One pot synthesis of L-xylose from formaldehyde with an improved benzoylformate decarboxylase

Junhui Zhou[a], Florian Bourdeaux[a], Tianwei Tan[b,c,e], Ulrich Schwaneberg[a,d*]*

a. Institute for Biotechnology RWTH Aachen University, Worringer Weg3, 52074, Aachen, Germany

b. Beijing Key Lab of Bioprocess, National Energy R&D Center for Biorefinery, Beijing 100029, PR China.

c. College of Life Science and Technology, Beijing University of Chemical Technology, Beijing 100029, PR China.

d. Beijing University of Chemical Technology, Beijing 100029, PR China.

e. State Key Laboratory of Green Biomanufacturing, Beijing 100029, PR China

[*]Corresponding author.
E-mail address:
Tianwei Tan: twtan@mail.buct.edu.cn
Ulrich Schwaneberg: u.schwaneberg@biotec.rwth-aachen.de

Methods

Chemicals

Formaldehyde (FALD) was purchased from Carl Roth (Karlsruhe, Germany). Glycolaldehyde (GALD), 1,3-dihydroxyacetone (DHA), glyceraldehyde (GCA), threose, erythrose and L-xylose were purchased from Merck (Darmstadt, Germany). All oligonucleotides were purchased from Eurofins Genomics (Ebersberg, Germany). 2×PCRBIO VeriFi™ Mix was purchased from PCR Biosystems (London, UK). All other chemicals, unless otherwise indicated, were of analytical grade and commercially available.

Thermal resistance prediction of single point variants

The structure of the PnBFD-M1 was obtained using ChimeraX to change the amino acid at specific positions. Force filed of FALD and TPP-FALD complex was generated by ACPYPE server. MD simulations were conducted with GROMACS (version 2022.3) using the Amber99SB-ildn force field and the TIP3P water model. Hydrogens were added to the side chains to mimic a pH value of 7.6. The enzyme (-substrate complex) was centered in a cubic box with 1 nm between the solute and the box. The system was neutralized with NaCl. Energy minimization was performed applying the steepest descent algorithm until a maximum force of 1000 kJ / (mol nm) on any atom was reached. The system was equilibrated by a 1 ns NVT run, followed by a 1 ns NPT run at 1 atm. Pressure and temperature were controlled using the Velocity-rescale and Parrinello-Rahman algorithms. All systems were simulated for 100 ns 340 K. The RMSD of the Cα atoms was calculated. The RMSF of each amino acid residue from 50ns to 100ns was calculated based on the MD trajectories. MD simulations of PnBFD-M1 at 310 K were done in our previous study,[1] the results such as RMSD of the Cα atoms, RMSF of amino acids from 50 ns to 100 ns, and hydrophobicity analysis of surface were published the previous report[1]. Additional analyses were carried out on MD trajectories of PnBFD-M1 at 310 K: identification of potentially important positions, H-bonds analysis, Rg calculation, and SASA calculation.

In rational design strategy 1, the residues with RMSF ≥ 2 Å in the β domain (179-338 amino acid residue) of PnBFD-M1 at 310 K were targeted to calculate the changes caused by point mutations using FoldX, DeepDDG, and I-Mutant 2.0. Among these three predictors, variants with ΔΔG ≤ +0.36 kcal/mol in FoldX, and with ΔΔG> 0 kcal/mol in DeepDDG and I-Mutant 2.0 were retained.

In rational design strategy 2, the RMSFs of the long loop residues in long loops in the β domain of PnBFD-M1 at 310 K and 340 K were compared. Residues with an RMSF at 340 K that is higher than or equal to 0.5 A compared to 310 K were then targeted for mutation to Proline.

Site-directed mutagenesis (SDM)

The plasmid pET-22b(+)-PnBFD-M1 was selected as the template for constructing the variants. The primers used for mutagenesis are listed in **Table S1**. The SDM variants was generated according to the PCR-based Quick-Change method. PCR reaction was performed with VerifiTM Mix under the following condition: Template 10ng (pET22b(+)-PnBFD-M1), F primer 10μM, R primer 10μM, VerifiTM Mix 25 μl were used; for PCR (72 ºC for 1 min, 1 cycle; 95 ºC, 15 s/65 ºC, 15 s/72 ºC, 4 min, 30 cycles; 72 ºC for 10 min, 1 cycle). The PCR product was digested

with the DpnI restriction enzyme to remove template DNA and subsequently purified using a PCR purification kit. The purified PCR products were transformed into E. coli BL21 (DE3) cell for further expression and screening.

Protein Expression and Purification

The gene of PnBFD-M1 and its variants, GALS-YM (GALS-F397Y/C398M), and FSA-TG (FSA-A129T/A165G) were cloned into the expression vector pET-22b(+) between NdeI and XhoI restriction sites, incorporating a C-terminal hex histidine tag for purification. All enzymes were expressed in *E.coli* BL21 (DE3). Cultures were first grown overnight at 37 °C in 4 mL LB medium containing 100 μg/mL ampicillin, then inoculated into 50 mL LB medium supplemented with the same antibiotic and cultured at 37 °C. Isopropyl β-D-1-thiogalactopyranoside (IPTG) was added when OD_{600} reached 0.6-0.8, then incubated at 18 °C, 20 h for protein expression. 50 μM IPTG was added to GALS-YM, PnBFD-M1 and its variants, and 200 μM IPTG was added to FSA-TG.

After incubation, the cells were harvested by centrifugation at 4 °C and 4000 rpm for 30 min and then resuspended in buffer 1 (50 mM NaPi, 300 mM NaCl, pH 7.5). The cells were lysed by sonication (15 s on, 15 s off, 70% power, 10 min in total), followed by centrifugation to remove cell debris. For FSA-TG, the supernatant was then subjected to purification using a Ni-IDA resin affinity column. The resin was washed sequentially with 30 mL of buffer 1 and 30 mL of buffer 2 (50 mM NaPi, 10 mM imidazole, 300 mM NaCl, pH 7.5) to remove non-target proteins, followed by elution of the target proteins using 15 mL of buffer 4 (50 mM NaPi, 100 mM imidazole, 300 mM NaCl, pH 7.5). The collected eluate was concentrated using an Amicon tube (10 KDa) to remove imidazole and change buffer. For GALS-YM, PnBFD-M1 and its variants, the supernatant was then subjected to purification using a Ni-NTA resin affinity column. The resin was washed sequentially with 30 mL of buffer 1 and 30 mL of buffer 3 (50 mM NaPi, 50 mM imidazole, 300 mM NaCl, pH 7.5) to remove non-target proteins, followed by elution of the target proteins using 15 mL of buffer 5 (50 mM NaPi, 250 mM imidazole, 300 mM NaCl, pH 7.5). The collected eluate was concentrated using an Amicon tube (30 KDa) to remove imidazole and change buffer. The protein content of the purified enzyme solution was measured using a microspectrophotometer equipped with a protein content detection function (Thermo).

