# Supplementary Information

# A cell-free artificial anabolic pathway for direct conversion of $CO_2$ to ethanol

Wanrong Dong[1,2#], Xiuling Ji[1#], Yuhong Huang[1]*, Yaju Xue[1], Boxia Guo[1], Dongbo Cai[2], Shouwen Chen[2]*, and Suojiang Zhang[1]*

[1] Beijing Key Laboratory of Ionic Liquids Clean Process, CAS Key Laboratory of Green Process and Engineering, State Key Laboratory of Multiphase Complex Systems, Institute of Process Engineering, Chinese Academy of Sciences, Beijing 100190, China
[2] State Key Laboratory of Biocatalysis and Enzyme Engineering, Environmental Microbial Technology Center of Hubei Province, College of Life Sciences, Hubei University, Wuhan 430062, China
*Corresponding authors: Yuhong Huang (yhhuang@ipe.ac.cn), Shouwen Chen (chenshouwen@hubu.edu.cn), Suojiang Zhang (sjzhang@ipe.ac.cn)

**Table of Contents**

## Materials and Methods

*Chemicals and agents.* Common chemicals were purchased from Sigma-Aldrich (Shanghai, China), SolarBio (Beijing, China) and Sinopharm Chemical Reagent Co., Ltd. The following enzymes: formate dehydrogenase from *C.boidinii* (CbFDH), formaldehyde dehydrogenase from *Pseudomonas sp.* (PsFaldDH), glucokinase from *S.cerevisiae* (Glk), acetate kinase from *E. coli* (Ack), phosphoglucose isomerase from *S.cerevisiae* (Pgi), glucose-6-phosphate dehydrogenase from *S.cerevisiae* (G6PD) and alcohol dehydrogenase from *S. cerevisiae* (ADH) were all purchased from Sigma-Aldrich (Shanghai, China).

*Strains and growth conditions. E. coli* BL21 (DE3) was grown at 37 ºC or 20 ºC in LB medium for protein expression.

*Plasmid construction and heterologous expression.* The gene encoding the required enzyme was sent to Shanghai Generay Biotechnology company for proceeding several steps, including codon optimization, gene synthesis, construction on the pETDuet-1 vector with a 6x His-tag, and heterologous expression in *E. coli* BL21 (DE3). For expression, the plasmids containing the constructed gene were first grown in 50 mL of LB medium at 37 ºC overnight with continuous shaking. Afterwards, 2 mL of the overnight culture was transferred to 250 mL of fresh LB medium and cultured until reaching an optical density at 600 nm ($OD_{600}$) of 0.6-1.0. At this point, cells were induced with 0.1 mM IPTG (isopropyl β-D-1-thiogalactopyranoside) and incubated at 20 ºC for 20 h. Following induction, the cells were harvested and resuspended in a lysis buffer consisting of 30 mM Tris-HCl (pH 7.4), 50 mM sodium phosphate, and 150 mM NaCl. To purify the enzymes, a HisTrap HP crude affinity column (5 mL, 17524802, Cytiva) and a HiTrap Desalting column (5 mL, 17140801, Cytiva) were used in conjunction with fast protein liquid chromatography (FPLC) on an ÄKTA™ pure 25 system. During the purification process, a running buffer A (50 mM sodium phosphate, 150 mM NaCl, and pH 7.4) and B (50 mM sodium phosphate, 150 mM NaCl, 500 mM imidazole and pH 7.4) were employed. The purified proteins were then stored at -80 ºC for future analysis. The purity of the enzymes was assessed using SDS-PAGE (sodium

dodecyl sulfate-polyacrylamide gel electrophoresis). Additionally, the protein concentration of the purified enzymes was determined using the Pierce™ BCA Protein Assay Kit (23225, Thermo Scientific), with bovine serum albumin (BSA) serving as a standard for comparison.

*The thermodynamic data analysis of pathway*. The total Gibbs energy change ($\Delta rG'^{m}$) and the max-min driving force (MDF) were calculated by the website of Quilibrator (http://equilibrator.weizmann.ac.il) using the ionic strength I = 0.15, pH 7.5, and 1 mM concentration of each metabolite[1].

*GC-MS analysis.* Ethanol was analyzed by Nexis GC-2030 gas chromatograph (GC) or gas chromatography-tandem mass spectrometry (GC-MS, Agilent Technologies). The GC was applied with an SH-Rtx-Wax polar column (60 m × 0.32 mm × 1.0 μm) and a GC-flame ionization detector (BID). The GC oven temperature was set to start at 40 °C for 3 min, and immediately follow by a ramp of 45 °C/min to 235 °C and a hold for 3 min. An injection volume of 0.5 μL with a split ratio of 25 in constant pressure mode with 9.52 psi at the inlet was used[2]. For full-scan data acquisition, the MS was set to scan from 20 to 200 atomic mass units. The ethanol concentration was calculated by using an external standard method with the corresponding calibration curve established by using the known concentrations of ethanol. Data analysis for GC-MS was performed with NIST Database.

***In vitro* activity assay of enzymes**

All assays were performed with 50 mM potassium phosphate buffer at pH 7.5 and conducted at 25 °C unless otherwise stated.

*Conditions for FDH Assays.* A reaction volume of 1 mL was made up of 1 mM NADH, 200 μg FDH, and a buffer solution saturated with $CO_2$ for 0.5 h in advance. Reactions were conducted for 1 h and the assay results were determined by the measurements of the consumption of NADH at 340 nm.

*Conditions for FaldDH Assays.* A reaction volume of 1 mL was composed of 1 mM NADH, 10 mM sodium formate, and 200 μg FaldDH. The assay results were determined by the measurements of the consumption of NADH at 340 nm.

*Conditions for Phi Assays.* A 500 μL reaction mixture was made up of 5 mM $MgCl_2$, 0.26 mM $NADP^+$, 5 mM R5P, 5 mM formaldehyde, 6.1 μg Pgi, 1.1 μg G6PD, 48 μg Rpi, 9.6 μg Hps and 0.0048 μg Phi. The assay results were determined by the measurements of the production of NADPH at 340 nm.

