Electronic Supplementary Material (ESI) for Green Chemistry.
This journal is © The Royal Society of Chemistry 2021

Supporting Information

Design of a Synthetic Enzyme Cascade for the *in vitro* Fixation of a C_1 Carbon Source to a Functional C_4 Sugar

Samed Güner,[a] Vanessa Wegat,[ab] Andre Pick,[c] and Volker Sieber*[abde]

[a] Chair of Chemistry of Biogenic Resources, Campus Straubing for Biotechnology and Sustainability, Technical University of Munich, Schulgasse 16, 94315 Straubing, Germany, E-mail: sieber@tum.de

[b] Straubing Branch BioCat, Fraunhofer IGB, Schulgasse 11a, 94315 Straubing, Germany

[c] CASCAT GmbH, Europaring 4, 94315 Straubing, Germany

[d] Catalytic Research Center, Technical University of Munich, Ernst-Otto-Fischer-Straße 1, 85748 Garching, Germany

[e] School of Chemistry and Molecular Biosciences, The University of Queensland, 68 Copper Road, St. Lucia 4072, Australia

Table of Contents

Experimental section

Reagents

Formaldehyde (FALD) was purchased from Carl Roth (Karlsruhe, Germany). Glycolaldehyde (GALD), dihydroxyacetone (DHA) and erythrulose (ERY) were purchased from Merck (Darmstadt, Germany). All enzymes for molecular cloning were purchased from New England Biolabs (Frankfurt, Germany). Genomic DNA was obtained from German Collection of Microorganisms and Cell Cultures (DSMZ). All oligonucleotides were purchased from Eurofins Genomics (Ebersberg, Germany). All other chemicals, unless otherwise indicated, were of analytical grade and commercially available.

Cloning, Expression and Purification of FLS

The formolase (FLS) "wild type" gene was synthesized as reported[1] by GeneArt (Regensburg, Germany) and cloned into the expression vector pET24a between NdeI and XhoI. This resulted in a C-terminal hexahistidine-tagged FLS. The plasmid was used to transform *E. coli* DH5α (Invitrogen, Germany) and correct sequence was confirmed by sequencing (Eurofins, Germany). Recombinant *E. coli* BL21 (DE3) harboring pET24a_FLS_CHIS$_6$ was cultured at 37° C in TB (Terrific Broth) media supplemented with 50 µg mL^{-1} kanamycin. When OD_{600} reached 0.8 the temperature was lowered to 25° C and 0.5 mM IPTG was added to induce protein expression. The cells were harvested after 24 h by centrifugation and stored at -80° C until further purification. The cell pellet was dissolved in binding buffer (50 mM sodium phosphate, pH 8.2, 500 mM sodium chloride, 5 mM imidazole). The cells were disrupted by sonication in the presence of lysozyme in an ice bath. The cell debris was removed by means of centrifugation at 40000 x *g* for 30 min. The supernatant was filtered through a 0.45 µm filter before application to Äkta Purifier. A Histrap column FF Crude 5 mL (GE Healthcare, Germany) was used to retain the target enzyme. The binding buffer was used to wash the column before elution with elution buffer (50 mM sodium phosphate, pH 8.2, 500 mM sodium chloride, 500 mM imidazole). Elution fraction was desalted (50 mM sodium phosphate, pH 8.2) using HiPrep desalting column 26/10 50 mL (Ge Healthcare, Germany). FLS was flash frozen in liquid nitrogen and stored at -80° C until further use.

Cloning, Expression, and Heat Purification of *Pfu*ADHd

The gene of the thermostable alcohol dehydrogenase ADHd (PF1960) was amplified as reported from the genomic DNA of *Pyrococcus furiosus*.[2] Using the oligonucleotides sense (5'-TATATACATATGAAAAGGGTAAATGCATTCAACGACCTTAAGC-3') and antisense (5'-TATATACTCGAGCACACACCTCCTTGCCATCTCTCTATCC-3'), containing the NdeI and XhoI restriction sites (underlined in the sequence). Native stop codon was removed to produce untagged protein. The purified gene was digested with NdeI-XhoI and cloned into pET24a (NdeI-XhoI-digested). The sequence of the expression clone was confirmed by sequencing. Recombinant *E. coli* BL21 (DE3) harboring pET24a_PfuADHd was cultured at 37° C in LB media supplemented with 50 µg mL^{-1} kanamycin. When OD_{600} reached 0.6, protein expression was induced with 0.5 mM IPTG. After incubation overnight at 37° C, cells were harvested by centrifugation. The cell pellet was dissolved in washing buffer (50 mM sodium phosphate, pH 8.2), and disrupted by sonication in the presence of lysozyme in an ice bath. The resulting cell lysate was heated for 30 min at 80° C, subsequently centrifuged for 30 min at 40000 xg. The supernatant was filtered through a 0.45 µm filter before desalting (100 mM potassium phosphate, pH 6.0) using HiPrep desalting column 26/10 50 mL (GE Healthcare, Germany).

Docking Studies

Docking studies were performed to identify promising target positons. The crystal structure of FLS (PDB ID: 4QQ8) was previously published with thiamine diphosphate (ThDP) as cofactor.[1] 2D structures of the ligands, such as dihydroxyethyl-ThDP, FALD, GALD, DHA and ERY were drawed by using ChemDraw Professional™ (Version 17.0). These structures were converted to 3D structures (.pdb) and energy minimized by using Chem3D™ (Version 17.0). Docking studies were performed by using YASARA[3], which rely on AutoDock VINA[4] to dock ligand to a receptor. The FLS as receptor (PDB ID: 4QQ8) was prepared according to YASARA standard protocol. *In silico* analysis was done by using PyMOL (Version 2.4.1).

In silico Engineering

FLS variants were generated *in silico* by using YASARA. Amino acid residues were swapped to desired amino acid and the protein was than energy minimized. Receptor was prepared according to YASARA standard protocol and used for further docking studies.

Double Mutant Library by Overlap Extension PCR

Oligonucleotides used for the generation of the semi-rational double mutant library of FLS are listed in Table S1. Overlap extension PCR protocol for introduction of degenerate bases at multiple codon locations was performed according to Williams *et al*., 2014.[5] Template used was pET24a_FLS_CHIS$_6$ with a length of 1749 bp. In order to saturate both positions simultaneously, three parts were first amplified individually (148, 276 and 1456 bp) and then combined using another PCR. The total fragment had a length of 1880 bp. Oligonucleotides 3a-e and 5a-e were mixed in advance in a ratio of 8:2:3:1:1 and then used as forward primer to amplify fragments 2 and 3, respectively. After verification by sequencing, the library was digested with NdeI-XhoI and was ligated again with pET24a.

