

Figure S25: Representative GC-FID chromatogram for the epoxidation of *trans*-β-methylstyrene catalyzed by MnPPIX@NB4H. The ratio of the signals at 6.66 and 8.44 minutes is 1/1.5. This corresponds to an *ee* of 20% for the (*R,R*) enantiomer.

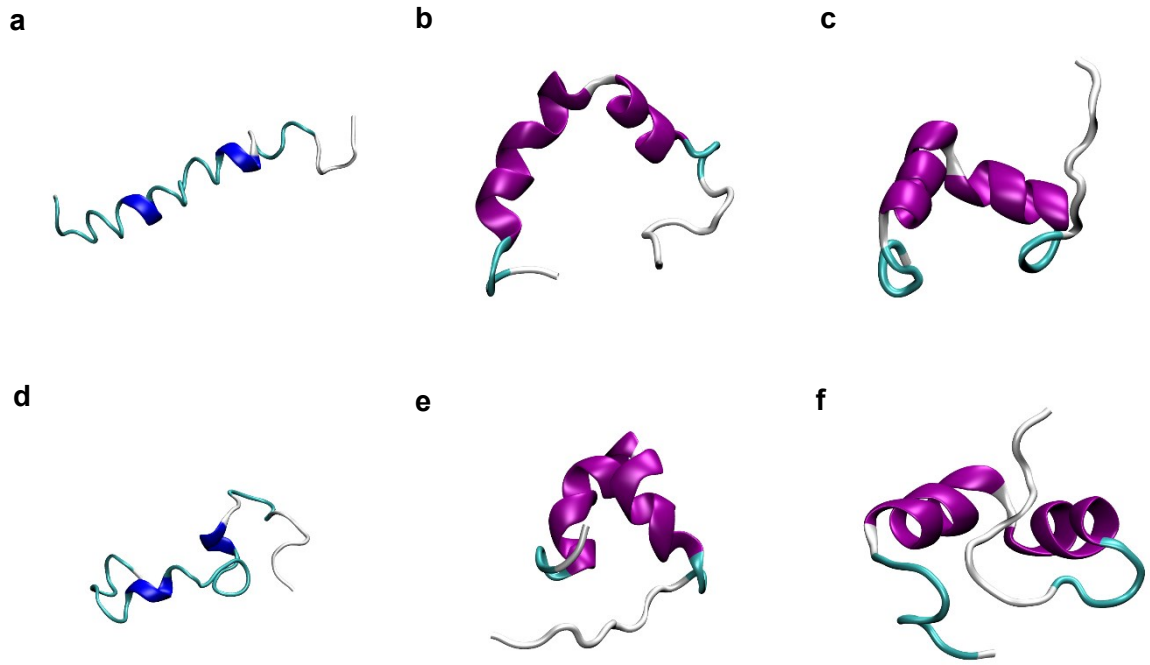

Figure S26: Structure of MH in the three force fields AMBER03, AMBER99SB-ILDN, and GROMOS54A7 (**a**)-(**c**) at the beginning of the MD simulation (t_0) and (**d**)-(**f**) at the end of the MD simulation after 100 ns (t_{100}). The colors indicate the secondary structure of the peptide. Cyan: Turn; Magenta: α-helix; Blue: 3-10 helix; light gray: Coil.

III. DNA and amino acid sequences

DNA sequence of the NB4 open reading frame

Codons of the amino acid positions 76 and 158 are highlighted.

ATGTGGAGCCACCCGCAGTTCGAAAAAAATCAACTGCAACAACTGCAAAATCCGGGCGAGAGTCCG
CCGGTTCATCCGTTCGTGGCACCGCTGTCCTATCTGCTGGGTACCTGGCGCGGCCAGGGTGAAGGCG
AGTATCCGACCATTCCGAGCTTTCGCTATGGCGAAGAGATCCGTTTCAGCCATTCGGGTAAACCGGT
GATTGCCTATACCCAAAAAACGTGGAAACTGGAATCGGGTGCACCGCTGCTGGCAGAGAGTGGTTA
TTTTCGCCCGCGTCCGGATGGTTCTATTGAAGTGGTTATCGCATGCTCGACCGGTCTGGTGGAAGTTC
AAAAAGGCACGTATAATGTGGATGAGCAGAGTATTAAACTGAAATCTGACCTGGTGGGCAACGCGT
CCAAAGTTAAAGAAATCAGCCGCGAATTCGAGCTGGTTGACGGTAAACTGAGTTATGTGGTTCGTCT
GAGCACGACCACGAATCCGCTGCAACCGCTGCTGAAAGCCATCCTGGACAAACTGTAA

