Electronic Supplementary Material (ESI) for Green Chemistry.

Supporting Information

ATP Regeneration by a Single Polyphosphate Kinase Powers Multigram-Scale Aldehyde Synthesis *In Vitro*

Michele Tavanti, Joseph Hosford, Richard C. Lloyd and Murray J. B. Brown*

Supporting Figures

Table S1. PPK2 enzymes included in the initially screened collection of enzymes. Newly identified protein sequences were assigned to a PPK2 class based on the protein sequence employed to carry out BLAST searches.

UniProt	Short Name	Organism	Transformation	Class	Origin Classification	Reference
M9XB82	mrPPK	*Meiothermus ruber*	AMP→ATP	PPK2-III	Thermophile	[1]
Q9RY20	drPPK	*Deinococcus radiodurans*	AMP→ATP	PPK2-III	Mesophile	[1]
A2SL03	mpPPK	*Methylibium petroleiphilum*	AMP→ATP	PPK2-III	Mesophile	This work
Q8DI82	tePPK	*Thermosynechococcus elongatus*	AMP→ATP	PPK2-III	Thermophile	[1]
Q83XD3	ajPAP	*Acinetobacter johnsonii*	AMP→ADP	PPK2-II	Mesophile	[2]
Q9HYF1	paPAP	*Pseudomonas aeruginosa*	AMP→ADP	PPK2-II	Mesophile	[3]
H0UPE2	tvPAP	*Thermanaerovibrio velox*	AMP→ADP	PPK2-II	Thermophile	This work
A0A1H2T5S6	mhPAP	*Methanohalophilus halophilus*	AMP→ADP	PPK2-II	Halophile	This work
Q6N140	rpPPK	*Rhodopseudomonas palustris*	ADP→ATP	PPK2-I	Mesophile	[3]
Q92SA6	rmPPK	*Rhizobium meliloti*	ADP→ATP	PPK2-I	Mesophile	[3]
A0A1I4IPN4	hrPPK	*Halogranum rubrum*	ADP→ATP	PPK2-I	Halophile	This work

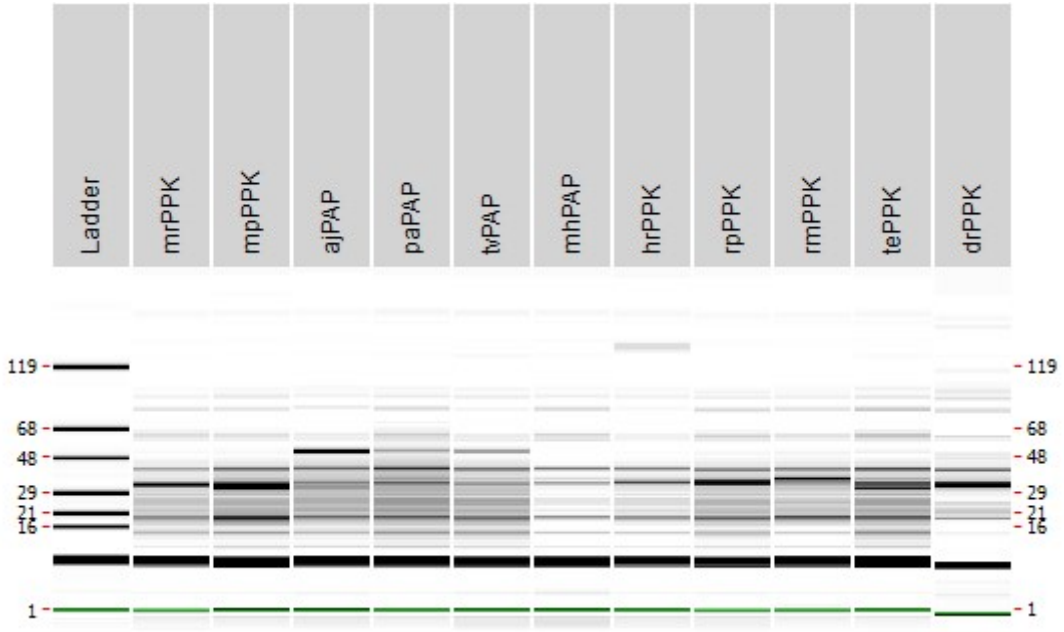

Figure S1. Capillary electrophoresis analysis of the soluble fraction obtained after lysis of E. coli cells expressing PPK2 enzymes listed in Table S1. Estimated molecular weights in kDa are indicated on each side of the figure. Bands corresponding to proteins with mass of ~30 kDa were expected for enzymes indicated as "PPK" (PPK2-I and PPK2-III), whereas bands corresponding to proteins with mass of ~50-60 kDa were expected for enzymes indicated as PAP (PPK2-II).

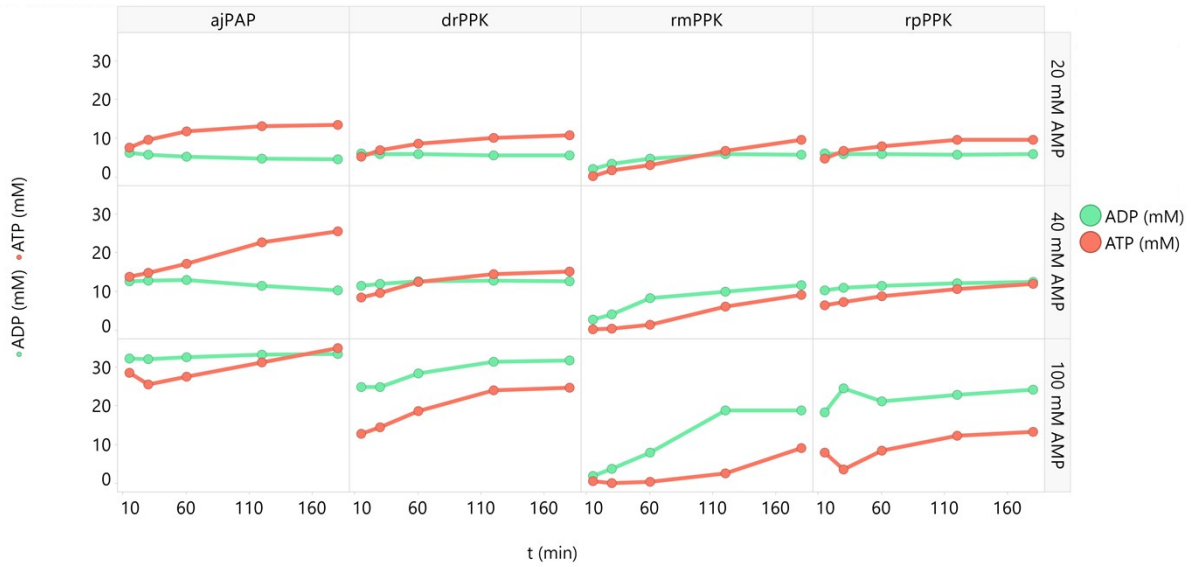

Figure S2. Reaction time course for AMP phosphorylation with different PPKs and at different AMP loadings.

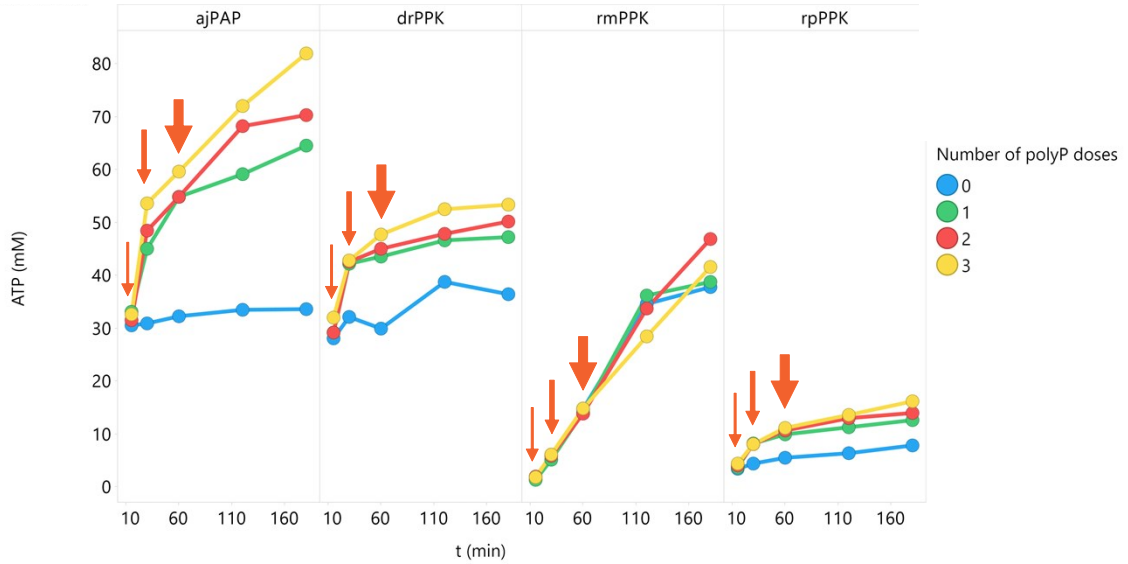

Figure S3. Reaction time course for AMP phosphorylation with different PPKs and polyP feeding. Reactions contained 200 mM polyP (based on phosphate units) at the beginning of the reaction. 100 mM polyP doses were added at 15, 30 and 60 minutes (see arrows). Reaction time course for biotransformations carried out without polyP feeding (blue), with 1 polyP dose (added at 15 minutes, green), with 2 polyP doses (added at 15 and 30 minutes, red) and 3 polyP doses (added at 15, 30 and 60 minutes, yellow) are shown.

Table S2. Desired performance criteria for ATP recycling and ATP formation rate extrapolated at 1 mg/mL ajPAP loading based on the initial ATP formation rate (30 min) observed at 10 mg/mL ATP-recycling enzyme loading. A linear relationship between enzyme loading and ATP formation rate was assumed. A model product molecular weight of 150 Da was considered for these calculations.