Table S1: Oligonucleotides for overlap extension PCR (oePCR) of pET24a_FLS_CHIS$_6$.

Part	Name	Oligonucleotide sequence	DC*	Ratio
1	oePCR_1_T7-fwd	GAATTGTGAGCGGATAACAATTCCC	-	-
1	oePCR_2_Part1_rev	CCATGCAGGCCAAACAGATG	-	-
2	oePCR_3a_H29NRT_fwd	CATCTGTTTGGCCTGCATGGNRTTCATATTGACACCATTTTTC	NRT	8
2	oePCR_3b_H29ATS_fwd	CATCTGTTTGGCCTGCATGGATSTCATATTGACACCATTTTTC	ATS	2
2	oePCR_3c_H291VAA_fwd	CATCTGTTTGGCCTGCATGGVAATCATATTGACACCATTTTTC	VAA	3
2	oePCR_3d_H29CCG_fwd	CATCTGTTTGGCCTGCATGGCCGTCATATTGACACCATTTTTC	CCG	1
2	oePCR_3e_H29ACT_fwd	CATCTGTTTGGCCTGCATGGACTTCATATTGACACCATTTTTC	ACT	1
2	oePCR_4_Part2_rev	CAGGGTGTTGGTTTCGTCATC	-	-
3	oePCR_5a_Q113NRT_fwd	GATGACGAAACCAACACCCTGNRTGCCGGTATTGATCAGG	NRT	8
3	oePCR_5b_Q113ATS_fwd	GATGACGAAACCAACACCCTGATSGCCGGTATTGATCAGG	ATS	2
3	oePCR_5c_Q113VAA_fwd	GATGACGAAACCAACACCCTGVAAGCCGGTATTGATCAGG	VAA	3
3	oePCR_5d_Q113CCG_fwd	GATGACGAAACCAACACCCTGCCGGCCGGTATTGATCAGG	CCG	1
3	oePCR_5e_Q113ACT_fwd	GATGACGAAACCAACACCCTGACTGCCGGTATTGATCAGG	ACT	1
3	oePCR_6_T7_rev	CTTTGTTAGCAGCCGGATCTC	-	-

*DC = degenerated codon

In total, the library contained 225 variants at DNA and amino acid level. An overview of degenerate codons (DC) and the corresponding codons as well as amino acids (one letter code) used are listed in Table S2.

Table S2: Detailed overview of used degenerated codons.

DC*	Codons	Amino acids
NRT	AAT, AGT, TAT, TGT, GAT, GGT, CAT, CGT	N, S, Y, C, D, G, H, R
ATS	ATG, ATC	M, I
VAA	AAA, CAA, GAA	K, Q, E
CCG	CCG	P
ACT	ACT	T

Determination of Protein Concentrations

The concentration of purified FLS and variants were determined by UV spectroscopy using a nanophotometer (P-class, Implen, USA). The molecular weight (M_W) as well as the extinction coefficient at 280 nm (ε_{280}) of all proteins were calculated with the ProtParam tool.[6] Hereby, the His_6-tag was considered for all proteins.

SDS-PAGE

Sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS-PAGE) was performed for protein separation as described by Ausubel and coworkers with modifications.[7] For sample preparation, 40 µl protein sample (approximately 1 mg mL^{-1}) was mixed with 10 µl 5x SDS loading buffer, spun down, and incubated for 10 min at 95° C. Then, 7.5 µl per sample were loaded on a 12% SDS-gel. Mini-PROTEAN® Tetra Vertical Electrophoresis Cell system was used, and the gel was run for about 45 min with 45 mA for each gel.

Thermal Shift Assay (Thermofluor Assay)

Protein melting points (T_m) were determined using the thermofluor assay. Here, SYPRO™ orange was used as fluorescent dye and instructions of Bovin *et al.* were followed.[8] The dye interacts with the hydrophobic core of the protein, which is exposed by unfolding under denaturing conditions. At this point, the dye becomes fluorescent, resulting in a sharp sigmoidal curve in the detection unit. The inflection point of the sigmoidal curve represents the melting temperature (T_m) at which 50 % of the protein is unfolded. An alternative representation is to plot the first derivative (dRFU dT^{-1}), where the apex equals T_m. Total well volume was 25 µL, containing 2 µL of 1 mg mL^{-1} enzyme, 2 µL of SYPRO Orange working stock and 21 µL of buffer. The temperature gradient was set to begin at 5° C and rise until 95° C in steps of 0.5° C per 30 s.

Biocatalysis

Biocatalysis of ERY directly from GALD was investigated using FLS_wt, FLS_B1 and FLS_B2 in comparison. Purified enzymes (stored at -80° C) were thawed at room temperature and concentrated using 10 kDa centrifugation filters (10 min, 10.000 x *g*). Reactions were prepared as duplicates (200 µL approach) in a 96-MTP. A aluminum sealing was used to avoid evaporation. The reaction mixture was prepared in 50 mM sodium phosphate buffer, pH 8.0 and contained 150 µM of enzyme, 25 g L^{-1} GALD, 2 mM $MgSO_4$ and 0.1 mM ThDP. After pre-incubation for 10 min at 30° C, reactions were started with additions of enzymes. Reactions were incubated at 30° C and 750 rpm shaking. Samples were taken after 0, 2, 4, 8 and 16 h and were analyzed by HPLC. The 96-MTP was sealed with an aluminum sealing.

Cascade Reaction

Reactions were performed in 1.5 mL reaction tubes with a total volume of 200 µL. A 1:1 ratio (45 µM each) of purified enzymes were applied to 3.0 g L^{-1} FALD in 50 mM sodium phosphate buffer (pH 8.2) supplemented with 2 mM $MgCl_2$ and 0.1 mM ThDP and were incubated for 22 h at 30° C. In doing so, FLS_A1, FLS_A2 or FLS_A3 was combined with FLS_B1 or FLS_B2, respectively. Analysis was performed by HPLC.

HPLC Analytics

For analysis, 50 µL of each sample was withdrawn and diluted 1:10 with 5.0 mM H_2SO_4. The mixtures were then filtered through 10 kDa filters. The filtrates were subjected to HPLC analysis. FALD, GALD, DHA and ERY were separated using an ion-exclusion column (Rezex™ ROA-Organic Acid H+ (8%), Phenomenex, Germany), run isocratically using 2.5 mM H_2SO_4 at 70° C for 20 min.