DNA sequence of the NB4H-MH open reading frame

The codon of amino acid position 158 is highlighted in green. The sequence of macaque histatin is highlighted in teal.

ATGTGGAGCCACCCGCAGTTCGAAAAAAATCAACTGCAACAACTGCAAAATCCGGGCGAGAGTCCG
CCGGTTCATCCGTTCGTGGCACCGCTGTCCTATCTGCTGGGTACCTGGCGCGGCCAGGGTGAAGGCG
AGTATCCGACCATTCCGAGCTTTCGCTATGGCGAAGAGATCCGTTTCAGCCATTCGGGTAAACCGGT
GATTGCCTATACCCAAAAAACGTGGAAACTGGAATCGGGTGCACCGCTGCTGGCAGAGAGTGGTTA
TTTTCGCCCGCGTCCGGATGGTTCTATTGAAGTGGTTATCGCATGCTCGACCGGTCTGGTGGAAGTTC
AAAAAGGCACGTATAATGTGGATGAGCAGAGTATTAAACTGAAATCTGACCTGGTGGGCAACGCGT
CCAAAGTTAAAGAAATCAGCCGCGAATTCGAGCTGGTTGACGGTAAACTGAGTTATGTGGTTCGTCT
GAGCACGACCACGAATCCGCTGCAACCGCACCTGAAAGCCATCCTGGACAAACTGGGCGGTGGCGG
TAGCGGCGGTGGCGGTAGCGCTGAAGCAGCTGCCAAAGAAGCGGCAGCGAAAGAAGCGGCGGCC
AAAGCCGAGAATCTGTACTTTCAGGGTGGCGGTGGCGGTAGCGGCGGTGGCGGTAGCGGCGGTGG
CGGTAGCGGCGGTGGCGGTAGCGATTCTCACGAAGAACGCCATCATGGTCGTCATGGTCACCACAA
GTATGGCCGCAAATTCCACGAGAAACATCACAGTCATCGTGGCTATCGCTCGAACTATCTGTACGAC
AACTGA

DNA sequences of adhesion-promoting peptides

Sequences were published previously.[6]