Biocatalyst Formulation	Lyophilised cell-free extract
Biocatalyst Loading (mg/mL)	< 5
ATP-Recycling Enzyme(s) Loading (mg/mL)	1
Final Product Concentration (g/L)	100
Product Molecular Weight (Da)	150
Final Product Concentration (M)	0.67
Reaction Time (h)	24
Desired Product Formation Rate (mM/min)	0.46
Observed ATP Formation Rate (mM/min) (@ 10 mg/mL)	1.8
Extrapolated ATP Formation Rate (mM/min) (@ 1 mg/mL)	0.18

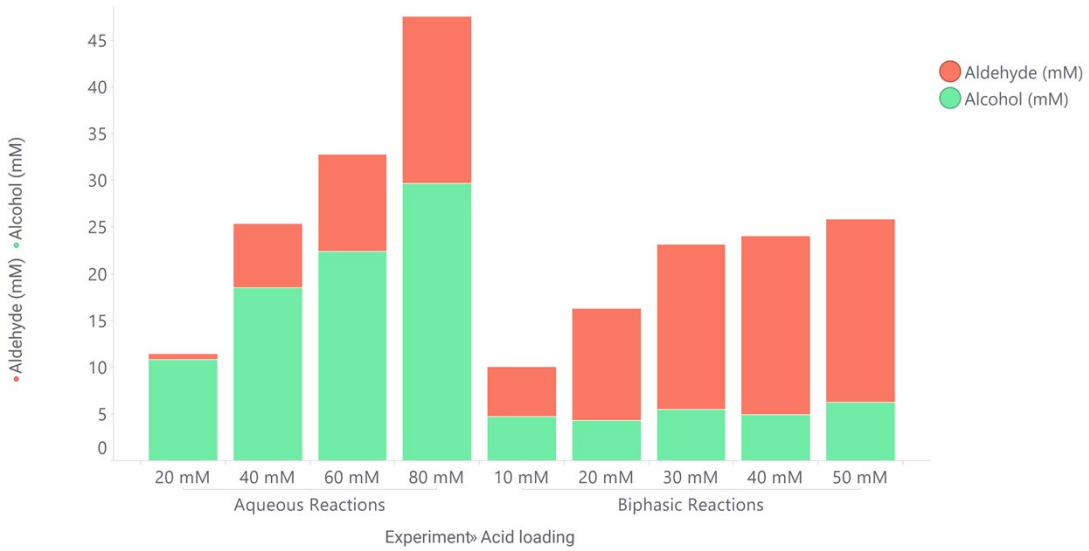

*Figure S4. Product concentrations after 4-methoxybenzoic acid reduction **1** by CAR33 with dual cofactor recycling in aqueous or biphasic reactions and at different substrate loadings. The addition of a toluene overlay (50% v/v) caused the alcohol:aldehyde ratio to drop dramatically from 94% to 26% for reactions carried out with 20 mM substrate, whereas a significant change was also observed for reactions carried out with 40 mM substrate, for which the percentage of total products represented by the alcohol went from 73% in aqueous conditions to 20% for biphasic reactions.*

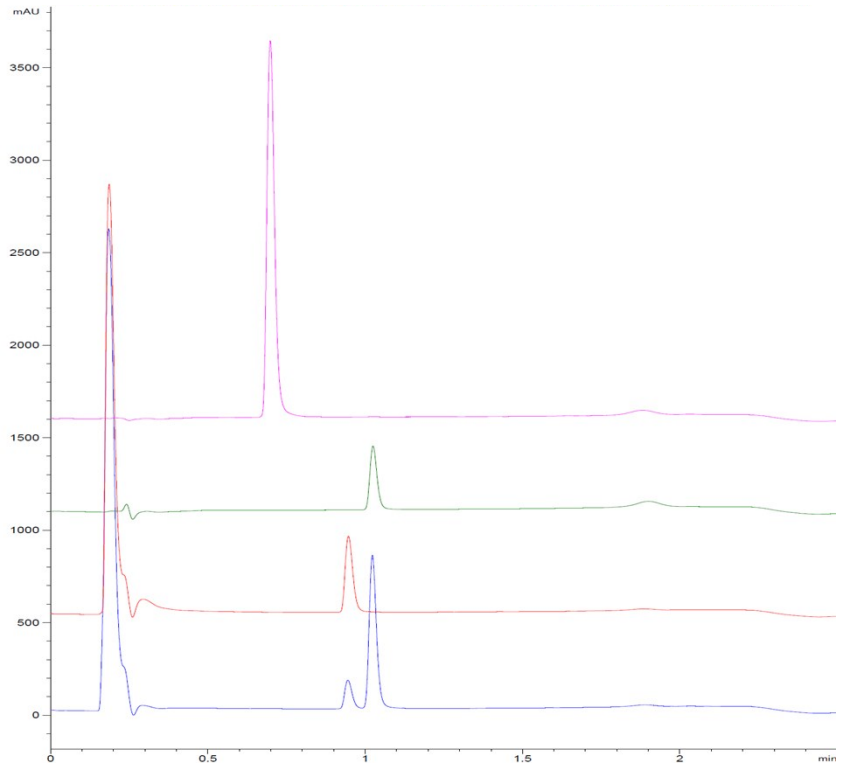

Figure S5. UPLC traces for 4-methoxybenzoic acid reduction with purified CAR33 and stoichiometric cofactors. The trace corresponding to the biotransformation (blue chromatogram) does not show detectable levels of alcohol overreduction product. UPLC traces corresponding to 4-methoxybenzyl alcohol (magenta), 4-methoxybenzaldehyde (green) and 4-methoxybenzoic acid (red) are also shown. Experiments (0.1 mL biotransformation volume) were carried out with 0.2 mg/mL purified CAR33, 10 mM NADPH, 10 mM ATP, 50 mM $MgCl_2$ and 10 mM 4-methoxybenzoic acid. Reactions were carried out in 1.5 mL Eppendorf tubes, 800 rpm, 24 h using and Eppendorf Thermomixer.

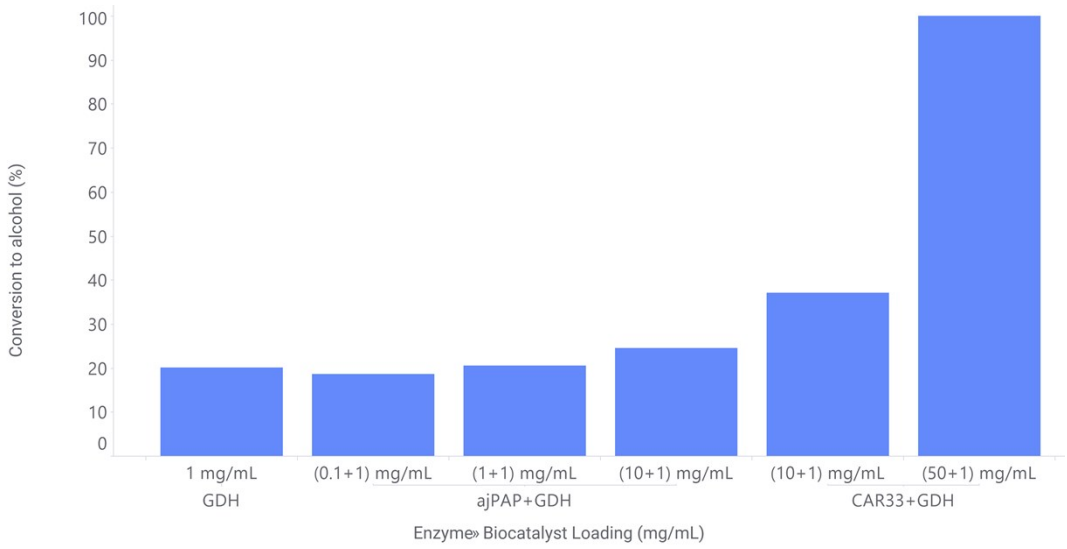

Figure S6. Aldehyde reduction control experiments with lyophilised cell free extract.

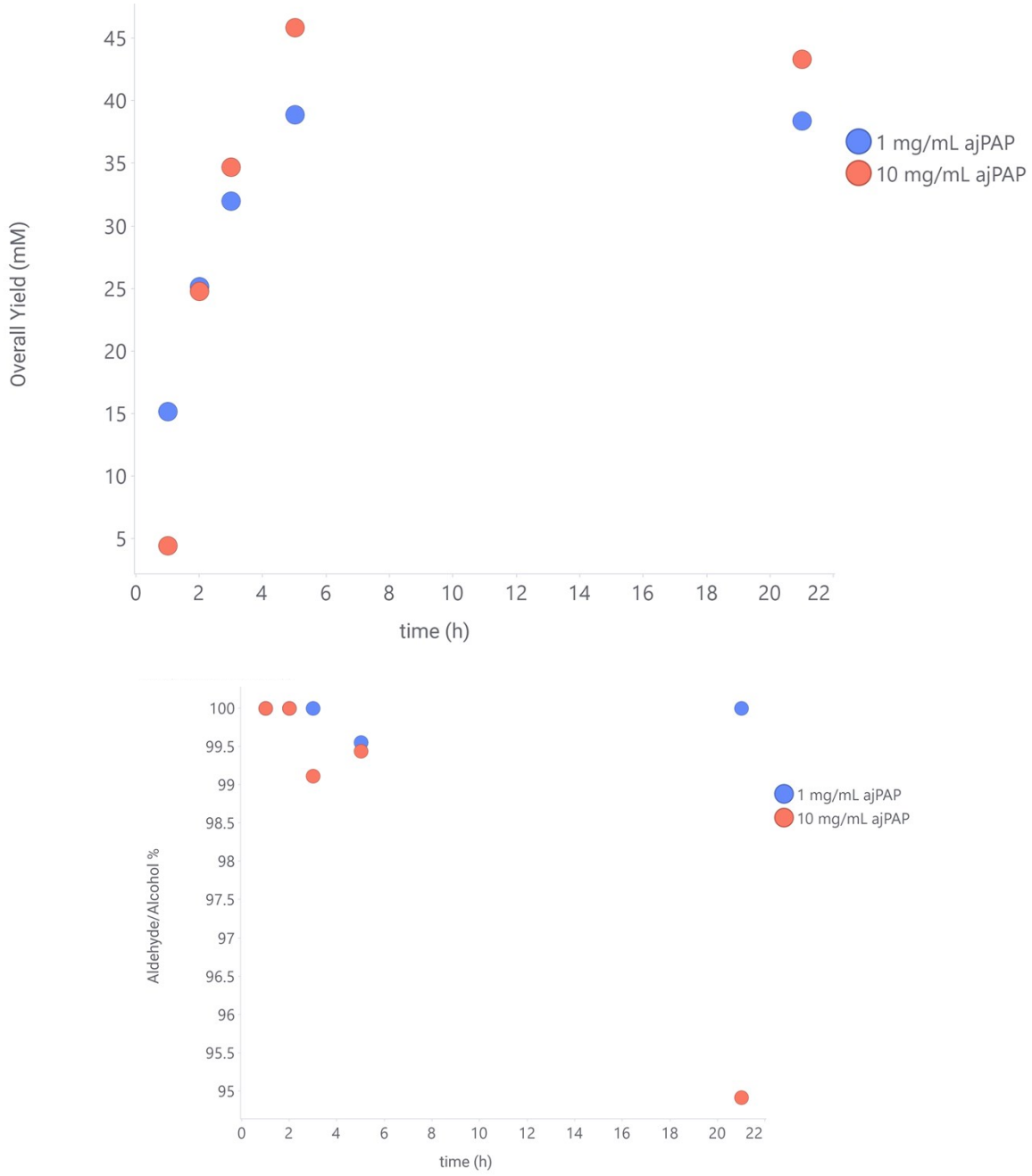

Figure S7. Time course for the CAR-catalysed 4-methoxybenzoic acid reduction with dual cofactor recycling on a 30 mL scale (150 mM substrate and 300 mM polyP, top) and product composition over time (bottom). The overall yield is the sum of aldehyde and alcohol yields.