Preliminary experiments

Conversion of Glycolaldehyde to Erythrulose

To determine whether the FLS “wild type” is capable of converting GALD to ERY, 10 µM (0.61 mg mL^{-1}) of FLS was applied to 20 mM GALD solely. Reaction mixtures (200 µL) were prepared in 50 mM sodium phosphate, pH 8.2 and supplemented with 2 mM $MgSO_4$ and 0.1 mM ThDP. As a control, enzyme addition was omitted. The samples were incubated for 60 h at room temperature, without shaking. Analysis was performed via HPLC. No change was determined for the control reaction. FLS showed partial consumption of GALD and a new peak appeared that overlaid with the standard for erythrulose.

Evaluation of the Active Site of FLS

The thiamine diphosphate-dependent (ThDP) benzaldehyde lyase was redesigned by protein engineering for the production of DHA from FALD. Therefore, several positions within the active pocket of the FLS have already been modified. In order to identify promising target positons, the focus was on amino acid residues in the active center that had not yet been investigated. The active site of the enzyme is formed at the interface of the monomers of the homodimer. The cofactor ThDP is thereby embedded (Figure S1-1A/B). A major part of the positions has already been investigated in the context of increased DHA production. Marked in green (Figure S1-2A/B) are I29, G394, N419 and W480 which showed a significant impact for the production of DHA. Marked in blue (Figure S1-3A/B) are non-investigated positions including H29 and Q113. Certain parts of secondary structure elements, in particular glycines, were also not changed. These are highlighted in magenta (Figure S1-4A/B).

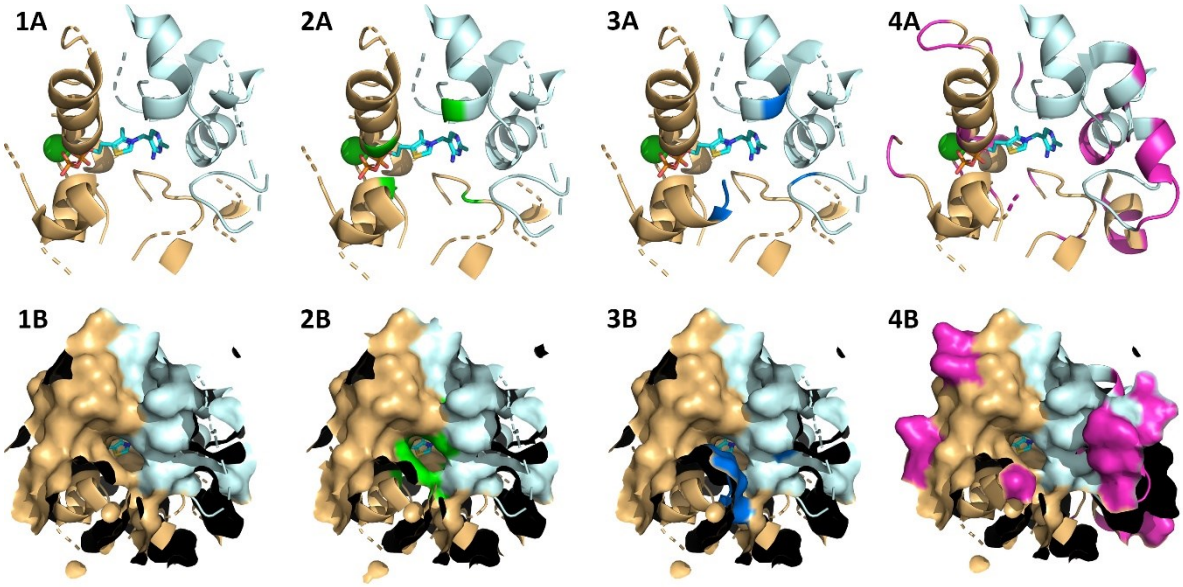

Figure S1: Preliminary *in silico* investigation of FLS. The active site of the enzyme is formed at the interface of the monomers (light brown and blue). The active center is illustrated as (A) cartoon or as (B) surface. (1) shows all surrounding amino acids within a radius of 12 Å. (2) marked in green are all positions with significant improvement in DHA production (3) marked in blue are all non-investigated positons in close proximity to the cofactor and (4) marked in magenta are further non-investigated positions.

High-Throughput Combinatorial Screening

To evaluate the variants in terms of GALD, DHA and ERY production, the double mutant library was screened using different assays. In total, each variant was tested under six conditions. All reactions were prepared in 96-well scale using 50 mM sodium phosphate, pH 8.2 supplemented with 2.0 mM $MgSO_4$, 0.1 mM ThDP and substrate. All approaches were incubated at 30° C in a plate shaker (TiMix 5 control, Edmünd Bühler GmbH, Germany). The following conditions were investigated: (1) Initial production of DHA (GDH assay; Kinetic measurement) started from 134 mM FALD. Simultaneous production of GALD and ERY based on 100 mM FALD (DPA-assay; Endpoint) after (2) 1 h and (3) 22 h. In order to distinguish between GALD and ERY formation activity, the reaction mixtures of (2) and (3) were assayed using GALD specific *Pfu*ADHd, also after (4) 1 h and (5) 22 h. The consumption of GALD was additionally investigated. For this purpose, the library was incubated with 100 mM GALD and the remaining GALD was detected after (6) 22 h by *Pfu*ADHd assay.

Statistics

If the library had been created using NNK (32 codons, 20 amino acids), a total of 1024 variants would have been generated. To achieve >95% library coverage, an oversampling factor of three must be applied. This would have required the screening of 35 plates (96-well). Taking all conditions into consideration, 210 plates would have had to be assayed. Therefore, we searched for possibilities to reduce library size, without reducing the number of allowed amino acids too strongly. It was important to include the template amino acids histidine and glutamine. This ensured that single mutations were also considered. Using NDT (12 codons, 12 amino acids) would not include glutamine. For this, it could be combined with VMA (6 codons, 6 amino acids). For using NDT/VMA (18 codons, 18 amino acids), 88 plates would still have had to be measured. Therefore, we decided to employ an individual library design that include histidine and glutamine. The aim was to generate the same number of variants at the DNA and amino acid level and to include as many amino acids as possible. Therefore, a total of 15 codons and amino acids were allowed for both positions (29 and 113), resulting in 225 variants. As controls, FLS wild type, empty vector (pET24a) and medium were measured on each plate as well. Thus, a total of 704 colonies (plus controls) were investigated on 8 plates. From this, a coverage of 95.7% can be calculated. Accordingly, 48 plates were measured.