Characterization of enzymatic activity and kinetic parameters

The standard enzyme activity assay for PnBFD-M1 and its variants was carried out under the following conditions: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM $MgSO_4$, 0.2 mg/mL purified enzyme, and 10 mM FALD. The reaction was incubated at 37 °C for 30 min. After incubation, 25 μL reaction mixture was mixed with 475 μL 1.0 g/L 2,4-dinitrophenylhydrazine (DNPH) solution (solution A: dissolve 1.56 g $NaH_2PO_4 \cdot H_2O$ in 50 mL of water; solution B: dissolve 0.1 g DNPH in 50 mL acetonitrile; then mix solutions A and B in a 1:1 (V: V) ratio) and derivatized at 60 °C for 30 min. The GALD production was then measured by high-performance liquid chromatography (HPLC). Samples not immediately analyzed were stored at -20°C for later use.

When measuring the initial reaction rates of PnBFD-M1, and PnBFD-M2 (Amino acid sequence was shown in **Table S2**) at different FALD concentrations, the reaction conditions were as follows: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM $MgSO_4$, 0.2 mg/mL purified

enzyme, and 10-200 mM FALD. The reactions were incubated at 37 °C for 150 min. At intervals, 25 μL or 10 μL samples were taken and mixed with 475 μL or 490 μL of 1.0 g/L DNPH solution, followed by derivatization at 60 °C for 30 min. The GALD production was subsequently measured using HPLC. The initial activity was calculated by fitting the linear area of the curve (first 20 min).

Characterization of thermal resistance parameters

The $t_{1/2}$, T_{opt}, and storage stability at 4 °C, and *T*m were determined to assess the thermal resistance of PnBFD-M1 and its variants.

To determine $t_{1/2}$ values, 0.2 mg/mL PnBFD-M1 or PnBFD-M2 was incubated at 37 °C, 40 C, 45 °C, and 50 °C, and the residual activities were measured at certain time intervals. The residual activity was determined as follows: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM $MgSO_4$, 0.2 mg/mL purified enzyme after heat treatment, and 10 mM FALD. The reaction was incubated at 37 °C for 30 min. After incubation, 25 μL reaction mixture was mixed with 475 μL 1.0 g/L DNPH solution and derivatized at 60 °C for 30 min and then measured by HPLC. The $t_{1/2}$ was calculated according to a first-order deactivation function (Equations 1 and 2).

$$A_t=A_0 \exp(-k_d t) \quad (1)$$

$$t_{1/2}=\ln 2/k_d \quad (2)$$

where A_t represents the specific activity after incubation for time of $_t$ (U/mg), A_0 is the initial specific activity at 0 min (U/mg), t is the incubation time (min), and k_d represents the deactivation rate constant (min^{-1}).

The T_{opt} of PnBFD-M1 or PnBFD-M2 was determined according to specific activity at different temperatures. The specific activity was determined as follows: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM $MgSO_4$, 0.2 mg/mL purified enzyme, and 10 mM FALD. The reaction was incubated at 25-60 °C for 30 min. After incubation, 25 μL reaction mixture was mixed with 475 μL 1.0 g/L DNPH solution and derivatized at 60 °C for 30 min and then measured by HPLC. The temperature at which the enzyme exhibits the highest activity is defined as T_{opt}.

To determine storage stability at 4 °C, PnBFD-M1 or PnBFD-M2 was stored at 4 °C, RT, or – 20 °C for few days, and the residual activities were measured at certain time intervals. The residual activity was determined as follows: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM $MgSO_4$, 0.2 mg/mL purified enzyme after storing at 4 °C, RT, or – 20 °C, and 10 mM FALD. The activity of fresh enzymes was set as 100% activity. The reaction was incubated at 37 °C for 30 min. After incubation, 25 μL reaction mixture was mixed with 475 μL 1.0 g/L DNPH solution and derivatized at 60 °C for 30 min and then measured by HPLC.

The *T*m of all enzymes were detected by protein thermostability analyzer NT48 (Nanotemper) after purification.

MD simulations

The structure of the PnBFD-M2 was obtained using ChimeraX to change the amino acid at specific positions. Force filed of FALD and TPP-FALD complex was generated by ACPYPE server. MD simulations were conducted with GROMACS (version 2022.3) using the Amber99SB-ildn force field and the TIP3P water model. Hydrogens were added to the side chains to mimic a pH value of 7.6. The enzyme-substrate complex was centered in a cubic box with 1 nm between

the solute and the box. The system was neutralized with NaCl. Energy minimization was performed applying the steepest descent algorithm until a maximum force of 1000 kJ / (mol· nm) on any atom was reached. The system was equilibrated by a 1 ns NVT run, followed by a 1 ns NPT run at 1 atm. Pressure and temperature were controlled using the Velocity-rescale and Parrinello-Rahman algorithms. All systems were simulated for 100 ns at 310 K. The RMSD of Cα was calculated. The RMSF of each amino acid residue from 50 ns to 100 ns was calculated based on the MD trajectories. GROMACS built in tools and pymol were used for trajectory analysis, secondary structure change. Structure visualizations were done using pymol.

L-xylose pathway construction

A one-pot method was used to construct the L-xylose pathway. Reaction solution containing PnBFD-M1 or PnBFD-M2, FSA-TG, 1 mM TPP, 2 mM $MgSO_4$, and 100 mM FALD was reacted at 37 °C for 10 h. Optimizing PnBFD-M1 and FSA-TG concentration to get the highest L-xylose production. Samples were taken at intervals, 40 μL samples were taken and terminated by 40 μL 0.2 M H_2SO_4 and then measured threose, erythrose, and L-xylose production by HPLC (with a RI detector and a Aminex HPX-87C column), 10 μL samples were mixed with 490 μL 1.0 g/L DNPH solution and derivatized at 60 °C for 30 min and then measured by HPLC (with a DAD detector and C18 column).

HPLC analysis

HPLC method for FALD, GALD, DHA, and GCA analysis: The analysis was performed using a Shimadzu HPLC chromatograph equipped with a DAD detector and a C18 column (Thermo, 4.6 × 250 mm, 5 μm). The mobile phase consisted of two solvents: acetonitrile and water, mixed in a 50:50 (v:v) ratio. The detection was performed at column temperature of 35 °C, with a flow rate of 0.5 mL/min and a wavelength of 368 nm. The retention times for GCA, GALD, and DHA were 5.2 min, 6.2 min, and 7.2 min, respectively for GCA, GALD, and DHA, and FALD were 5.2 min, 6.2 min, 7.2 min and 14.2 min, respectively.

HPLC method for threose, erythrose, and L-xylose analysis: The analysis was performed using a Shimadzu HPLC chromatograph equipped with a RI detector and a Aminex HPX-87C column (Bio-RAD, 300 × 7.8 mm,). The mobile phase was 5 mM H_2SO_4. The detection was performed at column temperature of 60 °C, with a flow rate of 0.6 mL/min. The retention times for threose, erythrose, and L-xylose were 11.3 min, 13.2 min, and 23.8 min, respectively.