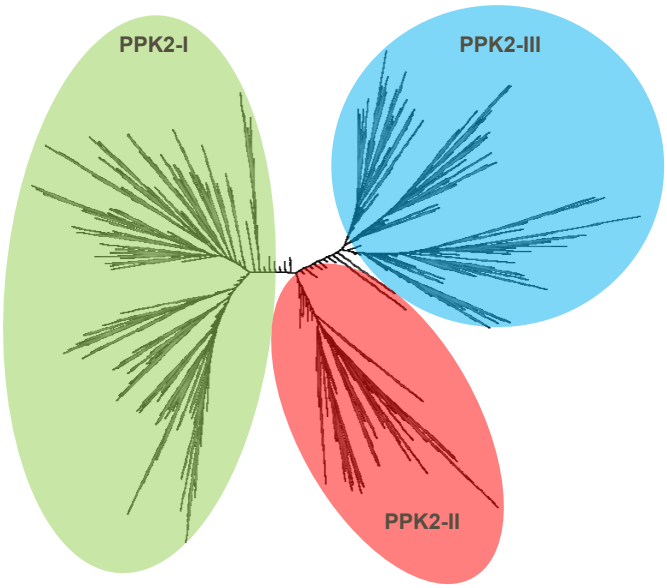

Figure S8. Phylogenetic analysis of PPK2 containing proteins (InterPro IPR022488). Clades were assigned based on the localisation of reviewed PPKs (UniProt) and the presence of signature residues. The localisation of the following PPK2-I was verified (UniProt IDs): A9CKA8, Q8NM65, A0QQV6, Q9I6Z1, Q5LU04, Q5LSN8, Q9I154, O05877, Q930V2, Q92ZU4, Q92SA6, Q5LX16. The localisation of the following PPK2-II was verified (UniProt IDs): Q9HYF1, Q886D9. The localisation of the following PPK2-III was verified (UniProt IDs): A1R8G0, Q11YW6, Q1IW43, Q9RY20, D7BBL3, M9XB82, Q6N140 and Q8DI82.[4]

Table S3. Previously reported PPK2-III enzymes included in the PPK2-III panel developed in this work.

UniProt	Short Name	Organism	Reference
M9XB82	mrPPK	*Meiothermus ruber*	[1]
Q9RY20	drPPK	*Deinococcus radiodurans*	[1]
A2SL03	mpPPK	*Methylibium petroleiphilum*	This work
A1R8G0	aaPPK	*Paenarthrobacter aurescens*	[4]
Q11YW6	chPPK	*Cytophaga hutchinsonii*	[4]
Q8DI82	tePPK	*Thermosynechococcus elongatus*	[1]
n.a.	chPPK truncated	*Cytophaga hutchinsonii*	[4]
D7BBL3	msPPK	*Meiothermus silvanus*	[1]
Q1IW43	dgPPK	*Deinococcus geothermalis*	[1]

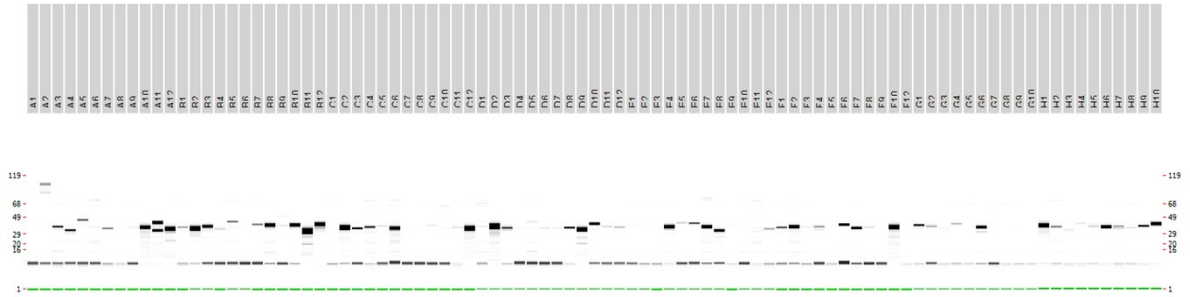

Figure S9. Capillary electrophoresis analysis of purified PPKs. The position of each protein in the panel is indicated on top of the figure. Estimated molecular weights in kDa are indicated on each side of the figure. Details about the identity of proteins in each well can be found in Table S4.

Table S4. Protein identities (UniProt ID) and well location for PPK2 enzymes included in the panel.

Protein Name	Well Location	UniProt_ID
PPK1	A01	A0A023D3G5_ACIMT
PPK9	A02	A0A404PG99_9ZZZZ
PPK17	A03	A0A1M4SFF4_9ACTN
msPPK (PPK26)	A04	D7BBL3_MEISD
PPK34	A05	A0A0E3QUX9_METBA
PPK43	A06	A0A1V4U2Y6_9EURY
PPK52	A07	A0A2V2N5B3_9EURY
PPK61	A08	A0A0F9PPZ6_9ZZZZ
PPK69	A09	A0A1J5T696_9ZZZZ
PPK77	A10	A0A3R2J6U0_9ZZZZ
PPK85	A11	A0A407LN77_9ZZZZ
aaPPK	A12	A1R8G0
PPK2	B01	A0A2W1NTC2_9FLAO
PPK10	B02	A0A023D3G5_ACIMT
PPK18	B03	A0A1V4HNG6_9BACL
PPK27	B04	E4U536_OCEP5
PPK35	B05	A0A0E3SF50_9EURY
PPK45	B06	A0A1V4UIP1_9EURY
PPK54	B07	A0A3A5WLP0_9EURY
PPK62	B08	A0A0F9VM92_9ZZZZ
PPK70	B09	A0A1J5TD51_9ZZZZ
PPK78	B10	A0A3R2T7T3_9ZZZZ
PPK86	B11	A0A408TAH8_9ZZZZ
chPPK	B12	Q11YW6
PPK3	C01	A0A148N6T5_9GAMM
PPK11	C02	A0A2G6QQJ7_9SPIO
PPK19	C03	A0A2W1NTC2_9FLAO
PPK28	C04	E8N0Z9_ANATU
PPK37	C05	A0A101H5Y8_9EURY
PPK46	C06	A0A1V4ZAW4_9EURY
PPK55	C07	A0A3N5SB38_9ARCH
PPK63	C08	A0A0W8EH77_9ZZZZ
PPK71	C09	A0A1J5TNI1_9ZZZZ
PPK79	C10	A0A3R2YN22_9ZZZZ
PPK87	C11	E6PTV2_9ZZZZ
drPPK	C12	Q9RY20
PPK4	D01	A0A1E7GDI9_9EURY
PPK12	D02	A0A3D5XRJ5_9FIRM

PPK20	D03	A0A3B8TX18_PSESP
PPK29	D04	A0A148N6T5_9GAMM
PPK38	D05	A0A139CH24_9EURY
PPK47	D06	A0A1V5IHL6_9EURY
PPK56	D07	A0A3N5VUA2_9ARCH
PPK64	D08	A0A160TD52_9ZZZZ
PPK72	D09	A0A2P2C180_9ZZZZ
PPK80	D10	A0A3R3U2L0_9ZZZZ
PPK88	D11	E6QVY4_9ZZZZ
mrPPK	D12	M9XB82
PPK5	E01	A0A2D6AAC9_9EURY
PPK13	E02	A0A4V3DS02_9MYCO
PPK21	E03	C7LYH8_ACIFD
chPPK Truncated	E04	n.a.
PPK39	E05	A0A1E7GDI9_9EURY
PPK48	E06	A0A2A2HQL7_9EURY
PPK57	E07	A7I4H2_METB6
PPK65	E08	A0A1J5PSL6_9ZZZZ
PPK73	E09	A0A381YJF6_9ZZZZ
PPK81	E10	A0A3R3ZC33_9ZZZZ
PPK89	E11	A0A404PG99_9ZZZZ
mpPPK	E12	A2SL03
PPK6	F01	M1QDU0_METMZ
PPK14	F02	F0SXQ4_SYNGF
PPK22	F03	A0A1H5N4P5_9MICO
dgPPK (PPK31)	F04	Q1IW43_DEIGD
PPK40	F05	A0A1Q7BC62_9ARCH
PPK49	F06	A0A2D6AAC9_9EURY
PPK58	F07	E1RJ42_METP4
PPK66	F08	A0A1J5QFM0_9ZZZZ
PPK74	F09	A0A382C2S6_9ZZZZ
PPK82	F10	A0A3R5FBT2_9ZZZZ
mCherry	F11	X5DSL3
tePPK	F12	Q8DI82
PPK7	G01	A0A1J5T696_9ZZZZ
PPK15	G02	A0A1D2P985_9CYAN
PPK24	G03	A0A0S7ASE9_MEIRU
PPK32	G04	A0A075MP24_9ARCH
PPK41	G05	A0A1Q7MEW4_9CREN
PPK50	G06	A0A2E2TV69_9EURY
PPK59	G07	M1QDU0_METMZ
PPK67	G08	A0A1J5QR14_9ZZZZ
PPK75	G09	A0A3B0WJP4_9ZZZZ
PPK83	G10	A0A403VGV9_9ZZZZ
mCherry	G11	X5DSL3
mCherry	G12	X5DSL3
PPK8	H01	A0A3R2YN22_9ZZZZ
PPK16	H02	A0A1I7N6I1_9BACT
PPK25	H03	A0A2A9HF20_9CHLR
PPK33	H04	A0A0E3P3I2_9EURY
PPK42	H05	A0A1V4TV38_9EURY
PPK51	H06	A0A2I0NXE8_9EURY
PPK60	H07	EHM53_03255
PPK68	H08	A0A1J5S232_9ZZZZ
PPK76	H09	A0A3R1QDU9_9ZZZZ