Procedure: Production of Library

Electrocompetent *E. coli* BL21 (DE3) was transformed by using the double mutant library. The obtained clones were robot-assisted transferred into individual wells of a 96 DWP using a colony picker (CP7200, Norgren Systems, United Kingdom), which were each filled with 1 mL TB media supplemented with 50 µg mL^{-1} kanamycin. Incubation was conducted overnight at 37° C and 1000 rpm using a plate shaker. Then, the optical density (OD_{710}) was measured automatically (Tecan Freedom Evo 200, Switzerland) and robot-assisted inoculation of the main cultures in 96-deep well plate (DWP) to an OD_{710} of 0.05 was done. To support initial cell growth, the culture plates were incubated at 37° C, 1000 rpm for 3.75 h followed by automated OD measurements. Subsequently, to induce gene expression, 1 mM IPTG was added using the 96-Multichannel-Arm (MCA) of the robot, prior to incubation for 16 h at 25° C, 1000 rpm. Following incubation, the biomass was harvested by centrifugation (15 min, 1000 rpm, 4° C) with subsequent cell lysis (Mand, Wu, Veach, & Kron, 2010) for 1.5 h.

Reactions

Reactions for initial measurement of DHA production were prepared separately and are described in the GDH assay section.

- FALD consumption was prepared as followed: 200 µL of library supernatant was transferred to a 96-DWP and 200 µL master mix (100 mM sodium phosphate, pH 8.2 supplemented with 4.0 mM $MgSO_4$, 0.1 mM ThDP and 100 mM FALD) was added. To avoid evaporation all plates were sealed. DWPs were incubated at 30° C in the plate shaker. After one hour, 100 µL of sample was taken for the DPA assay and another 100 µL for the *Pfu*AdhD assay. This was repeated after further 21 h incubation time.

- GALD consumption was prepared as followed: 100 µL of library supernatant was transferred to a 96-MTP and 100 µL master mix (100 mM sodium phosphate, pH 8.2 supplemented with 4.0 mM $MgSO_4$, 0.1 mM ThDP and 100 mM GALD) was added. To avoid evaporation all plates were sealed again. 96-MTPs were incubated at 30° C in the plate shaker. After 22 h, 100 µL of sample was taken for the *Pfu*AdhD assay.

GDH Assay

DHA production from FALD was measured using a coupled enzyme assay. 100 µL of supernatant containing FLS wild type or variants were combined with 100 µL assay mix consisting of 100 mM phosphate buffer (pH 8.0), 1 mM $MgSO_4$, 1.6 mM NADH, 268 mM FALD, 0.1 mM TPP and 100 µg mL^{-1} glycerol dehydrogenase. NADH concentrations were measured over a 1 h period at 30° C and 340nm.

DPA Assay

GALD was determined in 96-well scale in the presence of FALD and DHA by using a spectrophotometric chromogenic reagent, as previously published.[9] However, the assay also reacted equally with ERY, so that GALD and ERY were not distinguishable. Fortunately, this was not a serious issue, as GALD and ERY was distinguishable by using the *Pfu*ADHd assay. Chromogenic solution was prepared as followed: 3.0 g of diphenylamine was dissolved in 250 mL acetic acid, and then 7.5 mL concentrated sulfuric acid was added. 100 µL sample was transferred to a DWP and 300 µL of chromogenic solution was added to each well. DWPs were sealed with an aluminum foil and incubated at 90° C (water bath) for 15 min. After an additional 15 min on the bench (cool down), the plates were centrifuged at 4000 x *g* for 10 min. Finally, 200 µL were transferred to a 96-MTP and absorption at 650 nm was measured by using a spectrophotometer.

*Pfu*ADHd assay

In order to distinguish between GALD and ERY (DPA assay) the same reaction mixture was assayed with a thermostable alcohol dehydrogenase from *Pyrococcus furiosus* (*Pfu*ADHd). Initial experiments showed specific activity of *Pfu*ADHd with 50 mM GALD in presence and absence of equimolar FALD, DHA and ERY. The reduction master mix contained *Pfu*ADHd in 100 mM potassium phosphate, pH 6.0 supplemented with 2.0 mM NADH. A total of 3 L expression culture was used to prepare 170 ml *Pfu*ADHd (heat purified). Subsequently, 2 mM NADH was weighed and added to the mixture. Finally, 100 µL of sample (FALD or GALD consumption) was transferred to 96-MTP and 100 µL of reduction master mix was added. Oxidation of NADH was monitored at 30° C for 15 min at 340 nm using a spectrophotometer.

Selection of Variants

Using the combinatorial information from all the conditions studied, putative hits were selected. Altogether (for both reactions), 34 colonies were picked and sequenced. Thereby, 10 different amino acids were observed at position 29 and 9 different at position 113. Out of these, 16 variants (Table S3) were selected for pre-characterization.

Table S3: All variants selected for pre-characterization. Final selected variants are marked in bold.

ID	Position H29	Positon Q113	Combined	Name
01	H29	Q113	HQ	wild type
02	H29D	Q113M	DM	-
03	**H29I**	**Q113S**	**IS**	**FLS_B2**
04	H29K	Q113D	KD	-
05	H29G	Q113E	GE	-
06	H29D	Q113D	DD	-
07	**H29M**	**Q113G**	**MG**	**FLS_A1**
08	H29R	Q113G	RG	-
09	H29G	Q113R	GR	-
10	H29G	Q113G	GG	-
11	**H29R**	**Q113M**	**RM**	**FLS_A2**
12	H29E	Q113Q	EQ	-
13	H29I	Q113M	IM	-
14	H29Y	Q113G	YG	-
15	**H29I**	**Q113G**	**IG**	**FLS_B1**
16	H29E	Q113Y	EY	-

All in Table S3 listed FLS variants were expressed and following purified as described in the experimental section. FLS wild type (ID 01) and selected variants (ID 02-16) were diluted to approximately 1 mg mL^{-1} and SDS-PAGE was performed (Figure S2). Interestingly, some of the purified variants showed yellow coloration (also observed for FLS wild type) with different intensity. For some variants, no yellow coloration could be observed anymore.