*Conditions for Fpk Assay.* A 500 μL reaction was made up of 5 mM $MgCl_2$, 0.26 mM $NADP^+$, 1 mM thiamine pyrophosphate (TPP), 0.2 mM ADP, 10 mM glucose, 10 mM F6P, 1.6 μg F/Xpk, 30 μg Ack, 2 U Glk, and 2 U G6PD. The assay results were determined by the measurements of the NADPH generation at 340 nm.

*Conditions for Tal, Tkt, Rpe, and Rpi Assays.* A 500 μL reaction mixture was made up of 5 mM R5P, 5 mM $MgCl_2$, 1 mM TPP, 0.26 mM $NADP^+$, 0.4 U Glk and 0.4 U Pgi. For coupling reactions, high enzyme amounts were used: 260 μg Tal, 8 μg Tkt,14 μg Rpe, and 12 μg Rpi. The tested enzyme in each assay was used at the following levels: 0.5 μg Tal, 0.3 μg Tkt, 0.2 μg Rpe, and 0.1 μg Rpi. The assay results were determined by the measurements of the NADPH generation at 340 nm.

*Conditions for AlDH Assay.* A 500 μL reaction mixture was made up of 5 mM $MgCl_2$, 0.3 mM NADH, 1 mM butyryl-CoA, and 10 μg AlDH. The assay results were determined by the measurements of the decrease in NADH at 340 nm.

*Conditions for Pta Assays.* A 500 μL reaction mixture was made up of 0.2 mM NADH, 2 mM acetyl-phosphate, 5 mM CoA, 0.077 μg Pta, and 44 μg AlDH. The assay results were determined by the measurements of the NADPH generation at 340 nm.

*Screening of Hps.* The 500 μL reaction mixture was made up of 5 mM $MgCl_2$, 0.26 mM $NADP^+$, 5 mM R5P, 5 mM formaldehyde, 6.1 μg Pgi, 1.1 μg G6PD, 48 μg Rpi, 47 μg Phi and 1.9 μg Hps. The *in vitro* enzyme activities of Hps from different strains were determined by the measurements of the NADPH generation at 340 nm.

**Pathway assays**

All assays were performed with 50 mM potassium phosphate buffer at pH 7.5 and conducted at 25 °C unless otherwise stated. External standard method was used to establish the ethanol standard curve for the qualitative analysis of the product yield in the sample.

*Screening FDHs and FaldDHs.* A reaction volume of 1 mL was made up of 1 mM NADH, 200 μg CbFDH/ PaFDH/ ΔPaFDH48, and 200 μg PsFaldDH/ BmFaldDH/ SzFaldDH. During the reaction, the buffer solution was saturated with $CO_2$ 0.5 h in advance, and the prepared system reacts were bubbled with $CO_2$ for 1 h. The assay results were determined by the measurements of the NADH consumption at 340 nm.

*For Formaldehyde to Ethanol.* A reaction volume of 1 mL was made up of 0.2 mM NADH, 5 mM $MgCl_2$, 10 mM sodium formate, 0.5 mM TPP, 6 mM formaldehyde, 200 μg FDH, 150 μg Hps, 50 μg Phi, 300 μg Tkt, 500 μg Tal, 90 μg Fpk, 50 μg Rpe, 50 μg Rpi, 50 μg Pta, 300 μg AlDH and 0.1U ADH. The reaction time taken for formaldehyde to ethanol was 4 h. Tal was excluded for the control. Samples were analyzed by GC-MS.

*Concentration optimization for Fpk.* A reaction volume of 1 mL was made up of 0.2 mM NADH, 5 mM $MgCl_2$, 10 mM sodium formate, 0.5 mM TPP, 6 mM formaldehyde, 200 μg FDH, 150 μg Hps, 50 μg Phi, 300 μg Tkt, 500 μg Tal, 50 μg Rpe, 50 μg Rpi, 50 μg Pta, 300 μg AlDH and 0.1 U ADH. Fpk was set at different concentrations in this system. Samples were reacted for 4 h, then analyzed by GC-MS.

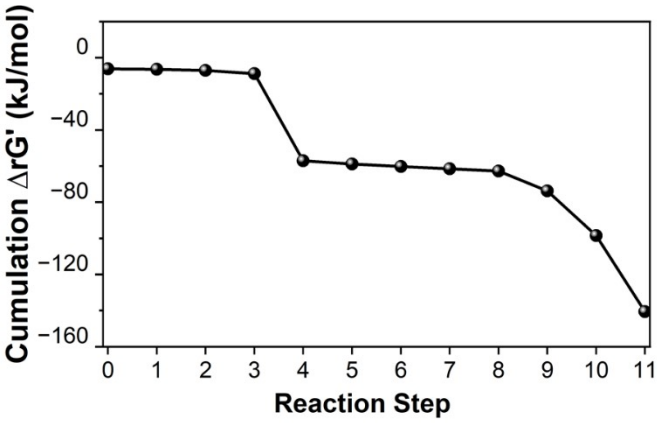
*For $HCO_3^-$ to Ethanol.* A reaction volume of 1 mL was made up of 500 mM $NaHCO_3$, 5 mM NADH, 5 mM $MgCl_2$, 0.5 mM R5P, 1 mM CoA, 0.5 mM TPP, and the amount of enzyme are as follows: 200 μg FDH, 200 μg FaldDH, 150 μg Hps, 30 μg Phi, 300 μg Tkt, 500 μg Tal, 30 μg Rpe, 30 μg Rpi, 140 μg Fpk, 30 μg Pta, 300 μg AlDH, 0.1 U ADH. The reactions were conducted at 37 ℃ for 6 h. Hps was excluded for the control. At each time point, samples were mixed with 8 M urea in a ratio of 10:1 to quench the reaction, then analyzed by GC-MS.