PPK84	H10	A0A405WRT5_9ZZZZ
mCherry	H12	X5DSL3

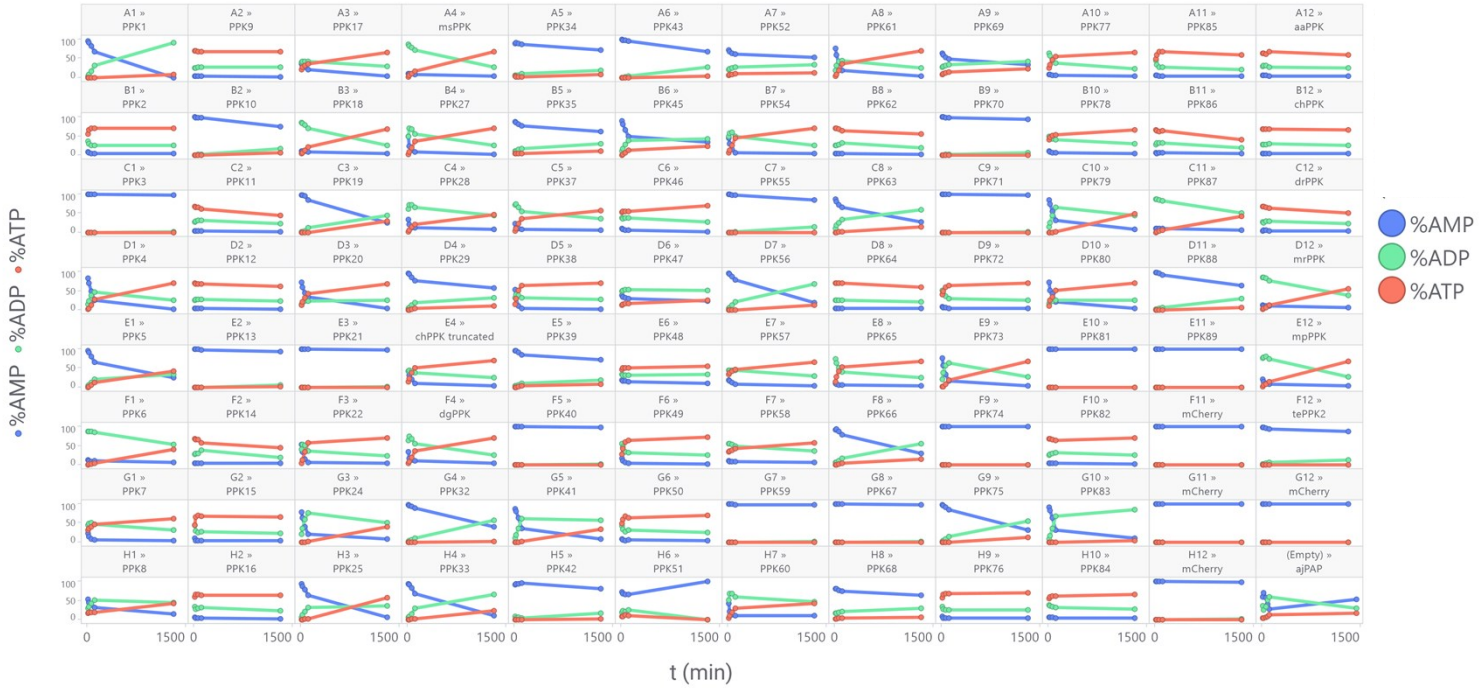

Figure S10.Time course for AMP phosphorylation by putative and characterised PPK2-III. Each panel corresponds to a specific PPK2 (plate location and assigned protein name indicated). The bottom right panel includes a purified protein experiment carried out with purified ajPAP (0.75 mg/mL) for comparison purposes. A negative control obtained from cells overexpressing mCherry was also included.

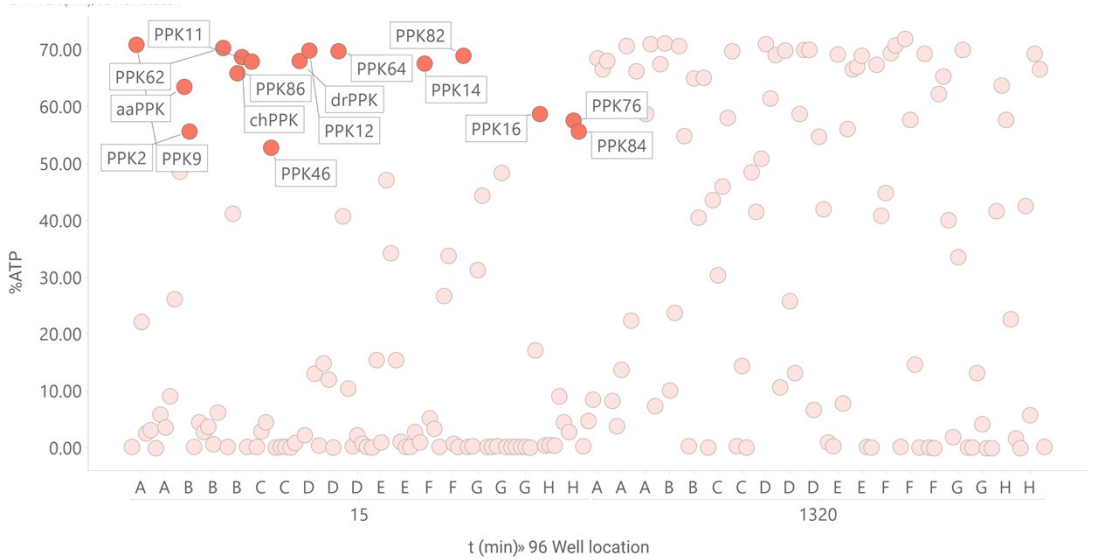

Figure S11. AMP conversion to ATP by putative and characterised PPK2-III. Data at 15 and 1320 minutes are shown. Enzymes showing higher than 50% conversion after 15 minutes are highlighted and labelled.

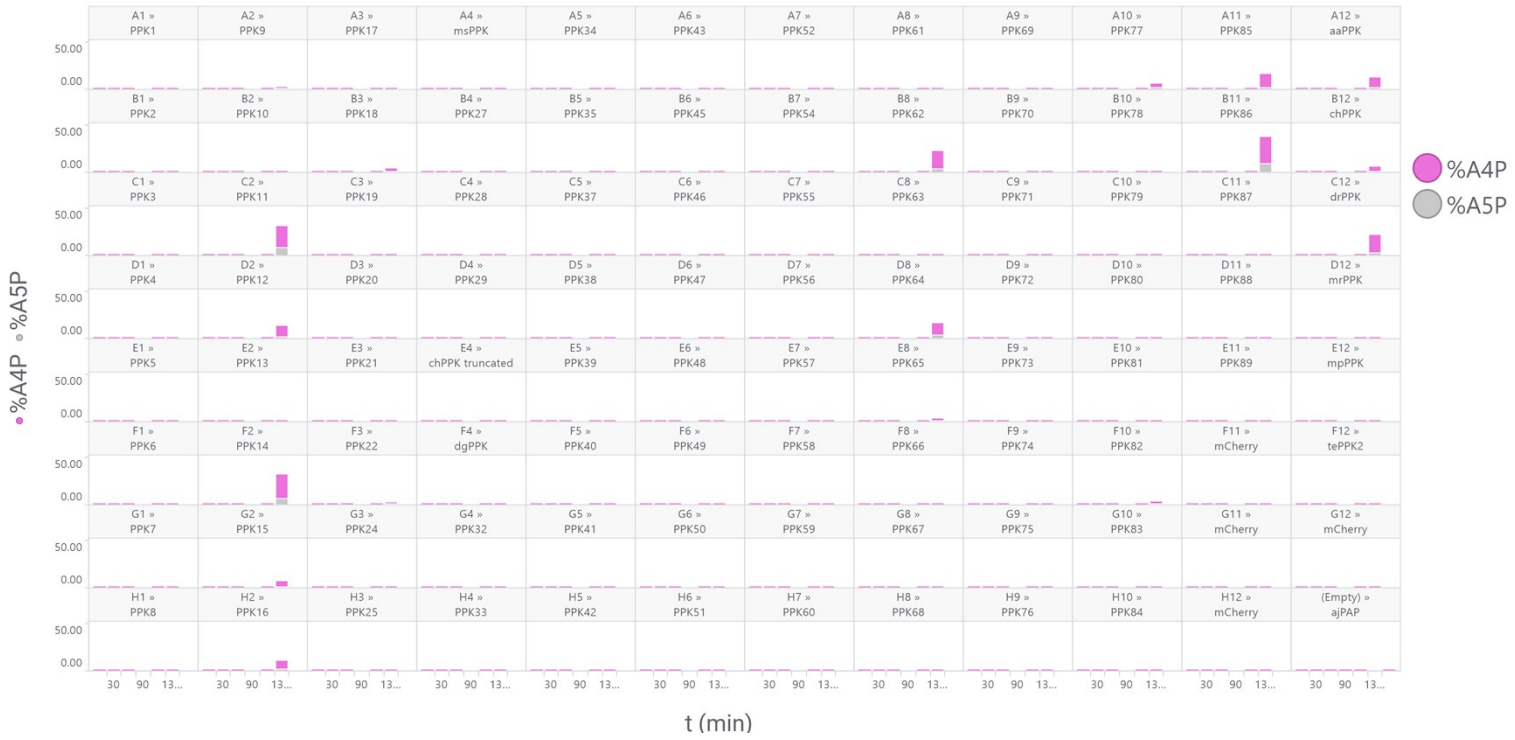

Figure S12. Unknown peaks detected after prolonged incubation times with purified putative PPK2-III. Peaks corresponding to unknown species were tentatively assigned to adenosine 5'-tetraphosphate (A4P) and adenosine 5'-pentaphosphate (A5P) based on retention times and previous literature.[5]