Scale-up and product characterization

The reaction scale was increased to 38 mL by coupling 8 mg/mL PnBFD-M2 with 4 mg/mL FSA-TG. After incubation at 40 °C for 10 h, the L-xylose concentration determined by HPLC was 10.01 mM, corresponding to a theoretical total amount of 57.11 mg. Proteins were precipitated by the addition of 40 mL of methanol, followed by centrifugation at 4000 rpm for 30 min. The resulting supernatant was lyophilized and re-dissolved in 8 mL of water. An aliquot (1.6 mL) was purified by HPLC to collect the xylose fraction and subsequently lyophilized, yielding 11.9 mg of product. HPLC measurement indicated that 9.3 mg of this product was L-xylose, corresponding to a purity of 79%. The overall recovery yield was calculated to be 82%, and the product was further

confirmed by mass spectrometry (MS). As FSA-TG has previously been demonstrated by NMR to produce L-xylose from FALD and GALD,[2] the product was assigned as L-xylose.

Tables and Figures

Table S1 Primers for PCR

Primers	Primer sequence from 5'-3'
K188I-F	acaaagtgATCgaatttgcccagcgtattaccgc
K188I-R	gcaaattcGATcactttgtccggatcgggc
Q192I-F	aggaatttgccATCcgtattaccgctagca
Q192I-R	aatacgGATggcaaattccttcactttgtccgga
K259I-F	tcgctggaaATCcaaattcaaggccacgatct
K259I-R	tgaatttgGATttccagcgaaccgatacccg
K259M-F	tcgctggaaATGcaaattcaaggccacgatct
K259M-R	tgaatttgCATttccagcgaaccgatacccg
D327Q-F	gcttattCAGcagcgcgaaaaaaataacacgcctca
D327Q-R	tcgcgctgCTGaataagctttaaggcctcaatta
D327E-F	gcttattGAGcagcgcgaaaaaaataacacgcctca
D327E-R	tcgcgctgCTCaataagctttaaggcctcaatta
K331R-F	cgcgaaCGGaataacacgcctcaacgatctcct
K331R-R	aggcgtgttattCCGttcgcgctgatcaataagct
Q336I-F	acacgcctATCcgatctcctatgacgaaagaaga
Q336I-R	aggagatcgGATaggcgtgttatttttttcgcg
Q336V-F	acacgcctGTCcgatctcctatgacgaaagaaga
Q336V-R	aggagatcgGACaggcgtgttatttttttcgcg
A282P-F	tgctagaagcCCTgcatggtctgacggcattg
A282P-R	ctgcccAGGtatccatgggtaatagcggaa
G283P-F	ggatagcaCCTcagtttatCcctgaaggctcaac
G283P-R	caggGataaactgAGGtgctatccatgggtaatagcgga
Q284P-F	ggatagcagggCCTtttatCcctgaaggctcaac
Q284P-R	caggGataaaAGGccctgctatccatgggtaatagcgga
I286P-F	ggatagcagggcagtttCCTcctgaaggctcaac
I286P-R	caggAGGaaactgccctgctatccatgggtaatagcgga
K331P-F	cgcgaaCCTaataacacgcctcaacgatctcct
K331P-R	aggcgtgttattAGGttcgcgctgatcaataagct
K188I/Q192I-F	gtgATCgaatttgccATCcgtattaccgctagca
K188I/Q192I-R	acgGATggcaaattcGATcactttgtccggat

Table S2 Amino acid sequences of PnBFD-M1, PnBFD-M2, GALS-YM, and FSA-TG

Enzymes	Amino acid sequences
PnBFD-M1	MRTVKEITFDLLRKLQVTTVVGNPGSTEETFLKDFPSDFNYVLALQEA SVVAIADGLSQSLRKPVIVNIHTGAGLGNAMGCLLTAYQNKTPLIITAG QQTREMLLNEPLLTNIEAINMPKPWVKWSYEPARPEDVPGAFMRAYA TAMQQPQGPVFLSLPLDDWEKLIPEVDVARTVSTRQGPDPDKVKEFA QRITASKNPLLIYGSDIARSQAWSDGIAFAERLNAPVWAAPFAERTPF PEDHPLFQGALTSGIGSLEKQIQGHDLIVVIGAPVFRYYPWIAGQFIPE GSTLLQVSDDPNMTSKAVVGDSLVSDSKLFLIEALKLIDQREKNNTPQ RSPMTKEDRTAMPLRPHAILEVLKENSPKEIVLVEECPSIVPLMQDVF RINQPDTFYTFASWGLGWDLPAAVGLALGEEVSGRNRPVVTLMGDG

	AFQYSVQGIYTGVQQKTHVIYVVFQNEEYGILKQFAELAQTPNVPGLD LPGLDIVAQGKAYGAKSLKVETLDELKTAYLEALSFKGTSVIVVPITKEL KPLFG
PnBFD-M2(M1-K188I/Q192I/A282P)	MRTVKEITFDLLRKLQVTTVVGNPGSTEETFLKDFPSDFNYVLALQEA SVVAIADGLSQSLRKPVIVNIHTGAGLGNAMGCLLTAYQNKTPLIITAG QQTREMLLNEPLLTNIEAINMPKPWVKWSYEPARPEDVPGAFMRAYA TAMQQPQGPVFLSLPLDDWEKLIPEVDVARTVSTRQGPDPDKVIEFAI RITASKNPLLIYGSDIARSQAWSDGIAFAERLNAPVWAAPFAERTPFPE DHPLFQGALTSGIGSLEKQIQGHDLIVVIGAPVFRYYPWIPGQFIPEGS TLLQVSDDPNMTSKAVVGDSLVSDSKLFLIEALKLIDQREKNNTPQRS PMTKEDRTAMPLRPHAILEVLKENSPKEIVLVEECPSIVPLMQDVFRIN QPDTFYTFASWGLGWDLPAAVGLALGEEVSGRNRPVVTLMGDGAFQ YSVQGIYTGVQQKTHVIYVVFQNEEYGILKQFAELAQTPNVPGLDLPG LDIVAQGKAYGAKSLKVETLDELKTAYLEALSFKGTSVIVVPITKELKPL FG
FSA-TG (FSA-A129T/A165G)	MELYLDTANVAEVERLARIFPIAGVTTNPSIIAASKESIWEVLPRLQKAI GDEGILFAQTMSRDAQGMVEEAKRLRDAIPGIVVKIPVTSEGLAAIKILK KEGITTLGTAVYSAAQGLLAALAGAKYVTPYVNRVDAQGGDGIRTVQ ELQTLLEMHAPESMVLAGSFKTPRQALDCLLAGCESITLPLDVAQQML NTPAVESAIEKFEHDWNAAFGTTHL
GALS-YM (GALS-F397Y/C398M)	MASVHGTTYELLRRQGIDTVFGNPGSNELPFLKDFPEDFRYILALQEA CVVGIADGYAQASRKPAFINLHSAAGTGNAMGALSNARTSHSPLIVTA GQQTRAMIGVEAGETNVDAANLPRPLVKWSYEPASAAEVPHAMSRAI HMASMAPQGPVYLSVPYDDWDKDADPQSHHLFDRHVSSSVRLNDQ DLDILVKALNSASNPAIVLGPDVDAANANADCVMLAERLKAPVWVAPS APRCPFPTRHPCFRGLMPAGIAAISQLLEGHDVVLVIGAPVFRYVFYD PGQYLKPGTRLISVTCDPLEAARAPMGDAIVADIGAMASALANLVEES SRQLPTAAPEPAKVDQDAGRLHPETVFDTLNDMAPENAIYLNESTSTT AQMWQRLNMRNPGSYYYMAAGGLGFALPAAIGVQLAEPERQVIAVIG DGSANYSISALWTAAQYNIPTIFVIMNNGTYGMLRWFAGVLEAENVPG LDVPGIDFRALAKGYGVQALKADNLEQLKGSLQEALSAKGPVLIEVST VSPVK