*One Step for $CO_2$ to Ethanol.* A reaction volume of 1 mL was made up of 5 mM NADH, 5 mM $MgCl_2$, 0.5 mM R5P, 1mM CoA, 0.5 mM TPP, and the amount of enzyme are as follows: 200 μg FDH, 200 μg FaldDH, 150 μg Hps, 30 μg Phi, 300 μg Tkt, 500 μg Tal, 30 μg Rpe, 30 μg Rpi, 140 μg Fpk, 30 μg Pta, 300 μg AlDH, 0.1 U ADH. During the reaction, the buffer solution was saturated with $CO_2$ 0.5 h in advance, and the reactions were conducted for 2 h. Hps was excluded for the control, and samples were

analyzed by GC-MS.

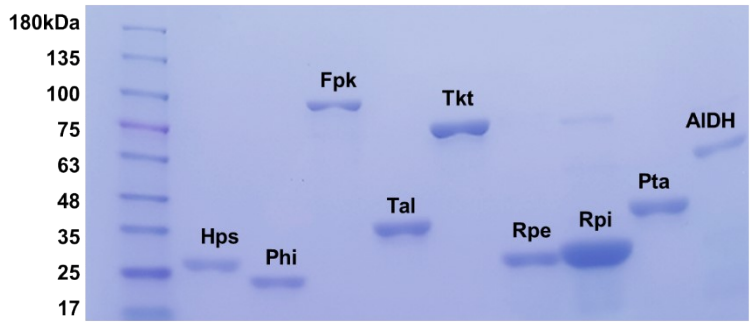
*Two-Step for $CO_2$ to Ethanol.* A reaction volume of 1 mL was made up of 5 mM NADH, 5 mM $MgCl_2$, 0.5 mM R5P, 1 mM CoA, 0.5 mM TPP, and the amount of enzyme are as follows: 200 μg FDH and 200 μg FaldDH, 150 μg Hps, 30 μg Phi, 300 μg Tkt, 500 μg Tal, 30 μg Rpe, 30 μg Rpi, 140 μg Fpk, 30 μg Pta, 300 μg AlDH, 0.1 U ADH. During the reaction, the buffer solution was saturated with $CO_2$ 0.5 h in advance, and Tkt, Tal, Rpe, Fpk, Pta, AlDH and ADH were added at 1 h delayed. Hps was excluded for the control. At each time point, samples were mixed with 8 M urea in a ratio of 10:1 to quench the reaction, then analyzed by GC-MS.

**Results**




| Reaction step | Substration | Production |
|---|---|---|
| 1 | $CO_2$ | CHOOH |
| 2 | CHOOH | $CH_2O$ |
| 3 | Ru5P, $CH_2O$ | H6P |
| 4 | H6P | F6P |
| 5 | F6P | E4P, AcP |
| 6 | F6P, E4P | G3P, S7P |
| 7 | G3P, S7P | X5P, R5P |
| 8 | X5P | Ru5P |
| 9 | R5P | Ru5P |
| 10 | AcP, CoA | AcCoA |
| 11 | AcCoA | $CH_3CHO$ |
| 12 | $CH_3CHO$ | $CH_3CH_2OH$ |

**Fig S1.** The results of thermodynamic data analysis in MDF-optimized concentrations.




| Enzyme | Mw (kDa) |
|---|---|
| Hps | 22.6 |
| Phi | 19.5 |
| Fpk | 92.6 |
| Pta | 34.6 |
| AlDH | 49.0 |
| Rpi | 22.8 |
| Rpe | 24.5 |
| Tkt | 72.2 |
| Tal | 35.1 |

**Fig S2.** SDS-PAGE of 9 recombinant purified enzymes for the CTE pathways. CTE,

$CO_2$ to ethanol; Hps, 3-hexulose-6-phosphate synthase; Phi, phosphohexulose isomerase; F/Xpk, phosphoketolase; Pta, phosphate acetyltransferase; Tal, transaldolase; Tkt, transketolase; Rpi, ribose-5 phosphate isomerase; Rpe, ribulose 5-phosphate epimerase; AlDH, acetaldehyde dehydrogenase.