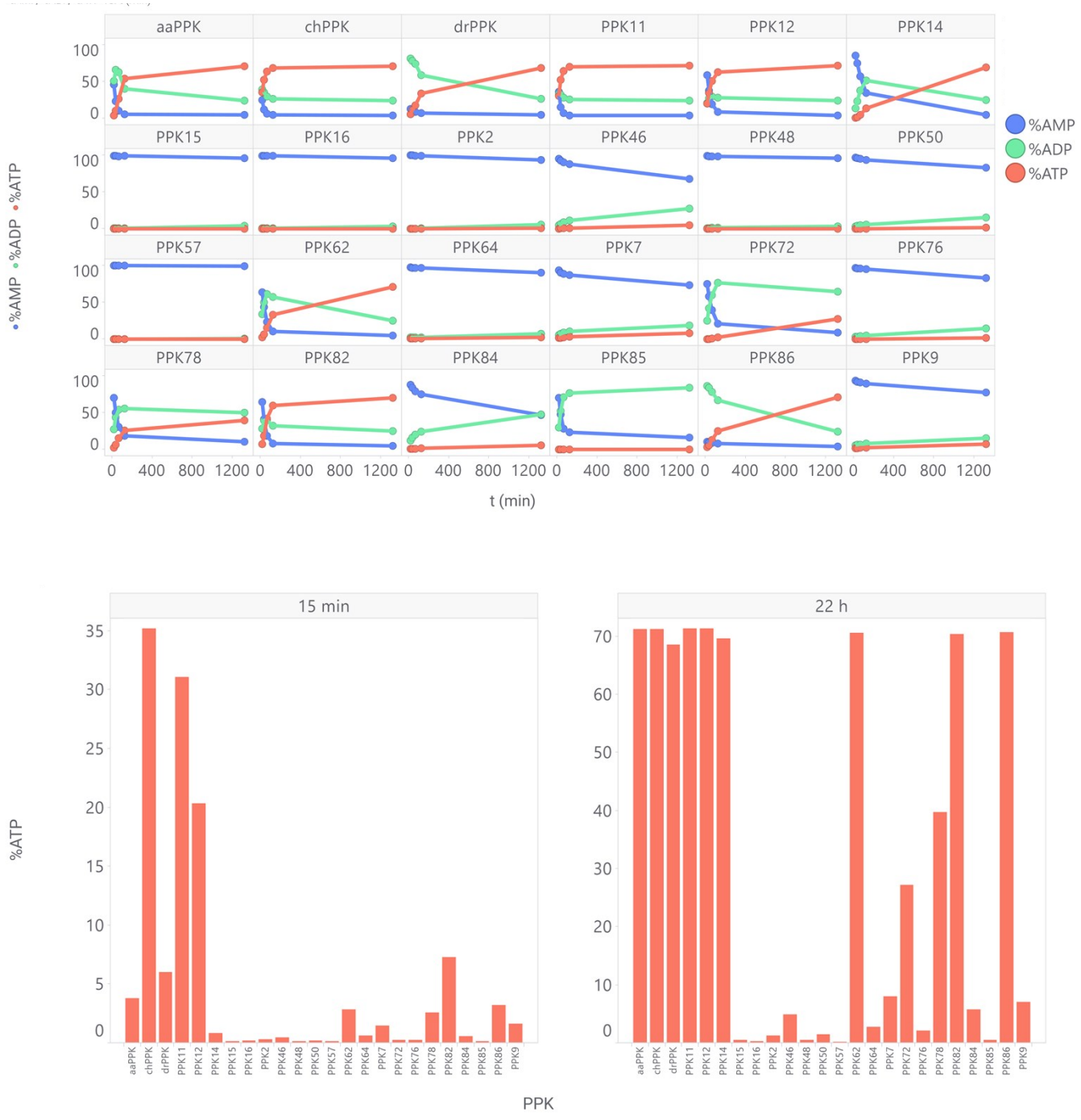

Figure S13. Time course for AMP phosphorylation by 24 putative and characterised PPK2-III selected after initial experiments (top panel) and AMP conversion to ATP data at 15 minutes and 22 h reaction times (bottom panel). Top: each panel corresponds to a specific PPK2 (assigned protein name indicated). Reactions (0.4 mL biotransformation volume, 30 °C, 800 rpm, 22 h) were carried out in 50 mM Tris buffer pH 7.5 and contained 30 µL of 100-fold diluted purified protein compared to Figure S10, 5 mM AMP, 50 mM polyP (based on phosphate units) and 50 mM $MgCl_2$.

Table S5. Novel and characterised PPK2-III enzymes selected after purified protein screening reactions. UniProt IDs, short name and sequence origin are indicated.

UniProt	Short Name	Origin
A1R8G0	aaPPK	*Paenarthrobacter aurescens*
Q11YW6	chPPK	*Cytophaga hutchinsonii*
A0A0F9VM92_9ZZZZ	PPK62	Marine Sediment Metagenome
Q9RY20	drPPK	*Deinococcus radiodurans*
A0A2G6QQJ7_9SPIO	PPK11	*Treponema* sp.
A0A3D5XRJ5_9FIRM	PPK12	*Erysipelotrichaceae* bacterium
A0A3R5FBT2_9ZZZZ	PPK82	Compost Metagenome

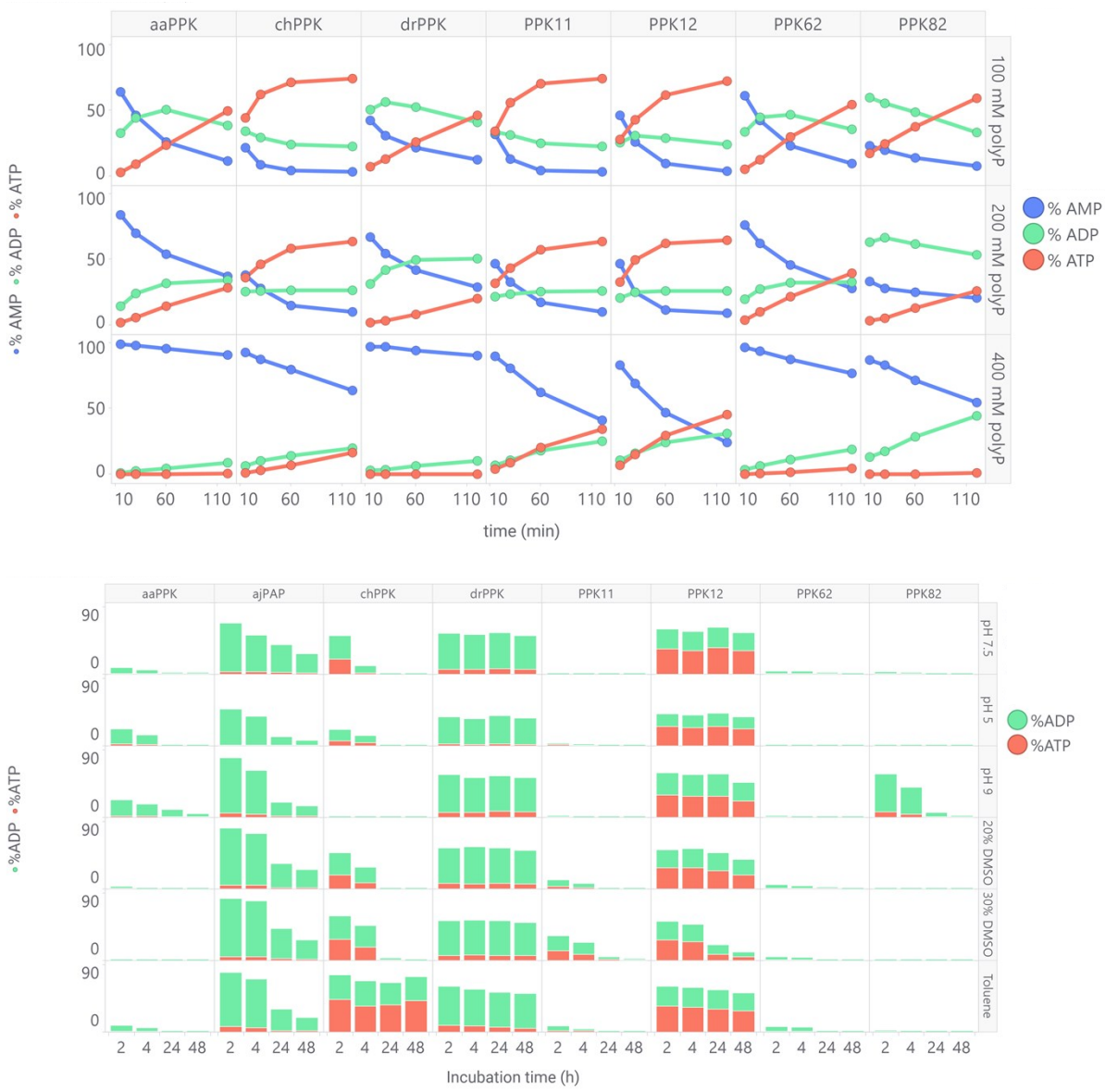

Figure S14. Reaction time course for AMP phosphorylation with selected PPK2 enzymes at different polyP loadings (top figure) and AMP conversion to ADP/ATP after 15 minutes reactions with enzymes pre-incubated for different amount of time (indicated in the x-axis) under different conditions (bottom figure). Reactions (0.4 mL biotransformation volume, 30 °C, 800 rpm) were carried out in 50 mM Tris buffer pH 7.5 and contained 50 µg/mL lyophilised cell free extracts, 5 mM AMP, 50 mM $MgCl_2$ and 100-400 mM polyP (100 mM polyP for stability assays, bottom figure).

	chPPK	tePPK	PPK12	msPPK	mrPPK	aaPPK	dgPPK	drPPK
chPPK	100	49.16	41.61	43.77	42.8	38.85	44.11	41.44
tePPK	49.16	100	41.61	42.26	44.32	39.71	42.59	40.68
PPK12	41.61	41.61	100	40.75	43.56	35.58	40.84	41.6
msPPK	43.77	42.26	40.75	100	57.58	47.74	47.51	41.76
mrPPK	42.8	44.32	43.56	57.58	100	41.89	49.62	45.83
aaPPK	38.85	39.71	35.58	47.74	41.89	100	48.87	48.12
dgPPK	44.11	42.59	40.84	47.51	49.62	48.87	100	65.04
drPPK	41.44	40.68	41.6	41.76	45.83	48.12	65.04	100