Table S3 Residues in the β-domain of PnBFD-M1 with RMSF ≥ 2 Å at 310K

Residues	RMSF (Å)	Residues	RMSF (Å)
K188	2.01	D327	2.03
Q192	2.11	E330	2.75
K198	2.46	K331	3.79
K259	2.02	N332	2.67
E288	2.36	N333	3.26
V306	2.03	Q336	2.04

Table S4 Protein concentration of PnBFD-M1 and PnBFD-M2 on multi-batch reuse experiments

Batch	Theoretical concentration (mg/mL)	PnBFD-M1 (mg/mL)	PnBFD-M2 (mg/mL)
1	0.500	0.500	0.500
2	0.488	0.429±0.015	0.422±0.014
3	0.475	0.390±0.015	0.397±0.024
4	0.463	0.312±0.012	0.348±0.012

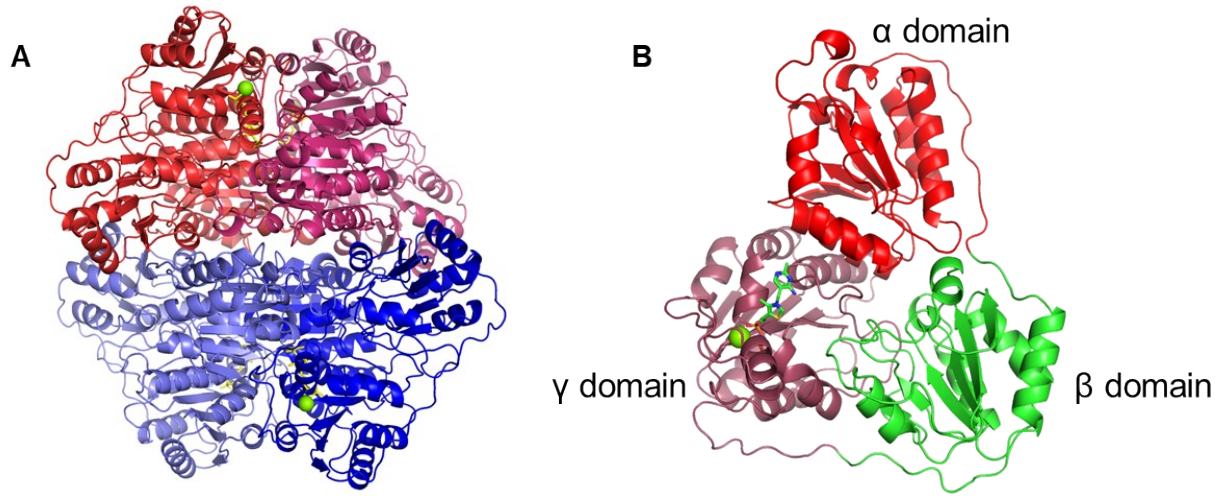


Fig. S1 (A) The tetramer structure of PnBFD-M1; Each catalytic unit is formed by a dimer: the red and dark-red chains form one active unit, and the blue and blue-gray chains form the other. TPP is shown as yellow sticks and Mg^{2+} as green spheres; **(B)** The structure of PnBFD-M1 monomer with the three main domains: α domain and the γ domain build up the catalytic site, while the β domain has a structural role by stabilizing the two other domains.

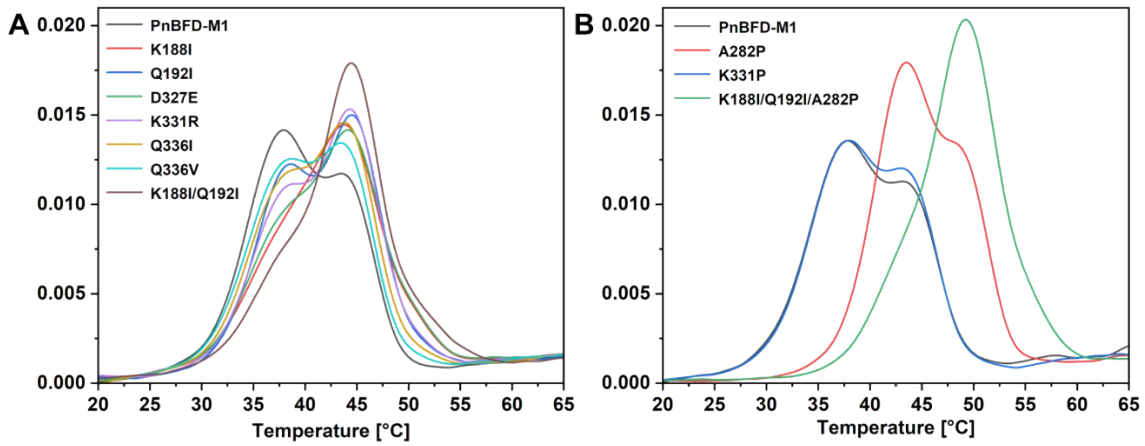


Fig. S2 (A) Thermal denaturation curves of PnBFD-M1 and variants with maintained or enhanced activity identified through the ΔΔG-based screening strategy, measured with the NanoTemper Prometheus NT.48; **(B)** Thermal denaturation curves of PnBFD-M1, variants with maintained or enhanced activity obtained by introducing proline into flexible loops, and the combined variants, measured with the NanoTemper Prometheus NT.48.