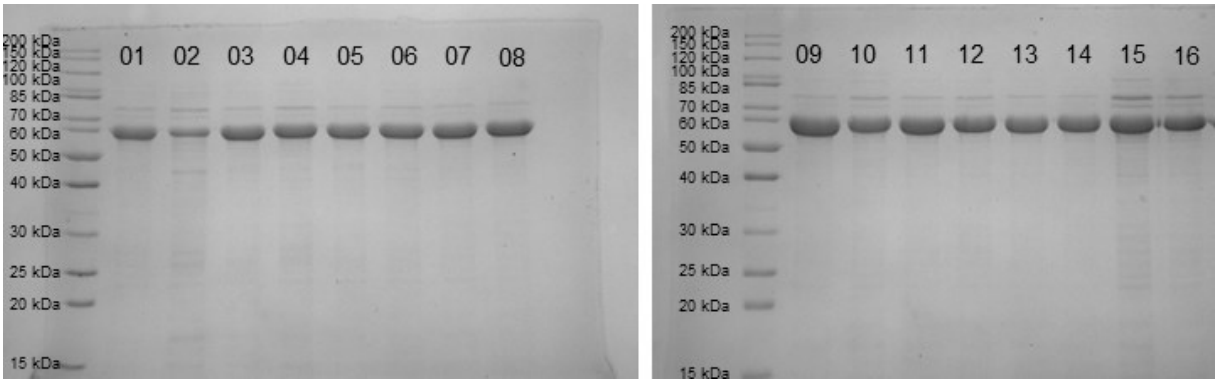

Figure S2: SDS-PAGE of purified variants for pre-characterization. ID of samples corresponds to Table S3

Characterization of FLS Variants

Characterization of FLS variants was performed via HPLC to allow simultaneous detection of FALD, GALD, DHA and ERY. All variants were investigated under three conditions to evaluate their putative behavior in a synthetic enzyme cascade. Therefore, 40 µM of FLS was applied to FALD (3.0 g L^{-1}), GALD (3.0 g L^{-1}) or a mixture (1.5 + 1.5 g L^{-1}). Reactions were performed in 50 mM sodium phosphate buffer (pH 8.2) supplemented with 2 mM $MgCl_2$ and 0.1 mM ThDP for 22 h at 30° C. The template FLS was not considered here because it primarily produced DHA from FALD. Analysis was performed via HPLC.

Further Experiments

Investigation of the Buffer System

The formation of DHA were observed in control reactions (without enzyme addition), when FALD and GALD was used as mixed substrates. Simonov *et al*., already described a catalytic activity of phosphate-based buffer systems in 2007.[10] Therefore, the influence of the buffer system was investigated in more detail, with special regard to the production of ERY. The buffer capacity used was increased 10-fold for this experiment. Furthermore, the influence of the pH value was also investigated. Accordingly, 4.5 g L^{-1} DHA in combination with 1.5 g L^{-1} FALD was incubated in 0.5 M sodium phosphate buffer (pH range 7.0 - 8.0; 0.2 steps) for 22 h at 30° C. As a control, the buffer system was replaced by water. The samples were then analyzed via HPLC, especially with regard to ERY formation (Figure S3). No formation of ERY or other products were observed in the water control. In addition, no consumption of DHA and FALD were found here. Using 0.5 M sodium phosphate, a formation of ERY was detected instead. The formation increased with increasing pH. However, several other products were also measured (Figure S4), which indicated a reaction with low selectivity. No standards for this side products were available. When only FALD, DHA and ERY are considered, a significant carbon loss was observed. At lower buffer capacities, the influence of the buffer was minimal. Nevertheless, an inert buffer system should be considered in future to further increase the selectivity.

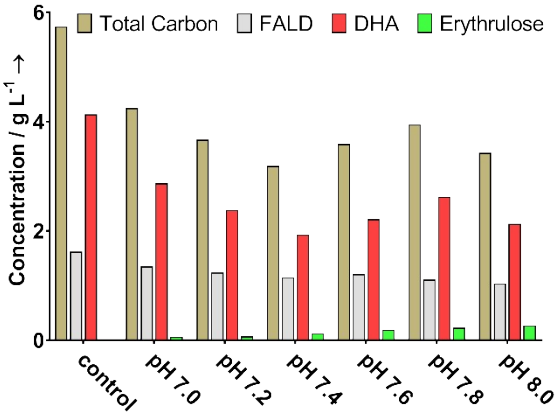

Figure S3: Catalytic activity of the sodium phosphate buffer (pH Range: 7.0 – 8.0). Total carbon is illustrated in black, FALD in grey, DHA in red and ERY in green.

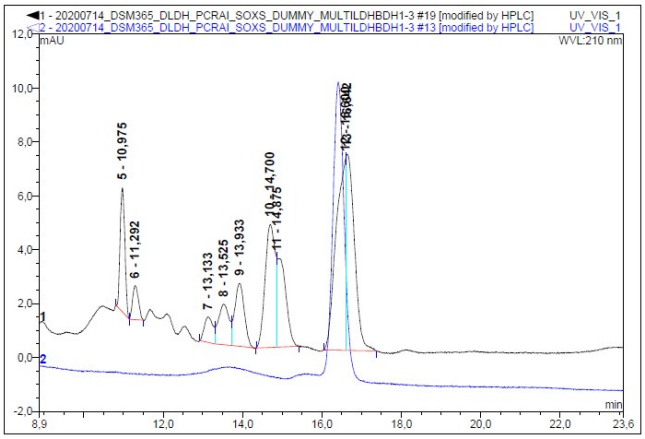

Figure S4: HPLC chromatogram (UV signal) of buffer-catalyzed reaction (0.5 M sodium phosphate, pH 8.0) in black compared to water control (blue). 4.5 g L^{-1} DHA was incubated in combination with 1.5 g L^{-1} FALD for 22 h at 30° C in water or 0.5 M sodium phosphate buffer, pH 8.0. For the control only the signal for DHA was detected, for the buffer reaction a lot of unidentified peaks were observed.

Kinetic Characterization of FLS “wild type” and variants

For kinetic characterization with GALD, approximately one-month-old purified FLS_wt, FLS_B1 and FLS_B2 (stored at -80° C) were thawed at room temperature. Reactions were prepared in a 96-MTP, containing 30 µM of respective FLS (1.84 mg mL^{-1}), 2 mM $MgSO_4$, 0.1 mM ThDP and varying GALD concentrations in 50 mM sodium phosphate buffer, pH 8.0. Thereby, GALD concentrations of 2.5, 5, 7.5, 10, 25, 50, 75, 100, 250, 500, 750 and 1000 mM were investigated. After pre-incubation of reaction mixture and enzymes separately for 10 min at 30° C, reactions were started with addition of enzyme. After mixing by pipetting directly t_0 samples were taken, proper diluted and prepared for HPLC analysis. After further 90 min incubation at 30° C and 750 rpm shaking in a MTP-Shaker, again samples were taken. The initial slope based on ERY production was used to determine catalytic activities. Subsequently, the data were fitted to determine the kinetic parameters.

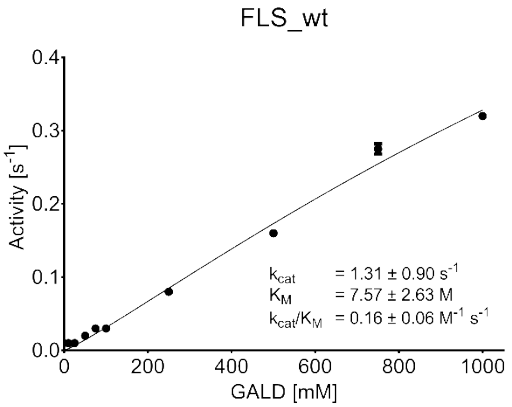

Figure S5: Michaelis-Menten kinetic of FLS_wt.