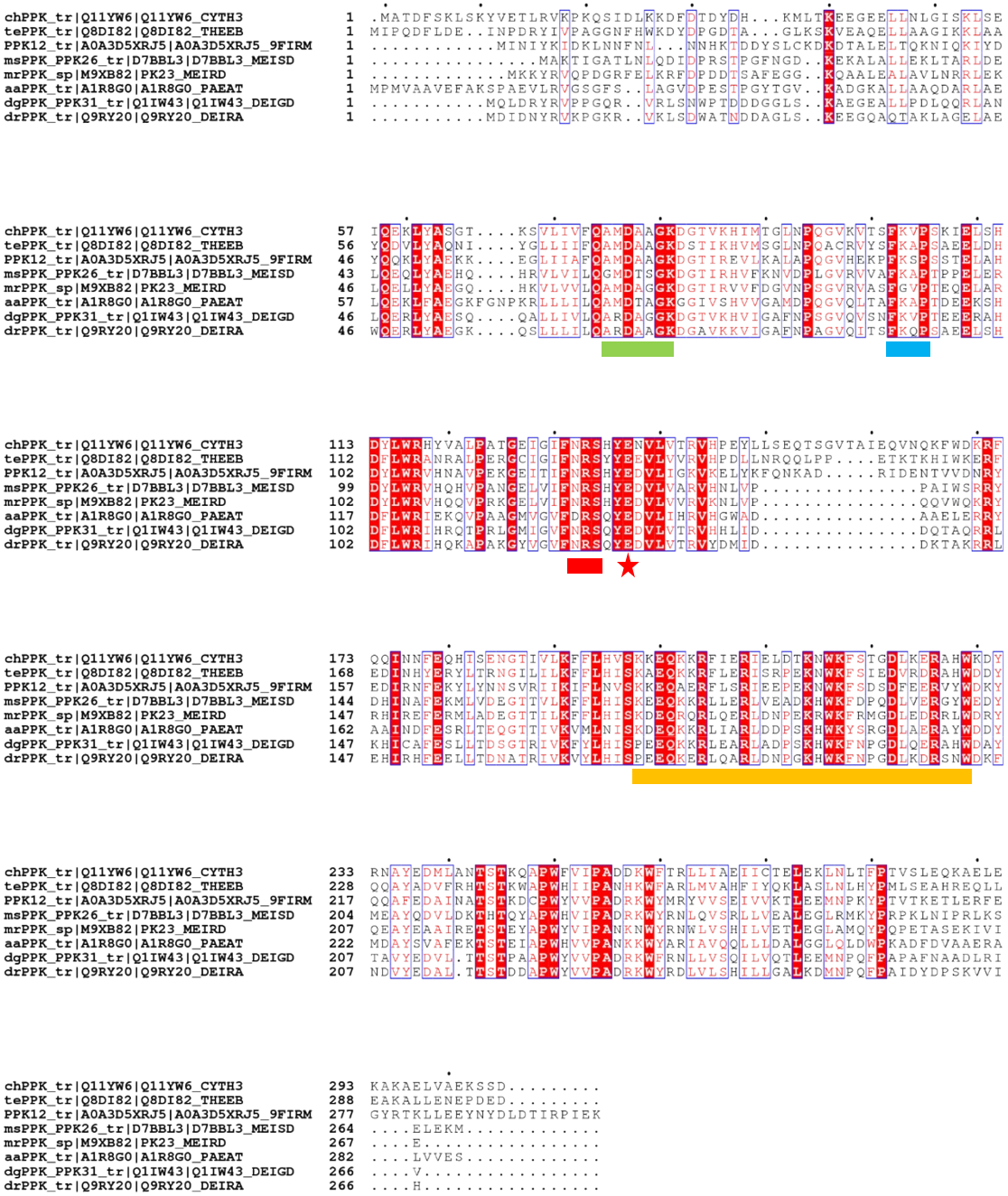

Figure S15. Percentage identity matrix (top) and sequence alignment between PPK12 and previously characterised PPK2-III enzymes. The signature residue for PPK2-III (Glu126^{PPK12}) is indicated by the red star. Walker A motif, base binding loop, Walker B motif and the lid domain are indicated by green, blue, red and orange lines, respectively.[4, 6] Strict identity is indicated by white characters in red boxes, similarity in a group of sequences (≥70%) is indicated by red characters in blue frames (similarity is assigned based on physico-chemical properties).Sequence alignments were produced with MAFFT[7] and visualised with ESPrint.[8]

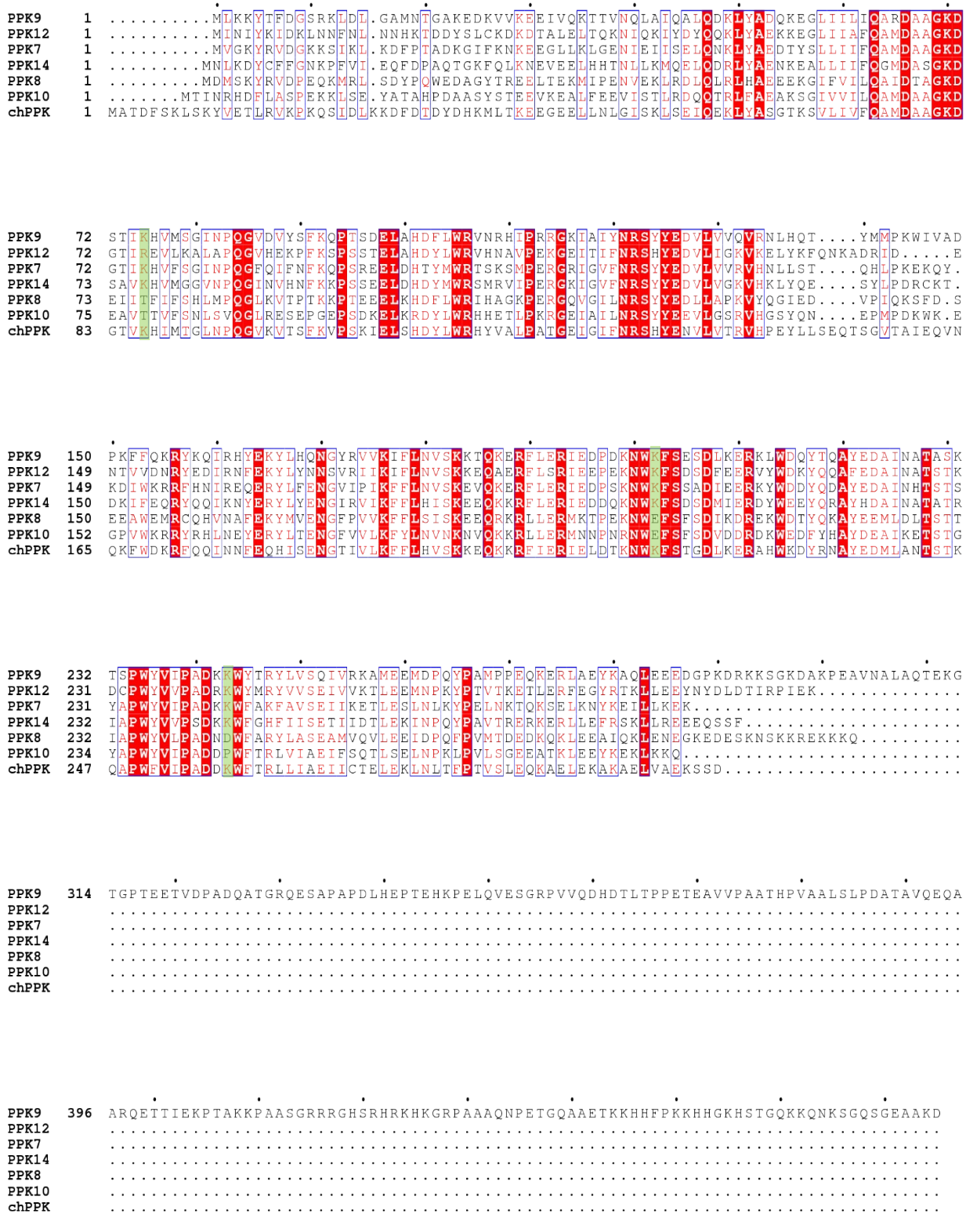

Figure S16. Sequence alignment between sequences falling in the clade comprising PPK12 and chPPK. Strict identity is indicated by white characters in red boxes, similarity in a group of sequences (≥70%) is indicated by red characters in blue frames (similarity is assigned based on physico-chemical properties). Residues discussed in the main text are shaded in green.

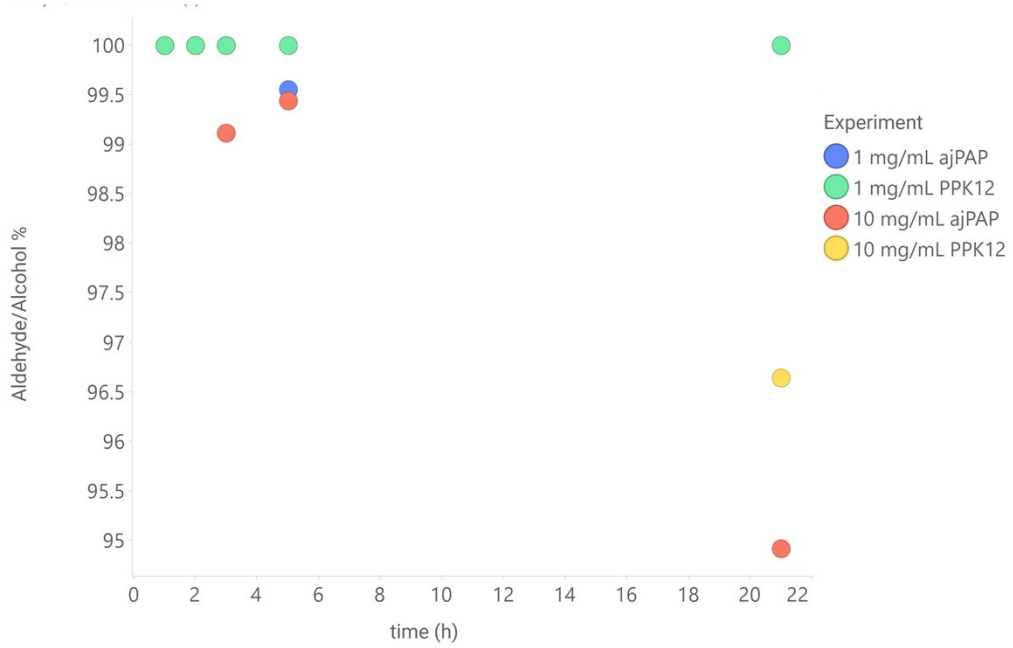

Figure S17. Aldehyde overreduction comparison for ajPAP and PPK12 powered carboxylate reduction.

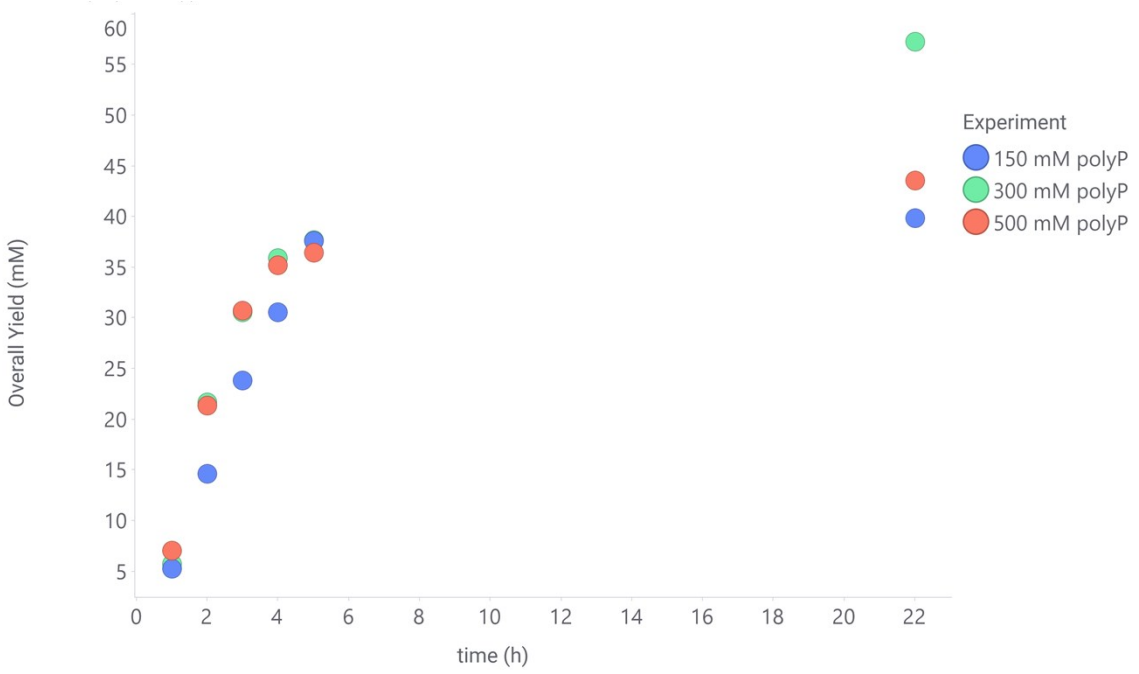

Figure S18. Time course for the CAR-catalysed 4-methoxybenzoic acid reduction with dual cofactor recycling on a 30 mL scale at different polyP loadings. Reactions (30 mL initial volume, 30°C, 400 rpm, 21 h) contained 10 mg/mL CAR33, 1 mg/mL PPK12, 1 mg/mL GDH-CDX-901, 75 mM 4-methoxybenzoic acid, 7% v/v DMSO, different polyP concentrations (see legend), 40 mM MgCl2, 1 mM $NADP^+$, 300 mM D-Glucose, 1 mM ATP, toluene (50% v/v). 4-methoxybenzoic acid and polyP were added over 5 h. Final concentrations take base additions into account. The overall yield is the sum of aldehyde and alcohol yields.