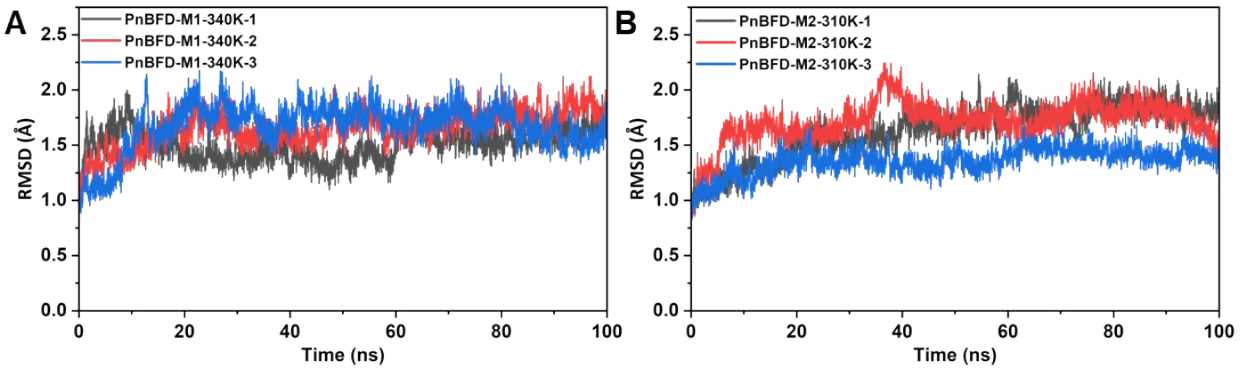


Fig. S3 RMSD of the Cα atoms of **(A)** PnBFD-M1-340K; and **(B)** PnBFD-M2-310K. RMSD of the Cα atoms of PnBFD-M1 at 310 K can be seen in our previous report[1].

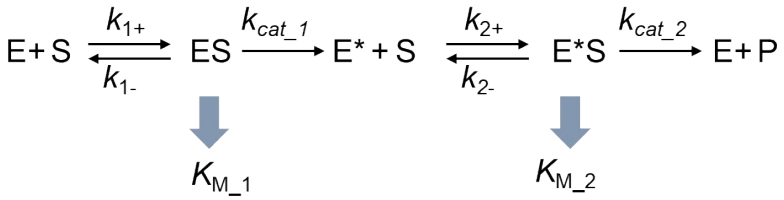


Fig. S4 Ping-pong kinetic mechanism of PnBFD-M1 and its variants. [1,3,4] This mechanism involves sequential substrate binding and product release, with the enzyme alternating between two chemically distinct states. Because only one substrate is bound at a time and multiple partial reactions contribute to the overall rate, determining true kinetic parameters (K_M and *k*cat) is inherently complex.

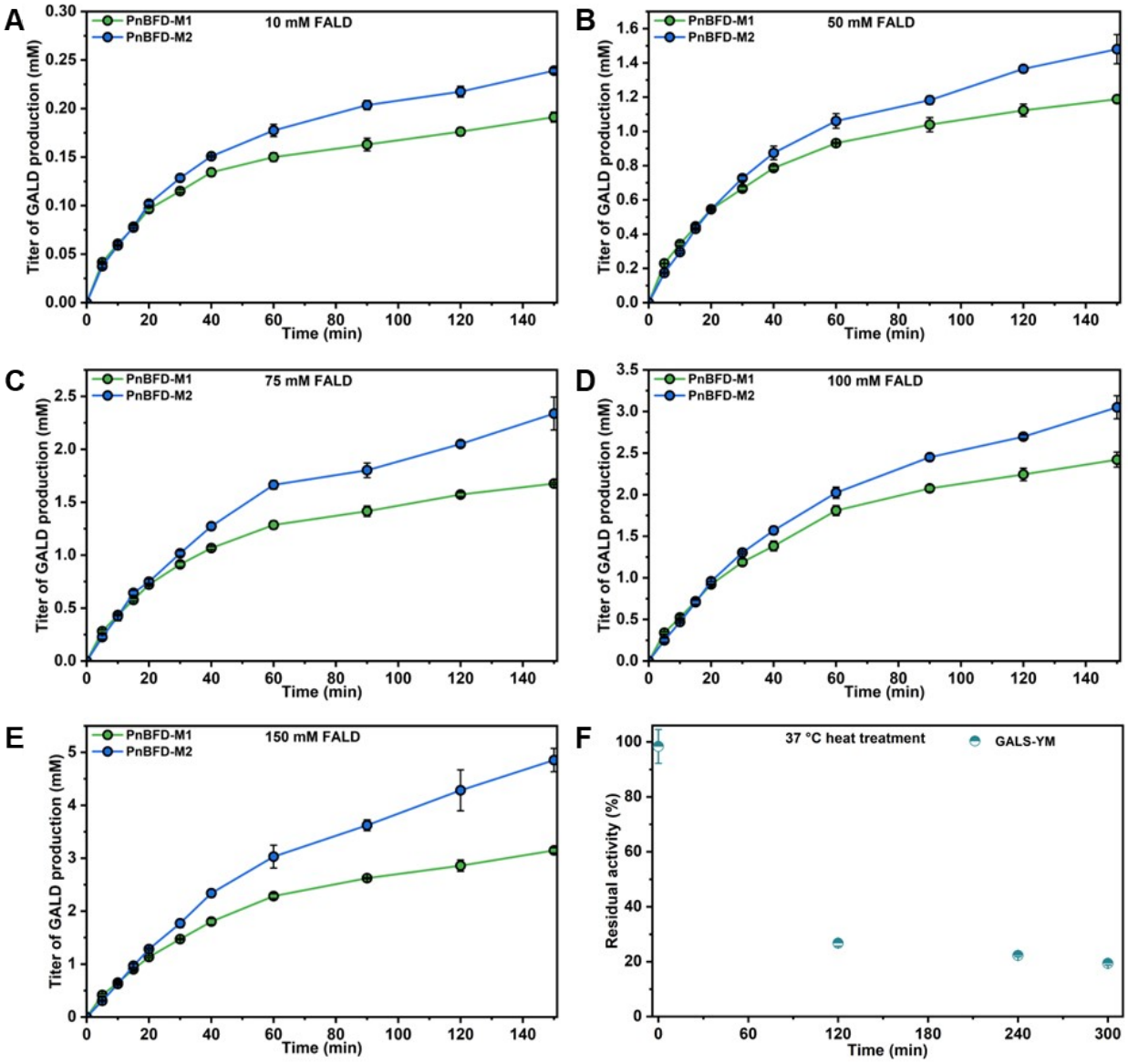


Fig. S5 (A)-(E) The GALD production for PnBFD-M1 and PnBFD-M2 was measured at different FALD concentrations; Activity assay: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM Mg^{2+}, 0.2 mg/mL purified enzyme, and 10-200 mM FALD concentration at 37 °C; **(F)** Residual activity of GALS-YM after heat treatment at 37 °C, after incubation the residual activity was measured, the activity of enzymes without heat treatment was set as 100%; Activity assay: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM Mg^{2+}, 0.2 mg/mL GALS-YM, and 10 mM FALD concentration; Error bars represent standard deviation of three technical replicates (n = 3).