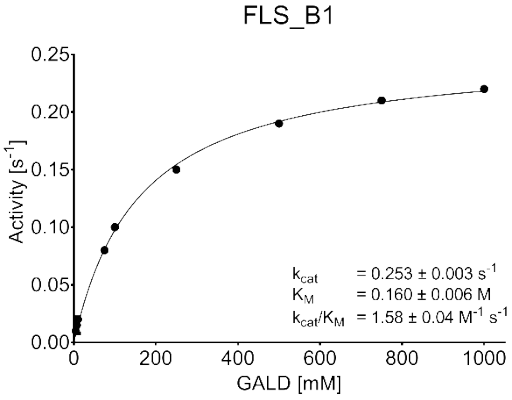

Figure S6: Michaelis-Menten kinetic of FLS_B1.

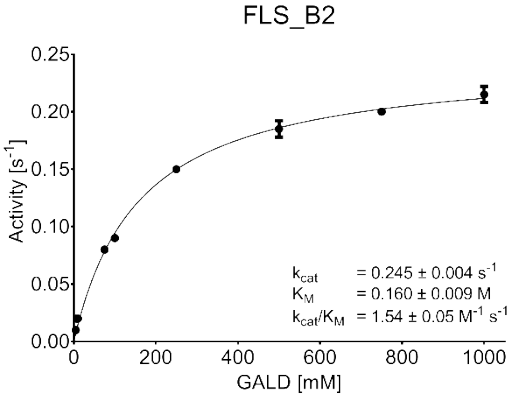

Figure S7: Michaelis-Menten kinetic of FLS_B2.

Influence of the Biocatalytic Process on the Stability of FLS_B2

To investigate whether structural changes of the biocatalysts occur during the biocatalytic process (Figure 5), samples from pre-process start ($t_{0\ h}$; before substrate addition), process start ($t_{0\ h}$; with 25 g L^{-1} GALD) and process end ($t_{16\ h}$) were further investigated. The experiments were performed only on FLS_B2, since the highest theoretical yield of 98% was observed by applying this variant.

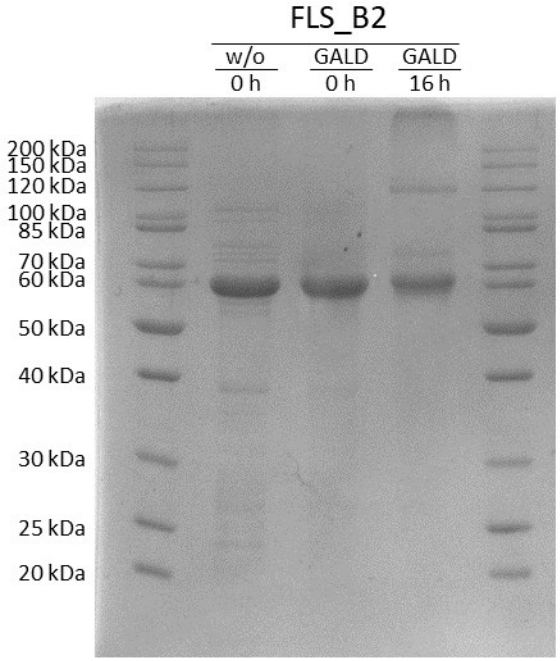

Figure S8: SDS-PAGE of biocatalysis samples of FLS_B2. Samples from pre-process start (w/o; 0 h), process start (GALD; 0 h) and process end (GALD; 16 h) were proper diluted to 1 mg mL^{-1} total protein and analyzed. FLS_B2 has a size of 61.4 kDa.

We first checked whether degradation of FLS_B2 occurs during the biocatalysis. For this purpose, an SDS-PAGE was performed (Figure S8). No significant degradation of FLS_B2 was observed. Although the intensity of the protein band appeared somewhat weaker at the end of the process, no fragments could be detected. Instead, a weak protein band at about 120 kDa was observed. This could be homodimers (122.8 kDa) of the enzyme. Since GALD (25 g L^{-1}; 416.3 mM) is clearly in excess compared to the biocatalysts (150 µM) used, the cross-linking of FLS_B2 might somehow be mediated by GALD without removing much substrate from the reaction. However, degradation of the biocatalyst during the process seems to be unlikely.

Next, we analyzed the same samples to determine whether conformational changes occurred during the biocatalytic process. Therefore, we investigated protein melting points (T_m) using a thermal shift assay. We assumed that any change of the enzyme during the process would also lead to a significant displacement of the T_m or to lower intensities.

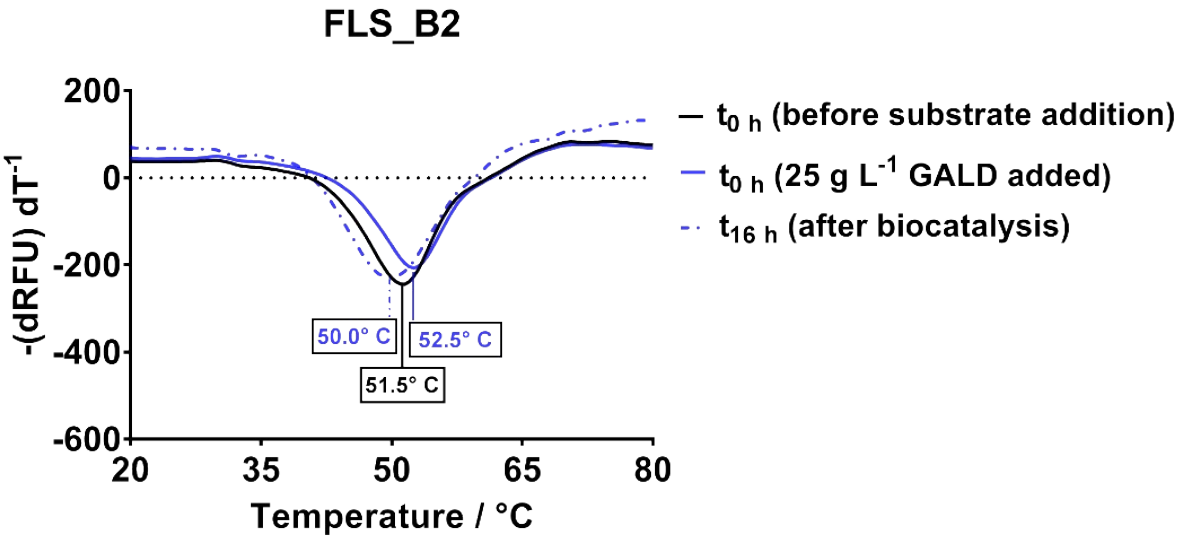

Figure S9: Thermofluor assay of biocatalysis samples of FLS_B2. Samples from pre-process start ($t_{0\ h}$; before substrate addition), process start ($t_{0\,h}$; 25 g L^{-1} GALD added) and process end ($t_{16\,h}$) were proper diluted to 1 mg mL^{-1} total protein and analyzed.