IV. References

1. L. J. Boucher, *J. Am. Chem. Soc.*, 1968, **90**, 6640.
2. K. Fukumoto, A. Onoda, E. Mizohata, M. Bocola, T. Inoue, U. Schwaneberg and T. Hayashi, *ChemCatChem*, 2014, **6**, 1229.
3. 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.
4. D. G. Gibson, L. Young, R.-Y. Chuang, J. C. Venter, C. A. Hutchison and H. O. Smith, *Nat. Methods*, 2009, **6**, 343.
5. A. R. Grimm, D. F. Sauer, T. Polen, L. Zhu, T. Hayashi, J. Okuda and U. Schwaneberg, *ACS Catal.*, 2018, **8**, 2611.
6. R. A. Meurer, S. Kemper, S. Knopp, T. Eichert, F. Jakob, H. E. Goldbach, U. Schwaneberg and A. Pich, *Angew. Chem. Int. Ed.*, 2017, **56**, 7380.
7. H. M. Key, P. Dydio, D. S. Clark and J. F. Hartwig, *Nature*, 2016, **534**, 534.
8. W. Kabsch, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 125.
9. C. Vonrhein, C. Flensburg, P. Keller, A. Sharff, O. Smart, W. Paciorek, T. Womack and G. Bricogne, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 293.
10. M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. Leslie, A. McCoy, S. J. McNicholas, G. N. Murshudov, N. S. Pannu, E. A. Potterton, H. R. Powell, R. J. Read, A. Vagin and K. S. Wilson, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 235.
11. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkoczi, V. B. Chen, T. I. Croll, B. Hintze, L. W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams and P. D. Adams, *Acta Crystallogr. D Struct. Biol.*, 2019, **75**, 861.
12. A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni and R. J. Read, *J. Appl. Crystallogr.*, 2007, **40**, 658.
13. K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.*, 2006, **62**, 1002.
14. P. Emsley, B. Lohkamp, W. G. Scott and K. Cowtan, *Acta Crystallogr. D Biol. Crystallogr.*, 2010, **66**, 486.
15. G. N. Murshudov, P. Skubak, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long and A. A. Vagin, *Acta Crystallogr. D Biol. Crystallogr.*, 2011, **67**, 355.
16. P. V. Afonine, R. W. Grosse-Kunstleve, N. Echols, J. J. Headd, N. W. Moriarty, M. Mustyakimov, T. C. Terwilliger, A. Urzhumtsev, P. H. Zwart and P. D. Adams, *Acta Crystallogr. D Biol. Crystallogr.*, 2012, **68**, 352.
17. N. W. Moriarty, R. W. Grosse-Kunstleve and P. D. Adams, *Acta Crystallogr. D Biol. Crystallogr.*, 2009, **65**, 1074.
18. *The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.*
19. J. A. Myers, B. S. Curtis and W. R. Curtis, *BMC Biophysics*, 2013, **6**, 4.
20. J. Zhao, J. G. Rebelein, H. Mallin, C. Trindler, M. M. Pellizzoni and T. R. Ward, *J. Am. Chem. Soc.*, 2018, **140**, 13171.
21. J. G. Rebelein, Y. Cotelle, B. Garabedian and T. R. Ward, *ACS Catal*, 2019, **9**, 4173.
22. E. Krieger, G. Koraimann and G. Vriend, *Proteins*, 2002, **47**, 393.
23. C. M. Bianchetti, G. C. Blouin, E. Bitto, J. S. Olson and G. N. Phillips, Jr., *Proteins*, 2010, **78**, 917.
24. A. A. Canutescu, A. A. Shelenkov and R. L. Dunbrack Jr., *Protein Sci.*, 2003, **12**, 2001.
25. S. Miyamoto and P. A. Kollman, *J. Comput. Chem.*, 1992, **13**, 952.
26. J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser and C. Simmerling, *J. Chem. Theory Comput.*, 2015, **11**, 3696.

27. J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157.
28. J. Yang, R. Yan, A. Roy, D. Xu, J. Poisson and Y. Zhang, *Nat. methods*, 2015, **12**, 7.
29. M. H. Olsson, C. R. Søndergaard, M. Rostkowski and J. H. Jensen, *J. Chem. Theory Comput.*, 2011, **7**, 525.
30. K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror and D. E. Shaw, *Proteins*, 2010, **78**, 1950.
31. N. Schmid, A. P. Eichenberger, A. Choutko, S. Riniker, M. Winger, A. E. Mark and W. F. van Gunsteren, *Eur. Biophys. J.*, 2011, **40**, 843.
32. T. J. Piggot, D. A. Holdbrook and S. Khalid, *J. Phys. Chem., B*, 2011, **115**, 13381.
33. H. Berendsen, J. Grigera and T. Straatsma, *J. Phys. Chem. A*, 1987, **91**, 6269.
34. U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee and L. G. Pedersen, *J. Chem. Phys*, 1995, **103**, 8577.
35. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19.
36. W. Humphrey, A. Dalke and K. Schulten, *J. Mol. Graphics*, 1996, **14**, 33.
37. A. R. Grimm, D. F. Sauer, M. D. Davari, L. Zhu, M. Bocola, S. Kato, A. Onoda, T. Hayashi, J. Okuda and U. Schwaneberg, *ACS Catal.*, 2018, **8**, 3358.