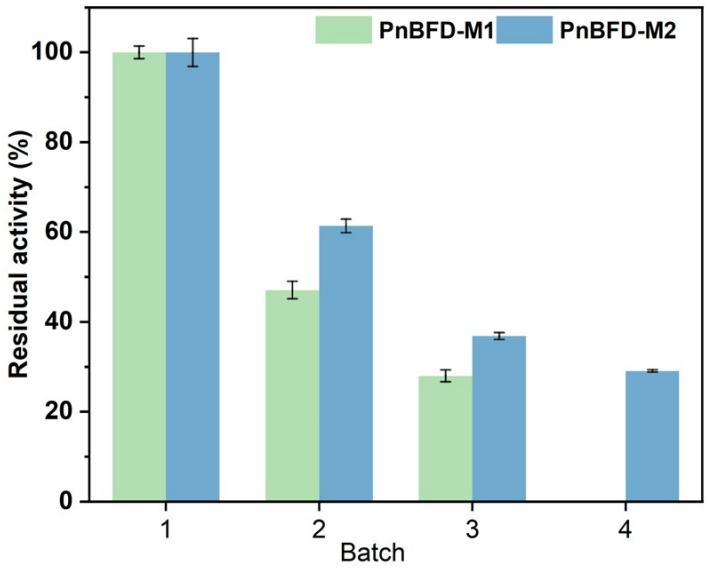


Fig. S6 Multi-batch reactions of PnBFD-M1 or PnBFD-M2. Activity assay: 50 mM bicine (pH 7.6), 300 mM NaCl, 1 mM TPP, 2 mM Mg^{2+}, 0.5 mg/mL PnBFD-M1 or PnBFD-M2 at first batch, and 10 mM FALD concentration, 37 °C for 30 min; After each reaction cycle, PnBFD-M1 or PnBFD-M2 was recovered using 30 kDa Amicon ultrafiltration tubes and washed three times with fresh bicine buffer (50 mM bicine (pH 7.6), 300 mM NaCl) to remove residual substrates and products. The recovered enzyme was then reused for the subsequent batch reaction under identical conditions. Error bars represent standard deviation of three technical replicates (n = 3).

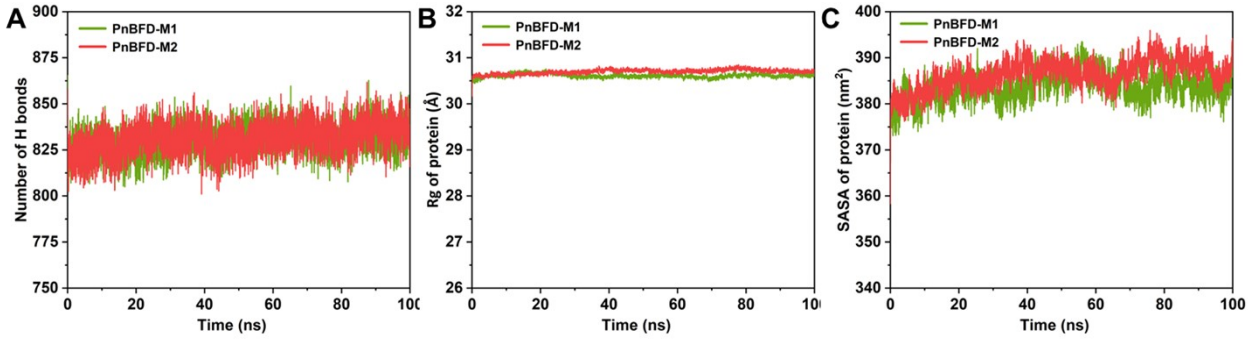


Fig. S7 (A) Number of H bonds, **(B)** Radius of gyration (Rg) of protein, and **(C)** Solvent accessible surface area (SASA) of protein of PnBFD-M1 and PnBFD-M2 over MD simulations.

Fig. S7 shows that the temporal profiles of hydrogen bonds (H-bonds), radius of gyration (Rg), and solvent accessible surface area (SASA) for PnBFD-M1 and PnBFD-M2 are nearly identical. The average number of H-bonds, Rg, and SASA for PnBFD-M1 are 830.17, 30.61 Å, and 383.74 nm^2, respectively, while those for PnBFD-M2 are 831.03, 30.70 Å, and 386.44 nm^2. These results indicate that PnBFD-M2 exhibits only minor differences in H-bonds, Rg, and SASA compared to PnBFD-M1.

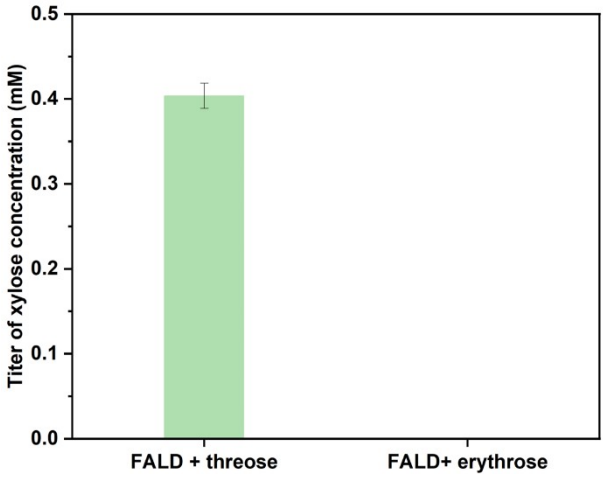


Fig. S8 Titer of xylose catalyzed by FSA-TG using FALD with threose or erythrose as substrate. Enzyme assay: 50 mM pH7.6 bicine, 300 mM NaCl, 0.1 mg/mL FSA-TG, and 10 mM FALD with 10 mM threose or 10 mM erythrose as substrates, 37 °C for 1 h; Error bars represent standard deviation of three technical replicates (n = 3).