Considering the different environments at the respective time points investigated (pre-process start: no aldehyde; process start: GALD in excess; process end: predominantly ERY), the observed marginal differences in protein melting points (Figure S9) are to be neglected. In addition, no loss of intensity (Decreased interaction between the dye and hydrophobic regions of the enzyme due to previous degradation) or increased baseline (e.g. due to already unfolded or partially unfolded proteins) could be detected. In addition, we would like to point out that the reaction mixture was clear at any time during the biocatalytic process. Therefore, FLS_B2 seems to be sufficiently stable over the biocatalytic process.

HPLC Chromatograms

Calibration

All calibration samples were prepared in a range between 1.0 and 100 mM and analyzed via HPLC. HPLC sample preparation required a 10-fold dilution, resulting in determined range of 0.1 – 10.0 mM. FALD and GALD were determined based on the signal of the RI detector. DHA and ERY were determined based on the signal of the UV detector.

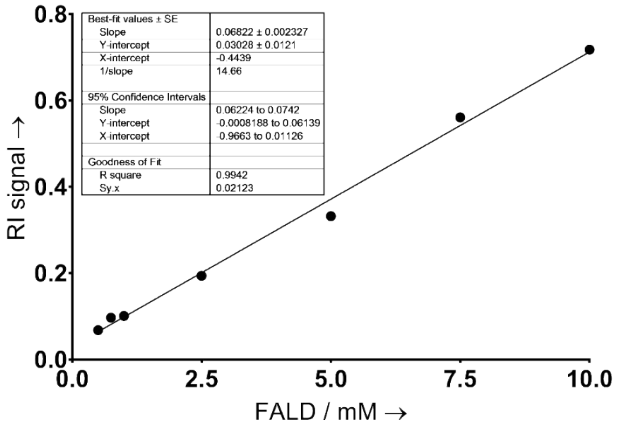

Figure S10: Measured calibration curve of formaldehyde in a range of 0.1 – 10.0 mM FALD.

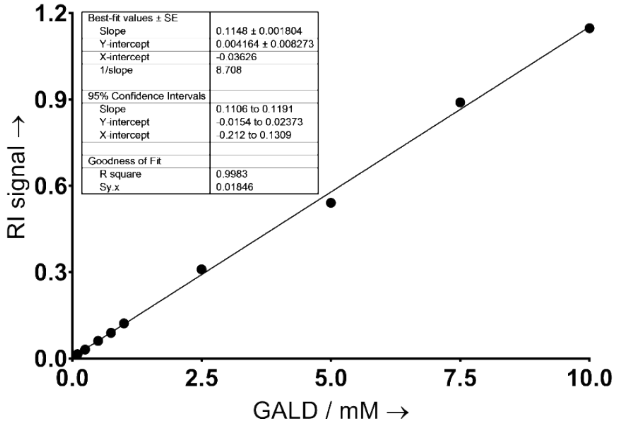

Figure S11: Measured calibration curve of glycolaldehyde in a range of 0.1 – 10.0 mM GALD

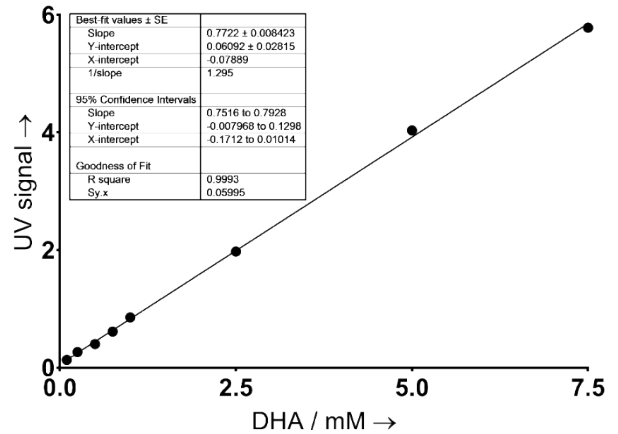

Figure S12: Measured calibration curve of dihydroxyacetone in range of 0.1 – 7.5 mM DHA.

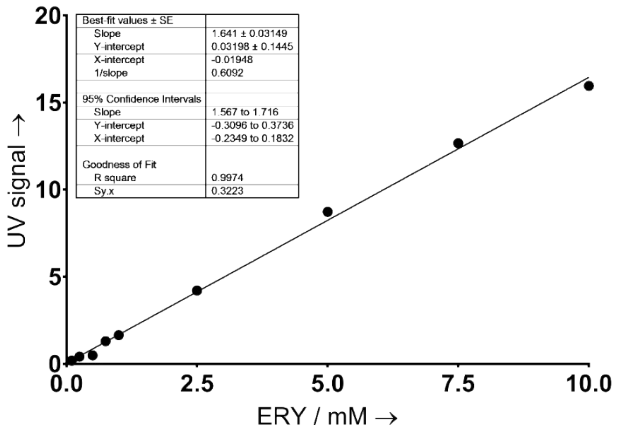

Figure S13: Measured calibration curve of erythrulose in range of 0.1 – 10.0 mM ERY.

Chromatograms of standards

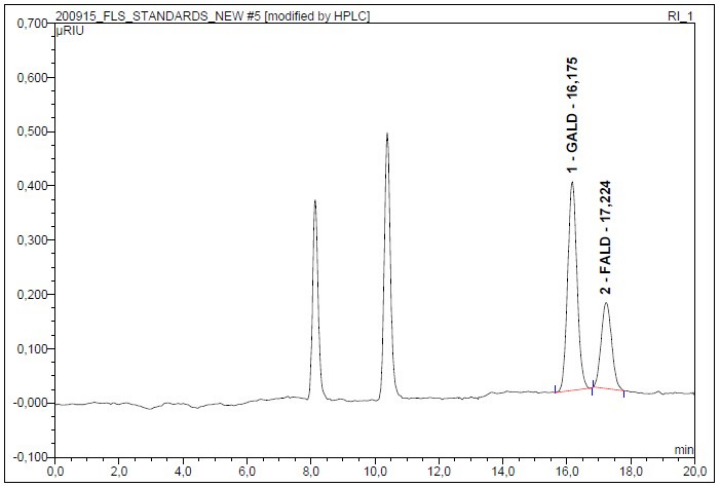

Figure S14: HPLC chromatogram of 1 mM standard 1 (FALD and GALD) with RI detector. GALD retention time of 16.2 and FALD retention time of 17.2 min.