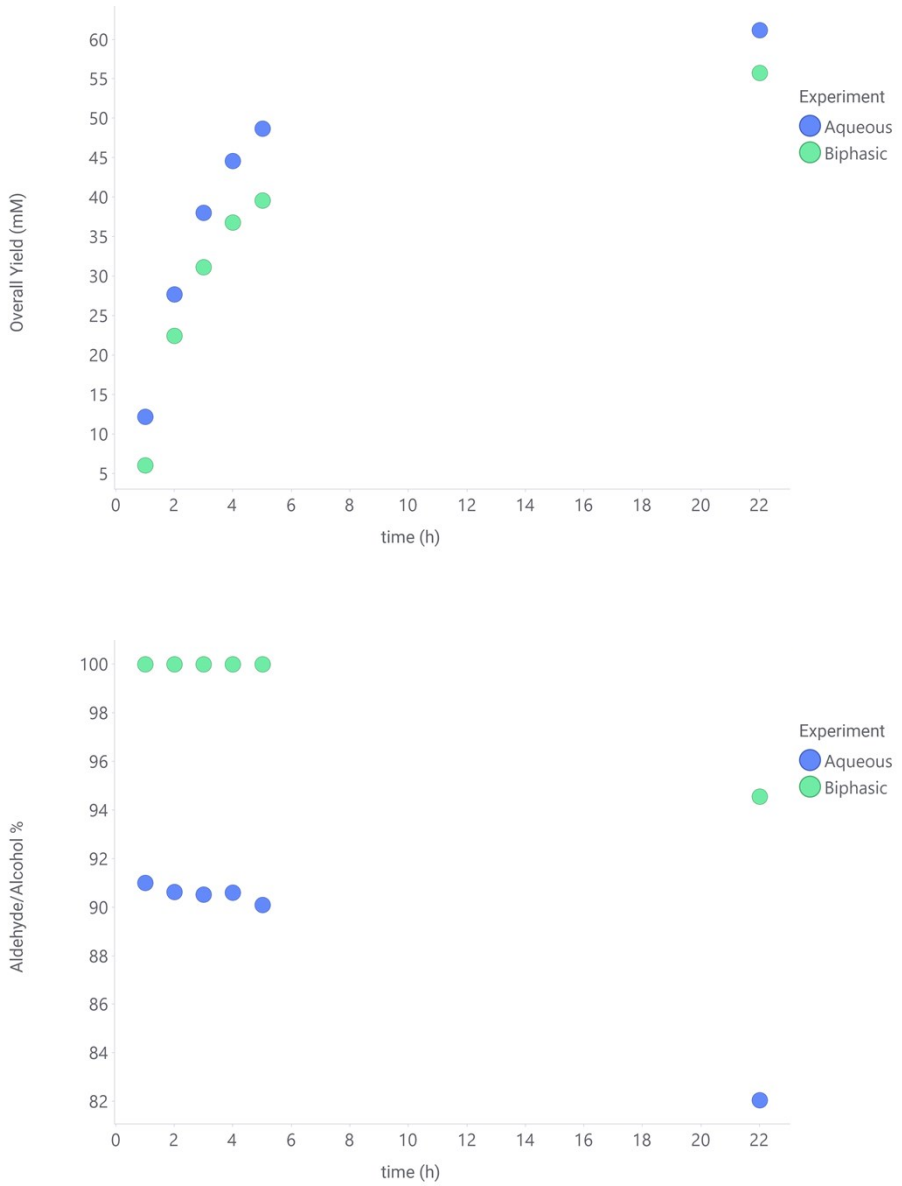

Figure S19. Time course for the CAR-catalysed 4-methoxybenzoic acid reduction with dual cofactor recycling on a 30 mL scale (top) and product composition over time (bottom). Reactions (30 mL initial volume, 30°C, 400 rpm, 21 h) contained 10 mg/mL CAR33, 1 mg/mL PPK12, 1 mg/mL GDH-CDX-901, 75 mM 4-methoxybenzoic acid, 7% v/v DMSO, 300 mM polyP, 40 mM MgCl2, 1 mM $NADP^+$, 300 mM D-Glucose, and 1 mM ATP. Toluene (50% v/v) was included for biphasic reactions. 4-methoxybenzoic acid and polyP were added over 5 h. Final concentrations take base additions into account. The overall yield is the sum of aldehyde and alcohol yields.

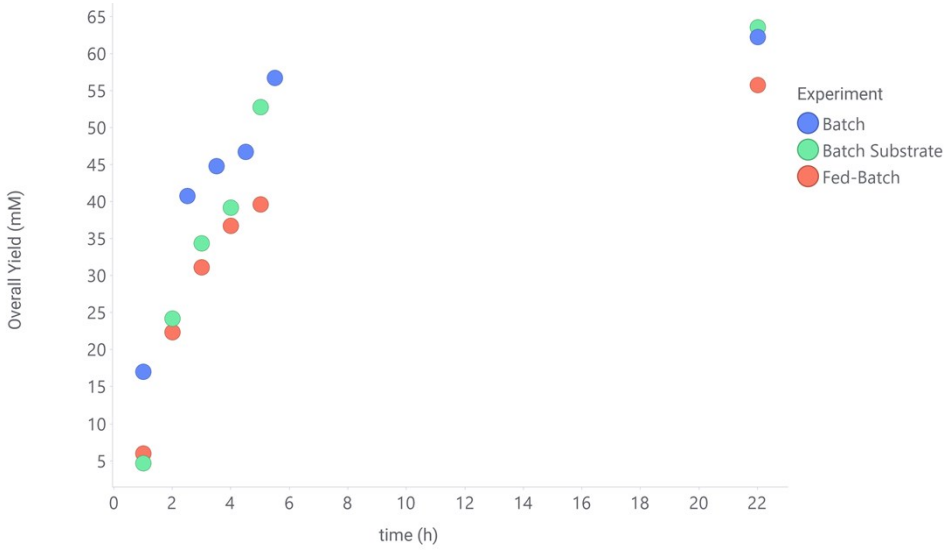

Figure S20. Time course for the CAR-catalysed 4-methoxybenzoic acid reduction with dual cofactor recycling on a 30 mL scale with different feeding strategies. Reactions (30 mL initial volume, 30°C, 400 rpm, 21 h) contained 10 mg/mL CAR33, 1 mg/mL PPK12, 1 mg/mL GDH-CDX-901, 75 mM 4-methoxybenzoic acid, 7% v/v DMSO, 300 mM polyP, 40 mM MgCl2, 1 mM $NADP^+$, 300 mM D-Glucose, 1 mM ATP, toluene (50% v/v). 4-methoxybenzoic acid and polyP were added over 5 h for "Fed-Batch" reactions, whereas polyP only was added over 5 h for "Batch Substrate" reactions. Both polyP and 4-methoxybenzoic acid were added as a single dose at the beginning of the reaction for "Batch" reactions. Final concentrations take base additions into account. The overall yield is the sum of aldehyde and alcohol yields.

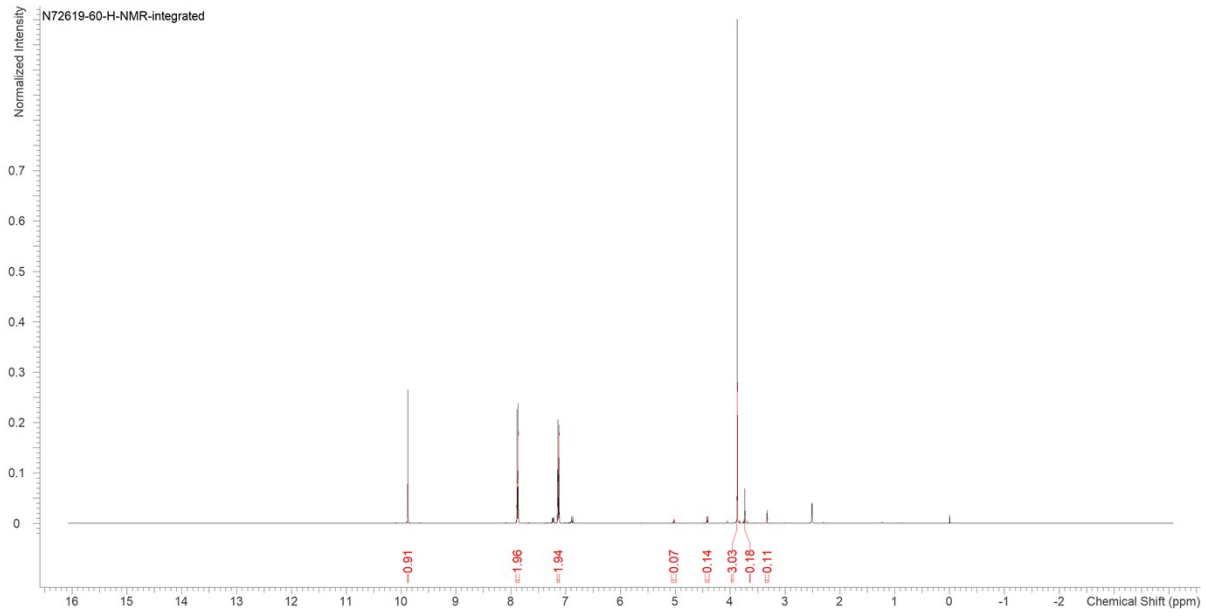

Figure S21. ^1H NMR (400 MHz, DMSO-d6) for isolated products after CAR-catalysed 4-methoxybenzoic acid reduction to 4-methoxybenzaldehyde on multigram-scale with ATP recycling by PPK12. Product characterisation was supported by spectra of known compounds.[9, 10]