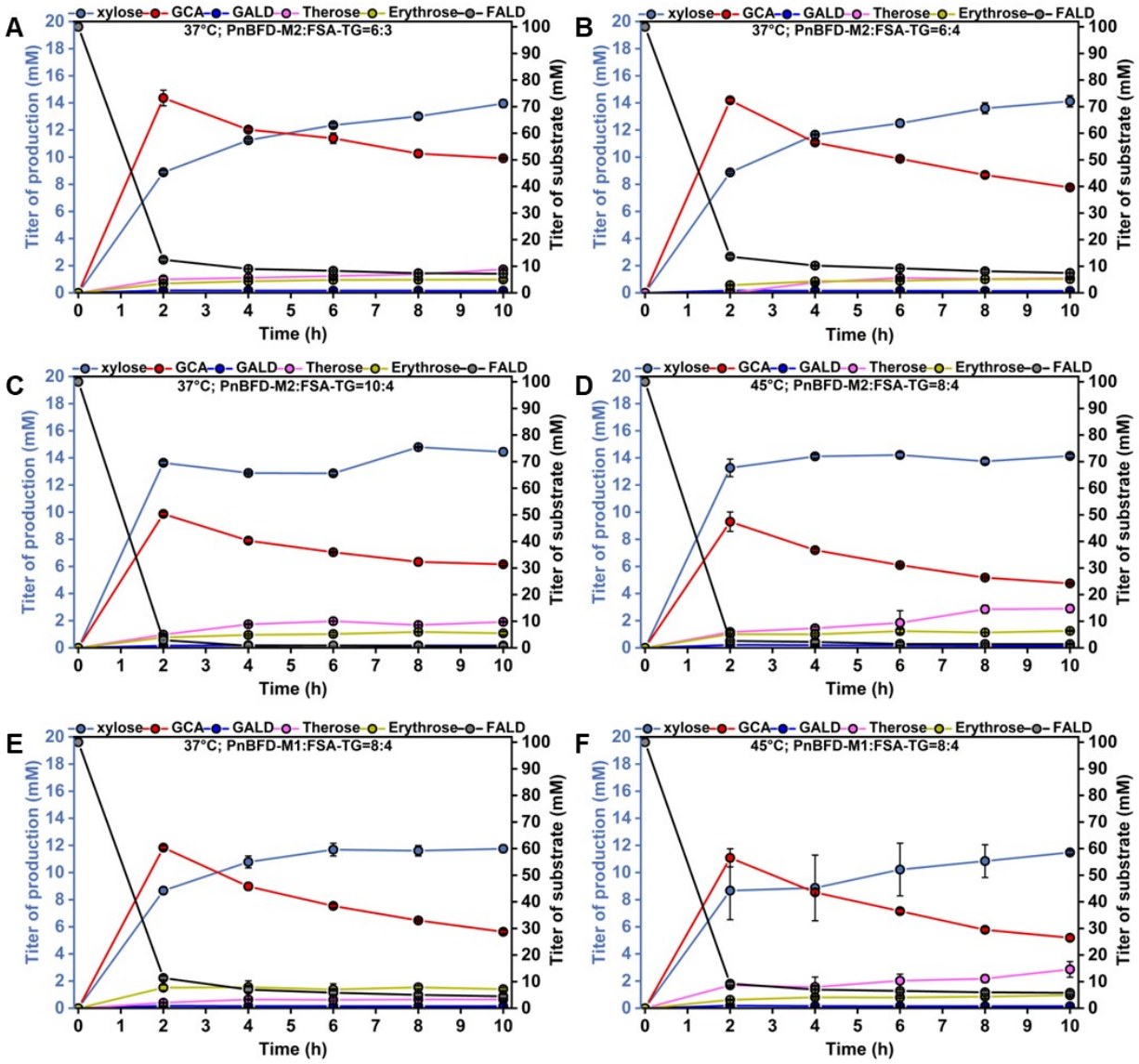
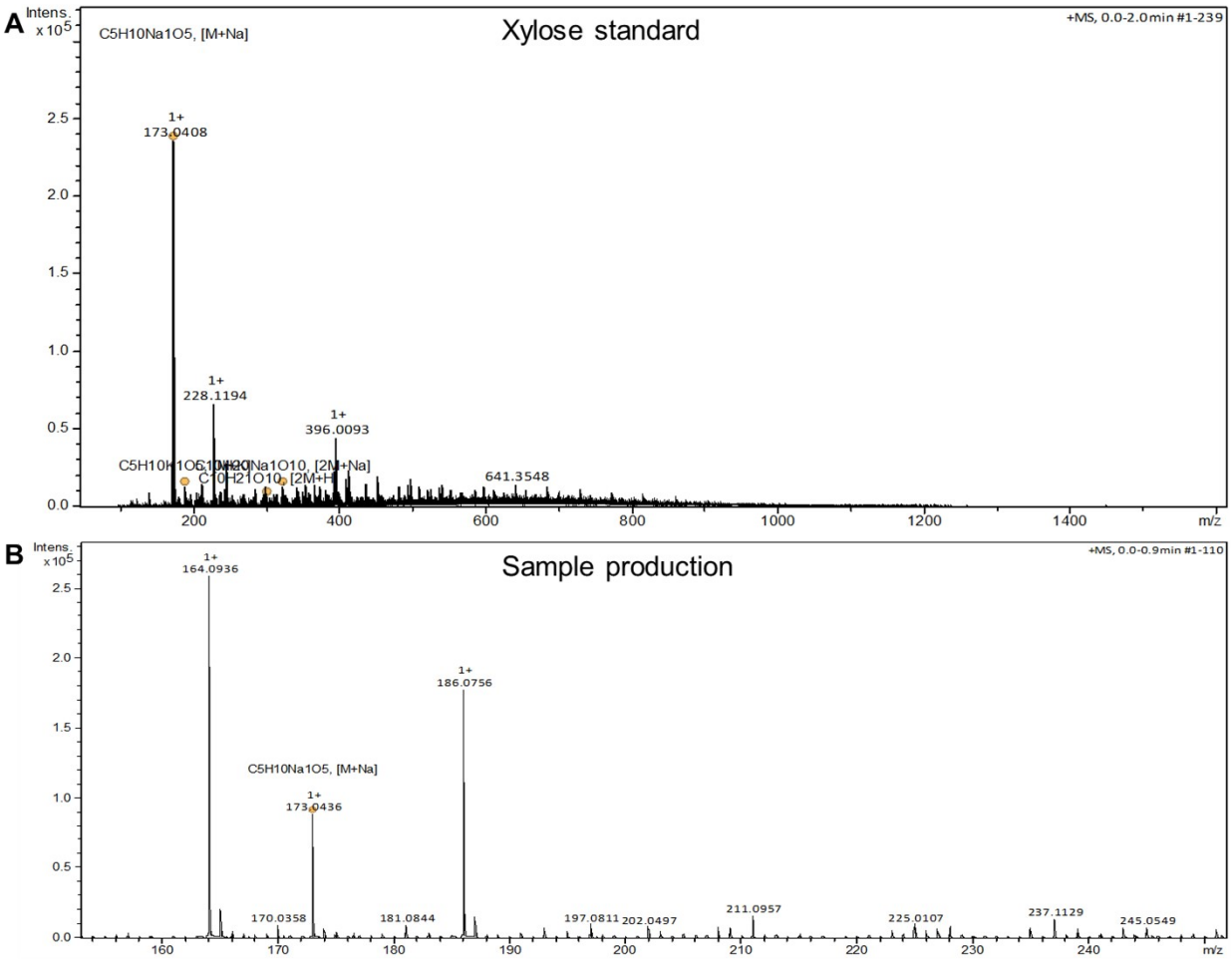


Fig. S9 Titer of products and substrate in L-xylose pathway of different ratios and different temperature; **(A)** 37 °C, 6 mg/mL PnBFD-M2 and 3 mg/mL FSA-TG; **(B)** 37 °C, 6 mg/mL PnBFD-M2 and 4 mg/mL FSA-TG; **(C)** 37 °C, 10 mg/mL PnBFD-M2 and 4 mg/mL FSA-TG; **(D)** 50 °C, 8 mg/mL PnBFD-M2 and 4 mg/mL FSA-TG; **(E)** 37 °C, 8 mg/mL PnBFD-M1 and 4 mg/mL FSA-TG; **(F)** 45 °C, 8 mg/mL PnBFD-M1 and 4 mg/mL FSA-TG; Activity assay: 50 mM pH7.6 bicine, 300 mM NaCl, 1 mM TPP, 2 mM Mg^{2+}, purified enzyme, and 100 mM FALD concentration, different temperature for 10 h; Error bars represent standard deviation of three technical replicates (n = 3).