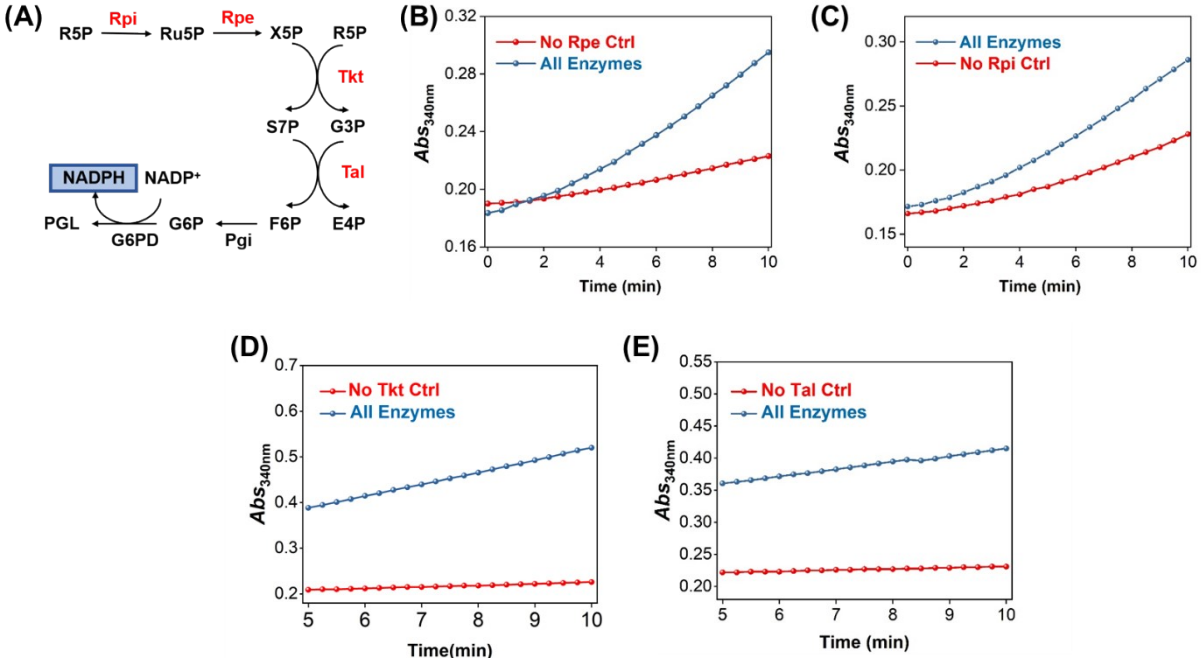



**Fig S3.** *In vitro* enzyme activity analysis of Rpi, Rpe, Tkt, Tal. Ru5P, ribulose-5 phosphate; F6P, fructose-6-phosphate; E4P, erythorse 4-phosphate; G3P, glyceraldehyde 3-phosphate; S7P, sedoheptulose 7-phosphate; R5P, ribose 5-phosphate; X5P, xylulose 5-phosphate.; PGL, 6-phosphogluconolactone; G6P, Glucose 6-phosphate. (A). The above enzyme activities were represented by the measurements of the production of NADPH. (B) *In vitro* enzyme activity of Rpe; (C) *In vitro* enzyme activity of Rpi; (D) *In vitro* enzyme activity of Tkt; (E) *In vitro* enzyme activity of Tal.

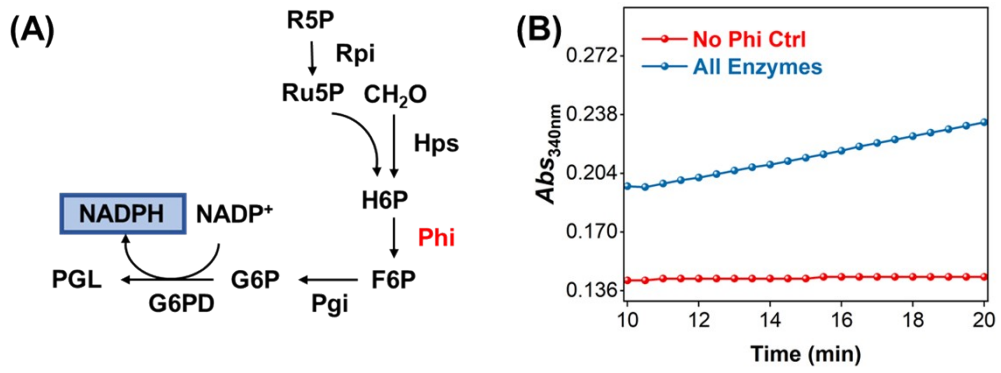


**Fig S4.** *In vitro* enzyme activity verification of Phi. (A) Pathway for Phi activity assays. (B) The activity of Phi *in vitro* represented by the measurements of the NADPH generation.

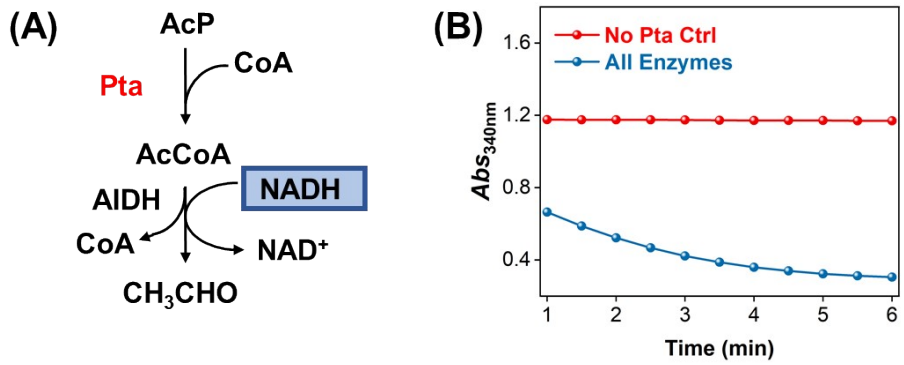



**Fig S5.** *In vitro* enzyme activity verification of Pta. CoA, coenzyme A; AcCoA, acetyl coenzyme A. (A) Pathway for Pta activity assays. (B) The activity of Pta *in vitro* depended on the measurements of the NADH consumption.

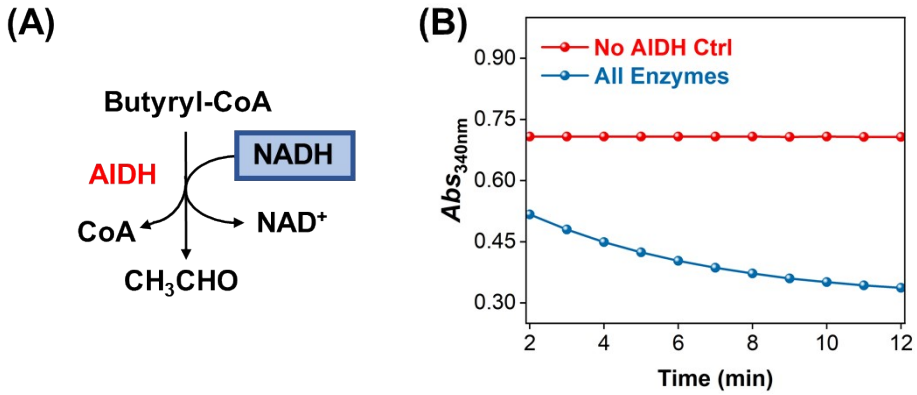



**Figure S6.** *In vitro* enzyme activity verification of AlDH. (A) Pathway for AlDH activity assays. (B) The activity of AlDH *in vitro* represented by the measurements of the consumption of NADH.