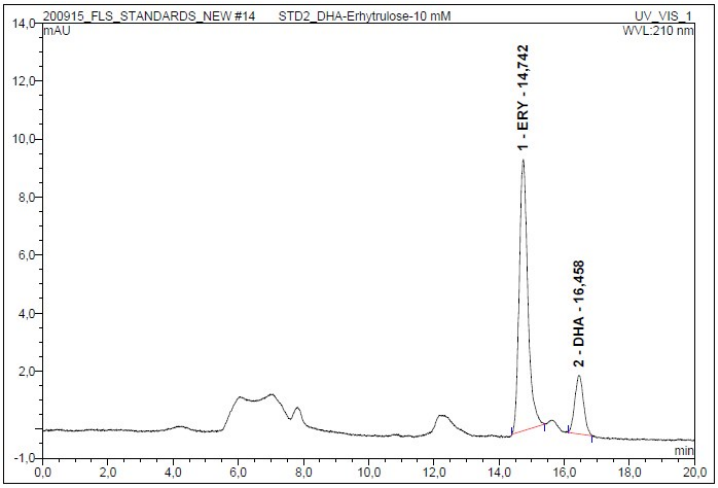

Figure S15: HPLC chromatogram of 1 mM standard 1 (DHA and ERY) with UV detector. ERY retention time of 14.7 and DHA retention time of 16.5 min.

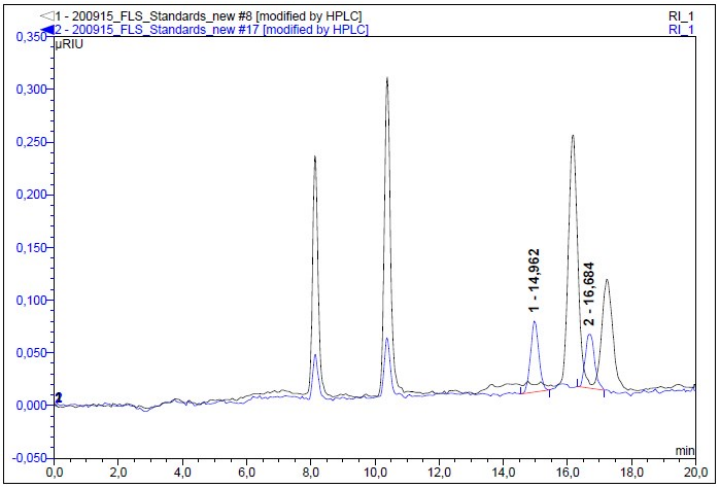

Figure S16: Overlay of chromatograms for standard 1 (Black) and 2 (Blue) to demonstrate successful separation of the corresponding peaks. Signal was detected by RI detector

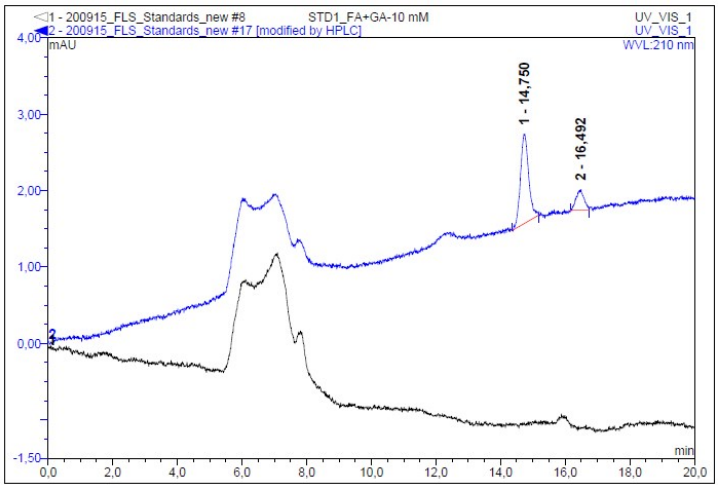

Figure S17: Overlay of chromatograms for standard 1 (Black) and 2 (Blue). Detection by UV detector. Only standard 1 (ERY and DHA) lead to a signal.

References

1 J. B. Siegel, A. L. Smith, S. Poust, A. J. Wargacki, A. Bar-Even, C. Louw, B. W. Shen, C. B. Eiben, H. M. Tran, E. Noor, J. L. Gallaher, J. Bale, Y. Yoshikuni, M. H. Gelb, J. D. Keasling, B. L. Stoddard, M. E. Lidstrom and D. Baker, *Proceedings of the National Academy of Sciences of the United States of America*, 2015, **112**, 3704–3709.

2 R. Machielsen, A. R. Uria, S. W. M. Kengen and J. van der Oost, *Applied and environmental microbiology*, 2006, **72**, 233–238.

3 E. Krieger and G. Vriend, *Bioinformatics (Oxford, England)*, 2014, **30**, 2981–2982.

4 O. Trott and A. J. Olson, *Journal of computational chemistry*, 2010, **31**, 455–461.

5 E. M. Williams, J. N. Copp and D. F. Ackerley, in *Directed Evolution Library Creation*, Springer, 2014, pp. 83–101.

6 E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M. R. Wilkins, R. D. Appel and A. Bairoch, in *The Proteomics Protocols Handbook*, ed. J. M. Walker, Humana Press, Totowa, NJ, 2005, pp. 571–607.

7 F. M. Ausubel, R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith and K. Struhl, *Inc., Hoboken, NJ*, 2003.

8 S. Boivin, S. Kozak and R. Meijers, *Protein expression and purification*, 2013, **91**, 192–206.

9 a) Y. P. Liu, H. Li, Y. Sun and Z. S. Zhou, *Chin J Pharm*, 2011, **42**, 834–837; b) X. Lu, Y. Liu, Y. Yang, S. Wang, Q. Wang, X. Wang, Z. Yan, J. Cheng, C. Liu, X. Yang, H. Luo, S. Yang, J. Gou, L. Ye, L. Lu, Z. Zhang, Y. Guo, Y. Nie, J. Lin, S. Li, C. Tian, T. Cai, B. Zhuo, H. Ma, W. Wang, Y. Ma, Y. Liu, Y. Li and H. Jiang, *Nature communications*, 2019, **10**, 1378;

10 A. N. Simonov, L. G. Matvienko, O. P. Pestunova, V. N. Parmon, N. A. Komandrova, V. A. Denisenko and V. E. Vas'kovskii, *Kinet Catal*, 2007, **48**, 550–555.