A
Intens.
x 10^5
C5H10Na1O5, [M+Na]
Xylose standard
+MS, 0.0-2.0min #1-239
1+
173.0408
1+
228.1194
1+
396.0093
641.3548
m/z
B
Intens.
x 10^5
Sample production
+MS, 0.0-0.9min #1-110
1+
164.0936
1+
186.0756
C5H10Na1O5, [M+Na]
1+
173.0436
170.0358
181.0844
197.0811
202.0497
211.0957
225.0107
237.1129
245.0549
m/z

242

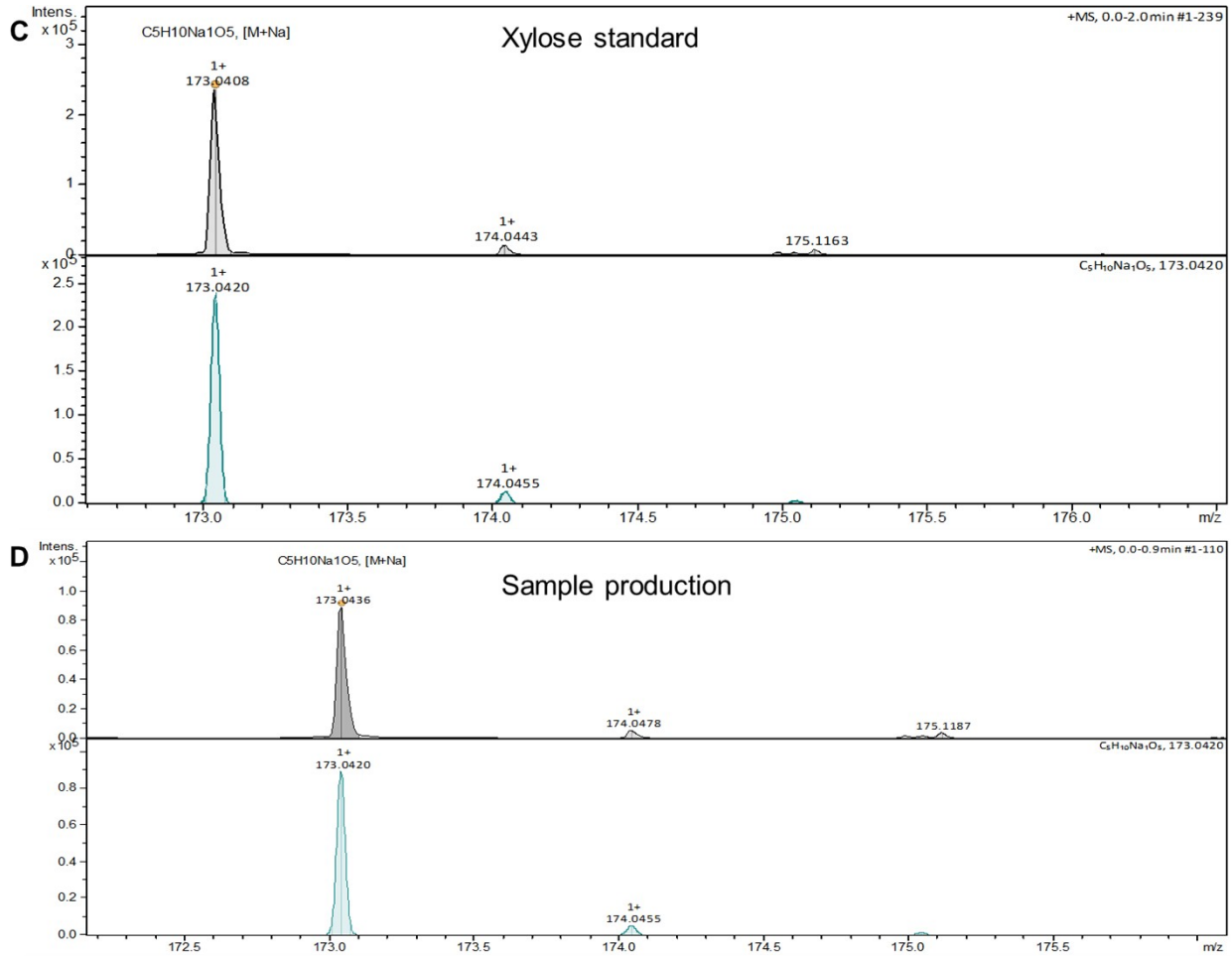


Fig. S10 MS spectra of xylose standard ((A) and (C)) and the enzymatic production sample ((B) and (D)). (A) and (C) show the full spectra, (B) and (D) show the expanded view around m/z = 173. Calculated mass (monoisotopic): L-xylose(+Na): 173.0421 Da (measured: Product: 173.0436 Da. Control: 173.0408 Da). Bicine(+H): 164.0917 Da (measured: 164.0936 Da). Bicine(+Na): 186.0737 Da (measured: 186.0756 Da). Masses were calculated with the web tool MolE - Molecular Mass Calculator v2.04 from https://mstoolbox.github.io/ by Jef Rozenski.

References

1. Zhou, J.; Bourdeaux, F.; Emann, V.; Tan, T.; Schwaneberg, U., A highly active and specific benzoylformate decarboxylase for glycolaldehyde production from formaldehyde. *ACS Catalysis* **2025**, 14469-14482.
2. Szekrenyi, A.; Garrabou, X.; Parella, T.; Joglar, J.; Bujons, J.; Clapes, P., Asymmetric assembly of aldose carbohydrates from formaldehyde and glycolaldehyde by tandem biocatalytic aldol reactions. *Nature Chemistry* **2015,** *7* (9), 724-9.
3. Liu, S.; Gong, X.; Yan, X.; Peng, T.; Baker, J. C.; Li, L.; Robben, P. M.; Ravindran, S.; Andersson, L. A.; Cole, A. B.; Roche, T. E., Reaction mechanism for mammalian pyruvate dehydrogenase using natural lipoyl domain substrates. *Archives of Biochemistry and Biophysics* **2001,** *386* (2), 123-35.
4. Bornadel, A.; Akerman, C. O.; Adlercreutz, P.; Hatti-Kaul, R.; Borg, N., Kinetic modeling of lipase-catalyzed esterification reaction between oleic acid and trimethylolpropane: a

263 simplified model for multi-substrate multi-product ping-pong mechanisms. *Biotechnology*
264 *Progress* **2013,** *29* (6), 1422-9.

265