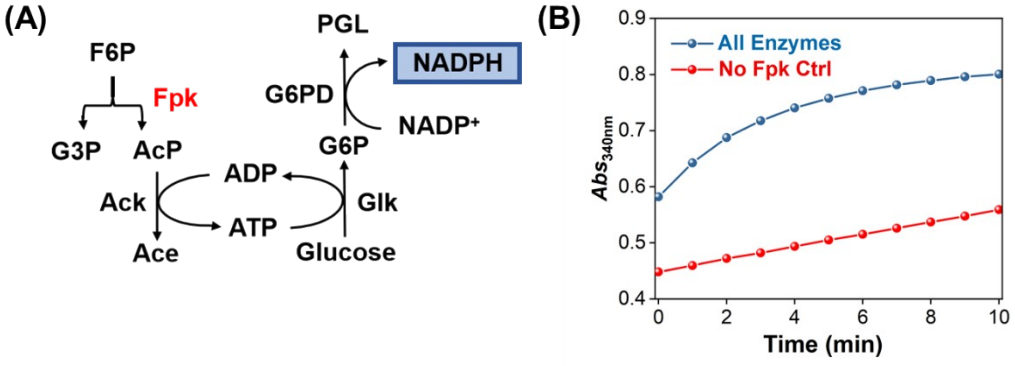


**Figure S7.** *In vitro* enzyme activity verification of Fpk. Ace, acetate; AcP, acetylphosphate; (A) Pathway for Fpk activity assays. (B) The activity of Fpk *in vitro* represented by the measurements of the generation of NADPH.

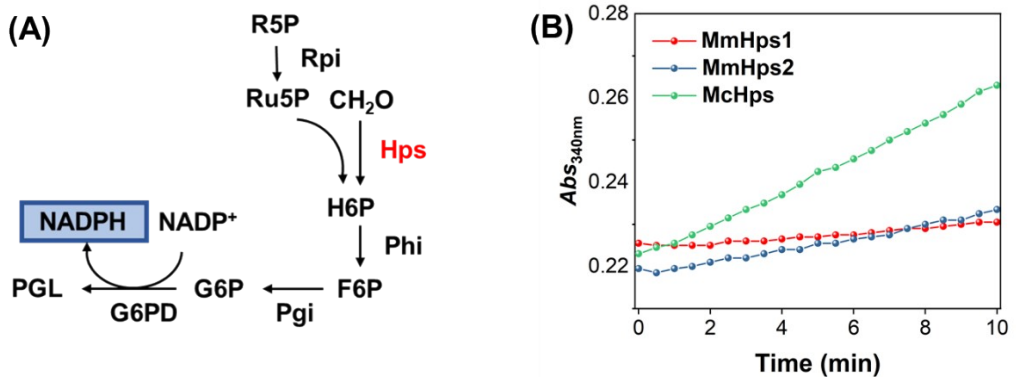



**Figure S8.** Comparison of *in vitro* enzyme activity of Hps screened by PPR. (A) Pathway for Hps activity assays. (B) The activity of Hps *in vitro* represented by the measurements of the NADPH generation.

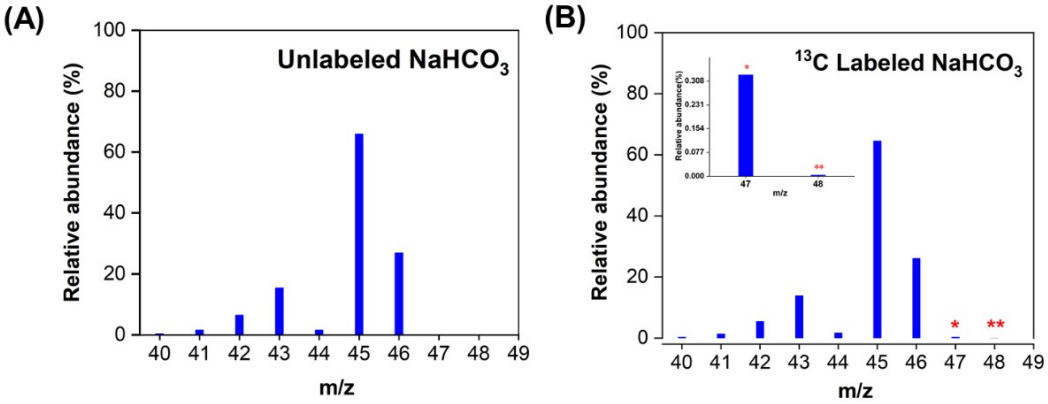



**Figure S9**. $^{13}$C Tracing from $^{13}$C-$NaHCO_3$ to ethanol. All spectra were normalized to the most abundant internal peak. (A) Mass spectrum of ethanol experimentally produced from unlabeled $NaHCO_3$ and R5P using the full CTE pathway. (B) Mass spectrum of ethanol experimentally produced from $^{13}$C-$HCO_3^-$ and unlabeled R5P using the full CTE pathway. One-labeled ethanol: [1-$^{13}$C]-ethanol, [2-$^{13}$C]-ethanol (one red asterisks); and double-labeled [1,2-$^{13}$C]-ethanol (two red asterisks).

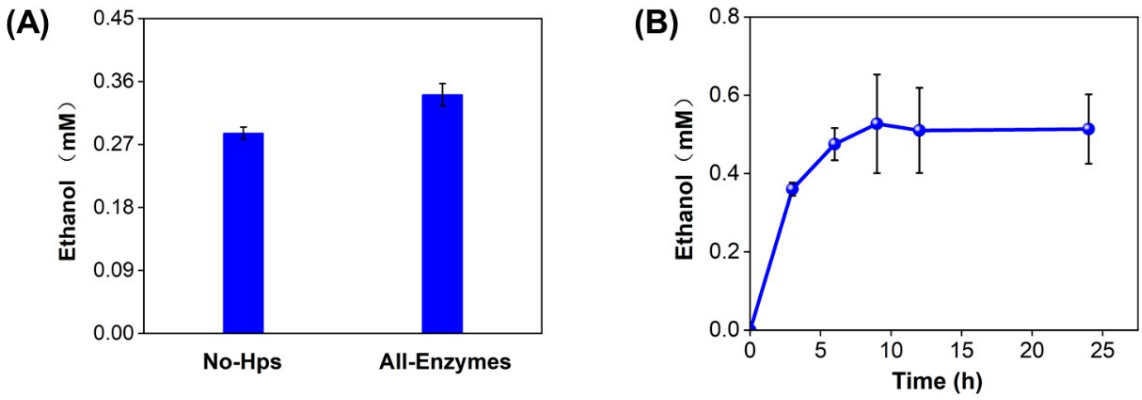



**Figure S10**. Yield of ethanol from $HCO_3^-$ through the full CTE pathway. (A) Ethanol production with Hps as the experimental variable. (B) Time course diagram of ethanol yield.

**Table S1.** Comparison of artificial *in vitro* carbon-fixation pathways

| Pathway | Substrate | Product | ATP/$CO_2$ (mol/mol) | Number of reactions | Carbon fixation rate [nmol/(min·mg)] | Oxygen tolerance | Ref |
|---|---|---|---|---|---|---|---|
| rGPS-MCG | $NaHCO_3$, crotonyl-CoA, phosphoenolpyruvate | acetyl-CoA, pyruvate, malate | 2.5 | 19 | 28.5 | Yes | [3] |
| CETCH | propionyl-CoA, $NaHCO_3$ | glyoxylate | 1 | 17 | 5.0 | Yes | [4] |
| POAP | sodium acetate $NaHCO_3$ | oxlate | 1 | 4 | 8.0 | No | [5] |
| ASAP | methanol (chemically hydrogenated by $CO_2$ and $H_2$) | starch | 0.5 | 11 | 22 | Yes | [4] |
| CTE | $CO_2$ | ethanol | 0 | 11 | 4.3 | Yes | This work |

**Table S2.** Summary of origins of CTE enzymes

| ABB. | Enzyme full name | Strain | Ref |
|---|---|---|---|
| PaFDH | Formate dehydrogenase | *Paracoccus* sp. MKU1 | [6] |
| ΔPaFDH48 | Formate dehydrogenase mutant | *Paracoccus* sp. MKU1 | [7] |
| SzFaldDH | Formaldehyde dehydrogenase | *Streptomyces zinciresistens* | [8] |
| BmFaldDH | Formaldehyde dehydrogenase | *Burkholderia multivorans* | [9] |
| BsHps | 3-hexulose-6-phosphate synthase | *Bacillus subtilis* | [2] |
| BmHps | 3-hexulose-6-phosphate synthase | *Bacillus methanolicus* | [10] |
| McHps | 3-hexulose-6-phosphate synthase | *Methylococcus capsulatus* | [2] |
| MmHps1 | 3-hexulose-6-phosphate synthase | *Methylocaldum marinu* | This work |
| MmHps2 | 3-hexulose-6-phosphate synthase | *Methylocaldum marinu* | This work |
| Phi | 6-phospho-3-hexuloisomerase | *Methylobacillus flagellatus* | [10] |
| Tal | Transaldolase | *Escherichia coli (JCL16)* | [2] |
| Tkt | Transketolase | *Escherichia coli (JCL16)* | [2] |
| Rpi | Ribose-5 phosphate isomerase | *Escherichia coli (JCL16)* | [2] |
| Rpe | Ribulose 5-phosphate epimerase | *Escherichia coli (JCL16)* | [2] |
| Fpk | Phosphoketolase | *Bifidobacterium adolescentis* | [2] |
| Pta | Phosphate acetyltransferase | *Bacillus subtilis (strain 168)* | [2] |
| AlDH | Aldehyde dehydrogenase | *Salmonella enterica* | [2] |

**Table S3.** List of combinatorial reactions.

| Enzyme | Reaction | Direction | EC number |
|---|---|---|---|
| FDH | NADH + $CO_2$ <=> $NAD^+$ + formate | ↔ | 1.2.1.2 |
| FaldDH | NADH + formate <=> $NAD^+$ + formaldehyde + $H_2O$ | ↔ | 1.2.1.46 |
| Hps | D-ribulose 5-phosphate + formaldehyde <=> D-arabino-hex-3-ulose 6-phosphate | → | 4.1.2.43 |
| Phi | D-arabino-hex-3-ulose 6-phosphate <=> D-fructose 6-phosphate | ↔ | 5.3.1.27 |
| Tal | D-glyceraldehyde 3-phosphate + D-sedoheptulose 7-phosphate <=> β-D-fructose 6-phosphate + D-erythrose 4-phosphate | ↔ | 2.2.1.2 |
| Tkt | D-glyceraldehyde 3-phosphate + D-sedoheptulose 7-phosphate <=> aldehydo-D-ribose 5-phosphate + D-xylulose 5-phosphate | ↔ | 2.2.1.1 |
| Rpi | aldehydo-D-ribose 5-phosphate <=> D-ribulose 5-phosphate | ↔ | 5.3.1.6 |
| Rpe | D-ribulose 5-phosphate <=> D-xylulose 5-phosphate | ↔ | 5.1.3.1 |
| Fpk | F6P + phosphate <=> acetyl phosphate + D-erythrose 4-phosphate + $H_2O$ | → | 4.1.2.9 |
| Xpk | phosphate + D-xylulose 5-phosphate <=> D-glyceraldehyde 3-phosphate + acetyl phosphate + $H_2O$ | → | 4.1.2.22 |
| Pta | acetyl-CoA + phosphate <=> acetyl phosphate + CoA | ↔ | 2.3.1.8 |
| AlDH | acetaldehyde + CoA + $NAD^+$ <=> acetyl-CoA + $H^+$ + NADH | ↔ | 1.2.1.10 |

**Table S4.** List of synthesized enzyme sequences.

| | |
|---|---|
| Rpi | MTQDELKKAVGWAALQYVQPGTIVGVGTGSTAAHFIDALGT<br>MKGQIEGAVSSSDASTEKLKSLGIHVFDLNEVDSLGIYVDGAD<br>EINGHMQMIKGGGAALTREKIIASVAEKFICIADASKQVDILGK<br>FPLPVEVIPMARSAVARQLVKLGGRPEYRQGVVTDNGNVILD<br>VHGMEILDPIAMENAINAIPGVVTVGLFANRGADVALIGTPDG<br>VKTIVK |
| Tkt | MSSRKELANAIRALSMDAVQKAKSGHPGAPMGMADIAEVLW<br>RDFLKHNPQNPSWADRDRFVLSNGHGSMLIYSLLHLTGYDLP<br>MEELKNFRQLHSKTPGHPEVGYTAGVETTTGPLGQGIANAVG<br>MAIAEKTLAAQFNRPGHDIVDHYTYAFMGDGCMMEGISHEVC<br>SLAGTLKLGKLIAFYDDNGISIDGHVEGWFTDDTAMRFEAYG<br>WHVIRDIDGHDAASIKRAVEEARAVTDKPSLLMCKTIIGFGSPN<br>KAGTHDSHGAPLGDAEIALTREQLGWKYAPFEIPSEIYAQWDA<br>KEAGQAKESAWNEKFAAYAKAYPQEAAEFTRRMKGEMPSDF<br>DAKAKEFIAKLQANPAKIASRKASQNAIEAFGPLLPEFLGGSAD<br>LAPSNLTLWSGSKAINEDAAGNYIHYGVREFGMTAIANGISLH<br>GGFLPYTSTFLMFVEYARNAVRMAALMKQRQVMVYTHDSIG<br>LGEDGPTHQPVEQVASLRVTPNMSTWRPCDQVESAVAWKYG<br>VERQDGPTALILSRQNLAQQERTEEQLANIARGGYVLKDCAG<br>QPELIFIATGSEVELAVAAYEKLTAEGVKARVVSMPSTDAFDK<br>QDAAYRESVLPKAVTARVAVEAGIADYWYKYVGLNGAIVGM<br>TTFGESAPAELLFEEFGFTVDNVVAKAKELL |
| Tal | MTDKLTSLRQYTTVVADTGDIAAMKLYQPQDATTNPSLILNA<br>AQIPEYRKLIDDAVAWAKQQSNDRAQQIVDATDKLAVNIGLEI<br>LKLVPGRISTEVDARLSYDTEASIAKAKRLIKLYNDAGISNDRI<br>LIKLASTWQGIRAAEQLEKEGINCNLTLLFSFAQARACAEAGV<br>FLISPFVGRILDWYKANTDKKEYAPAEDPGVVSVSEIYQYYKE<br>HGYETVVMGASFRNIGEILELAGCDRLTIAPALLKELAESEGAI<br>ERKLSYTGEVKARPARITESEFLWQHNQDPMAVDKLAEGIRKF<br>AIDQEKLEKMIGDLL |
| F/Xpk | MTSPVIGTPWKKLNAPVSEEAIEGVDKYWRAANYLSIGQIYLR<br>SNPLMKEPFTREDVKHRLVGHWGTTPGLNFLIGHINRLIADHQ<br>QNTVIIMGPGHGGPAGTAQSYLDGTYTEYFPNITKDEAGLQKF<br>FRQFSYPGGIPSHYAPETPGSIHEGGELGYALSHAYGAVMNNP<br>SLFVPAIVGDGEAETGPLATGWQSNKLINPRTDGIVLPILHLNG<br>YKIANPTILSRISDEELHEFFHGMGYEPYEFVAGFDNEDHLSIH<br>RRFAELFETVFDEICDIKAAAQTDDMTRPFYPMIIFRTPKGWTC |

| | |
|---|---|
| | PKFIDGKKTEGSWRSHQVPLASARDTEAHFEVLKNWLESYKP<br>EELFDENGAVKPEVTAFMPTGELRIGENPNANGGRIREELKLP<br>KLEDYEVKEVAEYGHGWGQLEATRRLGVYTRDIIKNNPDSFRI<br>FGPDETASNRLQAAYDVTNKQWDAGYLSAQVDEHMAVTGQ<br>VTEQLSEHQMEGFLEGYLLTGRHGIWSSYESFVHVIDSMLNQH<br>AKWLEATVREIPWRKPISSMNLLVSSHVWRQDHNGFSHQDPG<br>VTSVLLNKCFNNDHVIGIYFPVDSNMLLAVAEKCYKSTNKINA<br>IIAGKQPAATWLTLDEARAELEKGAAEWKWASNVKSNDEAQI<br>VLAATGDVPTQEIMAAADKLDAMGIKFKVVNVVDLVKLQSA<br>KENNEALSDEEFAELFTEDKPVLFAYHSYARDVRGLIYDRPNH<br>DNFNVHGYEEQGSTTTPYDMVRVNNIDRYELQAEALRMIDAD<br>KYADKINELEAFRQEAFQFAVDNGYDHPDYTDWVYSGVNTN<br>KQGAISATAATAGDNE |
| Phi | MNKYQELVVNKLTNVINNTAEGYDDKILSMVDAAGRTFLGG<br>AGRSLLVSRFFAMRLVHAGYQVSMVGEVVTPSIQAGDLFIVIS<br>GSGSTETLMPLVRKAKSQGAKVIVISMKAQSPMAELADLVVPI<br>GGNDAHAFDKTHGMPMGTIFELSTLWFLEATIAKLIDQKGLTE<br>EGMRAIHANLE |
| Pta | MADLFSTVQEKVAGKDVKIVFPEGLDERILEAVSKLAGNKVL<br>NPIVIGNENEIQAKAKELNLTLGGVKIYDPHTYEGMEDLVQAF<br>VERRKGKATEEQARKALLDENYFGTMLVYKGLADGLVSGAA<br>HSTADTVRPALQIIKTKEGVKKTSGVFIMARGEEQYVFADCAI<br>NIAPDSQDLAEIAIESANTAKMFDIEPRVAMLSFSTKGSAKSDE<br>TEKVADAVKIAKEKAPELTLDGEFQFDAAFVPSVAEKKAPDSE<br>IKGDANVFVFPSLEAGNIGYKIAQRLGNFEAVGPILQGLNMPV<br>NDLSRGCNAEDVYNLALITAAQAL |
| BsHps | MELQLALDLVNIPEAIELVKEVEQYIDVVEIGTPVVINEGLRAV<br>KEIKEAFPQLKVLADLKIMDAGGYEIMKASEAGADIITVLGAT<br>DDATIKGAVEEAKKQKKKILVDMINVKDIESRAKEIDALGVDY<br>ICVHTGYDLQAEGKNSFEELTTIKNTVKNAKTAIAGGIKLDTLP<br>EVIQQKPDLVIVGGGITSAADKAETASKMKQLIVQG |
| McHps | MARPLIQLALDTLDIPQTLKLASLTAPYVDIFEIGTPSIKHNGIA<br>LVKEFKKRFPNKLLLVDLKTMDAGEYEATPFFAAGADITTVLG<br>VAGLATIKGVINAANKHNAEVQVDLINVPDKAACARESAKAG<br>AQIVGIHTGLDAQAAGQTPFADLQAIAKLGLPVRISVAGGIKAS<br>TAQQVVKTGANIIVVGAAIYGAASPADAAREIYEQVVAASA |

| | |
|---|---|
| BmHps | MELQLALDLVNIEEAKQVVAEVQEYVDIVEIGTPVIKIWGLQA<br>VKAVKDAFPHLQVLADMKTMDAAAYEVAKAAEHGADIVTIL<br>AAAEDVSIKGAVEEAKKLGKKILVDMIAVKNLEERAKQVDEM<br>GVDYICVHAGYDLQAVGKNPLDDLKRIKAVVKNAKTAIAGGI<br>KLETLPEVIKAEPDLVIVGGGIANQTDKKAAAEKINKLVKQGL |
| MmHps1 | MARPLIQLALDSLDVNQTLKLAGLAAPYVDIFEIGTPCIKHNGV<br>SLVKELRKKYPNKLILVDLKTMDAGEYEATPFYAAGADICTVL<br>GVSGLPTIAGVIKAANAHNAEVQVDLINVPDKIACARESAKLG<br>AHIIGVHTGLDAQAAGQTPFADLQAISRLKLPVRISVAGGIKQS<br>TVQQVARAGANIVVVGAAIYGAASPADAAREICELASAA |
| MmHps2 | MARPLIQLALDSLDIDQTLRLANITAPYIDIFEIGTPCIKHNGIAL<br>VKELKQRYPDKLILVDLKTMDAGEYEATPFFAAGADICTVLG<br>VSGLPTIAGVIKAAKAHNAEVQVDLINVPDKLVCAREAAKLG<br>AHIIGVHTGLDAQAAGQTPFADLQAIAKLGLPVRISVAGGINQ<br>VTVRQVAKTGADIIVVGAAIYGAPSPSDAAREIRELVEGKHHK<br>FIMSKLAGVLGSTDARYEARLTTMLDRASRVFIAGAGRSGLVA<br>KFFGMRLMHGGYDAYIVGEVVTPSIRKGDLFIVISGSGETETM<br>LAYTKRAKQMGANIGLITTKDSSTIGDLADVVFRIGSPEQYRK<br>VIGMPMGTTFELSTLLLLEATVSHIIHAKKIPEEQMRTRHANLE |
| AlDH | MNTSELETLIRTILSEQLTTPAQTPVQPQGKGIFQSVSEAIDAAH<br>QAFLRYQQCPLKTRSAIISAMRQELTPLLAPLAEESANETGMG<br>NKEDKFLKNKAALDNTPGVEDLTTTALTGDGGMVLFEYSPFG<br>VIGSVAPSTNPTETIINNSISMLAAGNSIYFSPHPGAKKVSLKLIS<br>LIEEIAFRCCGIRNLVVTVAEPTFEATQQMMAHPRIAVLAITGG<br>PGIVAMGMKSGKKVIGAGAGNPPCIVDETADLVKAAEDIINGA<br>SFDYNLPCIAEKSLIVVESVAERLVQQMQTFGALLLSPADTDK<br>LRAVCLPEGQANKKLVGKSPSAMLEAAGIAVPAKAPRLLIALV<br>NADDPWVTSEQLMPMLPVVKVSDFDSALALALKVEEGLHHT<br>AIMHSQNVSRLNLAARTLQTSIFVKNGPSYAGIGVGGEGFTTFT<br>IATPTGEGTTSARTFARSRRCVLTNGFSIR |

## References


1. E. Noor, H. S. Haraldsdottir, R. Milo and R. M. Fleming, *PLoS Comput. Biol.*, 2013, **9**, e1003098.
2. I. W. Bogorad, C. T. Chen, M. K. Theisen, T. Y. Wu, A. R. Schlenz, A. T. Lam and J. C. Liao, *Proc. Natl. Acad. Sci. USA*, 2014, **111**, 15928-15933.

3. S. Luo, P. P. Lin, L.-Y. Nieh, G.-B. Liao, P.-W. Tang, C. Chen and J. C. Liao, *Nat. Catal.*, 2022, **5**, 154-162.
4. S. Bierbaumer, M. Nattermann, L. Schulz, R. Zschoche, T. J. Erb, C. K. Winkler, M. Tinzl and S. M. Glueck, *Chem. Rev.*, 2023, DOI: 10.1021/acs.chemrev.2c00581.
5. L. Xiao, G. Liu, F. Gong, H. Zhu, Y. Zhang, Z. Cai and Y. Li, *ACS Catal.*, 2021, **12**, 799-808.
6. X. Ji, Y. Xue, Z. Li, Y. Liu, L. Liu, P. K. Busk, L. Lange, Y. Huang and S. Zhang, *Green Chem.*, 2021, **23**, 6990-7000.
7. Y. Huang, Y. Xue, X. Ji and S. Zhang, CN202211598631.9, 2022.
8. Y. Huang, X. Ji, Y. Xue, B. Guo and S. Zhang, CN202211599488.5, 2022.
9. R. K. Singh, R. Singh, D. Sivakumar, S. Kondaveeti, T. Kim, J. Li, B. H. Sung, B.-K. Cho, D. R. Kim, S. C. Kim, V. C. Kalia, Y.-H. P. J. Zhang, H. Zhao, Y. C. Kang and J.-K. Lee, *ACS Catal.*, 2018, **8**, 11085-11093.
10. J. E. N. Muller, F. Meyer, B. Litsanov, P. Kiefer, E. Potthoff, S. Heux, W. J. Quax, V. F. Wendisch, T. Brautaset, J. C. Portais and J. A. Vorholt, *Metab. Eng.*, 2015, **28**, 190-201.