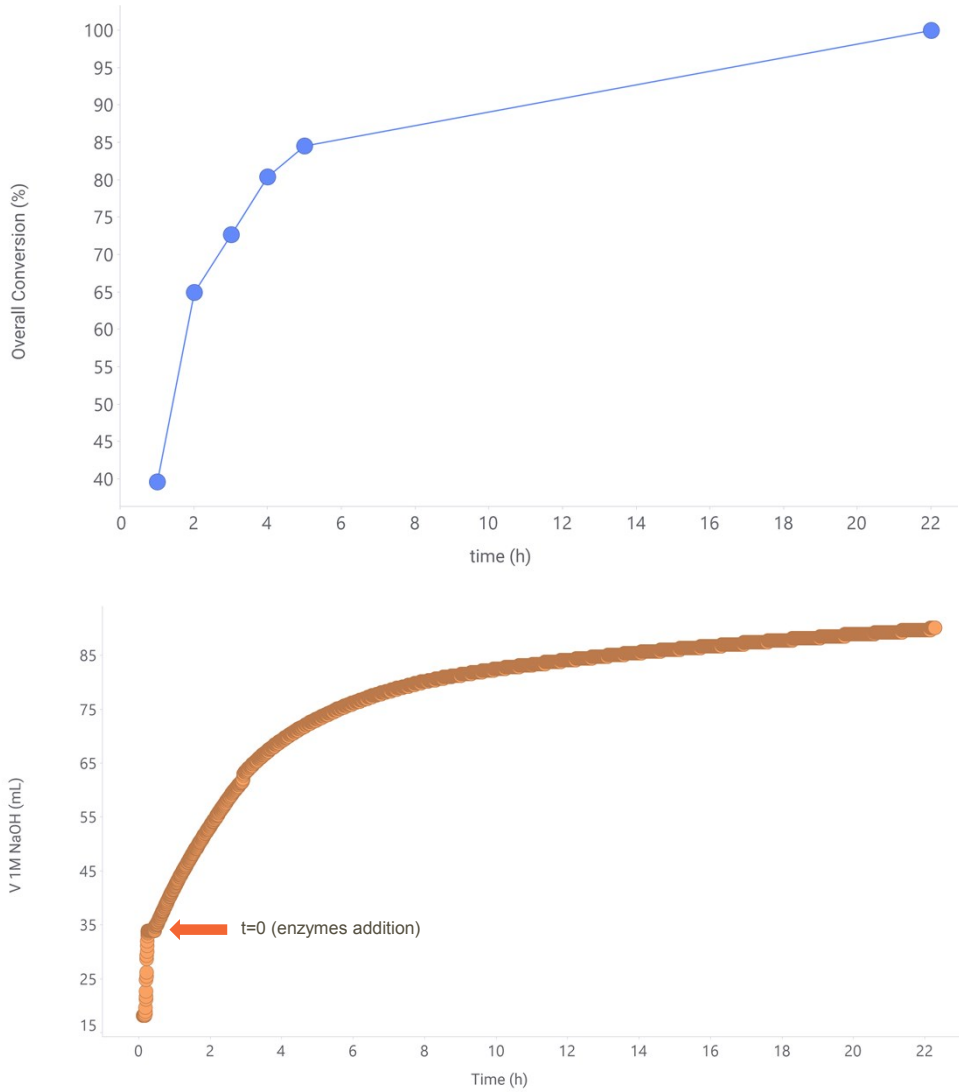

Figure S22. Time course for the CAR-catalysed 4-methoxybenzoic acid reduction on a 0.4 L scale with PPK12 ATP regeneration (top) and pH regulation profile (bottom). Reactions (0.4 L final volume, 30°C, 400 rpm, 22 h) contained 7.6 mg/mL CAR33, 0.76 mg/mL PPK12, 0.76 mg/mL GDH-CDX-901, 37.5 mM $MgCl_2$, 0.75 mM NADP+, 225 mM D-Glucose, 0.75 mM ATP, 235 mM polyP, 65 mM 4-methoxybenzoic acid, 4.3% v/v DMSO and toluene (150 mL, 37.5% v/v). Reported concentrations refer to the final reaction volume (400 mL final volume). The overall conversion is the sum of aldehyde and alcohol conversion values.

Material and Methods

General

NADP disodium salt was obtained from Bontac Bio-Engineering. Other compounds were generally obtained from Sigma-Aldrich and were of the highest purity. Sodium hexametaphosphate (crystalline, +200 mesh, 96%, Sigma 305553) was employed as polyphosphate source. Polyphosphate concentrations correspond to the concentration of phosphate units (MW 101.96). NMR spectra were collected using a Bruker Avance 400 Ultrashield instrument and analyzed using ACD Spectrus Processor. UPLC analysis was carried out on an Agilent 1200 Infinity System equipped with G4204A Quaternary Pump, G1316C Column Compartment, G4226A Autosampler and G1314E variable wavelength detector. Competent cells, enzymes for molecular biology and Gibco™ Terrific Broth were purchased from ThermoFisher Scientific. Overnight Express™ Instant TB Medium was purchased from Sigma-Aldrich. Unless stated otherwise, protein production in shake flasks was carried out in Kuhner shakers whereas small-scale biotransformations were carried out using Infors Multitron shakers. Plates were either sealed with gas permeable membranes for protein production growth or with aluminium-based seals for storage/heat shock (Agilent PlateLoc sealer). Whenever possible, transfer and aliquoting operations were carried out using a Beckman Biomek FX liquid handling robot.

Experiment IDs are indicated for matching with raw data that will be submitted to 4TU.ResearchData after manuscript revision and acceptance.

UPLC Methods

The following ion-pairing method was employed for phosphorylated nucleotide analysis:

AXP_dev_6.M

Time	Flow Rate (mL/min)	%A	%B
0.00	1	95	5
1.3	1	59	41
1.40	1	10	90
2.00	1	10	90
2.10	1	95	5
3	1	95	5

Zorbax RRHD Eclipse Plus C18 column (50mm x 2.1mm i.d.1.8µm packing diameter)

60 °C

1 µL injection vol.

260 nm

Mobile phase A = 50 mM hexylammonium acetate (HAA) in water

Mobile phase B = 65:35 IPA/MeCN

For the preparation of mobile phase A, 1M HAA was prepared as follows:

- 5.72 mL Acetic Acid (Sigma, A6283-100mL)
- 13.21 mL Hexylamine (Sigma, 219703-100mL)
- 81.07 mL water (Sigma, 34877-2.5L)

Acetic acid was dissolved in 50 mL H_2O with vigorous stirring. Hexylamine was then added and the solution stirred for 10 minutes. The remaining water was then added. pH was adjusted to 7 with hexylamine or acetic acid. The following solution was then added to 1900 mL water to generate mobile phase A.

For the analysis of carboxylate reduction reactions, the following method was employed:

CAR_dev_50%_220.M

Time	Flow Rate (mL/min)	%A	%B
0.00	1	90	10
1.50	1	50	50
1.90	1	50	50
2.00	1	90	10
2.50	1	90	10

Acquity UPLC CSH C18 column (50mm x 2.1mm i.d.1.7µm packing diameter)

40 °C

0.3 µL injection vol.

220 nm

A = 0.1% (v/v) TFA in water

B = 0.1% (v/v) TFA in acetonitrile

Conversion values for ATP synthesis reactions are area to area based, whereas conversion data for CAR reactions are corrected using relative response factors. Yields for these reactions have been calculated based on the conversion value and the concentration of starting material.

Construction of Panels

Sequences listed in Table S1 were identified after literature review and BLAST searches. For database mining, ajPAP, rmPPK and mrPPK sequences were used to search for PPK2 enzymes from extremophiles. The selected PPK sequences were assigned to a specific PPK2 class based on the query sequence employed.

PPK2-III panel was constructed starting from PPK2 related proteins (Interpro IPR022488, 28000 proteins). Sequences were filtered by length (150-600 amino acids) using Galaxy web plaform (https://usegalaxy.org/). The obtained sequences were clustered at 90% sequence identity using CD-HIT[11] (11779 clusters obtained). Sequences were retrieved using UNIPROT, exported in Excel and fragments/poorly annotated sequences filtered out. Sequences for characterised PPK2 enzymes[4] were added for later clustering verification of PPK2 enzymes belonging to different classes. A multiple sequence alignment was performed with MAFFT (standard parameters for large data-sets).[7] A phylogenetic tree was constructed (UPGMA) and visualised in Archaeopteryx.[12] Sequences clustering with reported PPK2-III enzymes were retrieved using UNIPROT (4004 sequences) and divided based on taxonomy (38 sequences from metagenomes, 34 from Archaea and 3930 from Bacteria). Bacterial sequences were further clustered using h-CD-HIT (3 steps, 80, 70 and 60% sequence identity) and a phylogenetic tree constructed with the obtained 686 sequences after sequence alignment. Sequences clustered in 4 major clades. Each individual cluster was manually inspected with a major focus on extremophiles and finally 27 bacterial sequences selected for collation with archaeal and metagenomic sequences.

Molecular Biology and Gene Acquisition

DNA sequences encoding proteins listed in Table S1 were codon optimised for *E. coli* expression using the GeneArt codon optimisation tool and obtained as gblocks® from IDT. DNA sequences were cloned in pET28a (Novagen) between NdeI and XhoI restriction sites using NEBuilder® HiFi DNA Assembly Cloning Kit according to manufacturer instructions. Sequences encoding the N-terminal 6xHis tag proteins were verified by Sanger sequencing (GeneWiz).

The sequences encoding the selected proteins for the PPK2-III panel were codon optimised for *E. coli* expression, synthesised and cloned by Twist Bioscience. The cloning strategy and receiving plasmid was the same as above. Previously reported PPK2-III included in the panel were obtained from IDT and cloned as indicated above.

Shake Flask Protein Expression for Small Scale Experiments

BL21(DE3) cells (10 or 50 µL aliquots) were transformed with 50-100 ng plasmid (1 µL) encoding the desired protein. The transformation protocol followed manufacturer instruction. Outgrowth was carried out after addition of 300 or 900 µL SOC for 10 or 50 µL aliquots of competent cells respectively. 300 µL of the transformation volume were plated in agar plates containing 50 µg/mL kanamycin and plates incubated overnight at 37 ºC. The following day, 50 mL LB medium containing 50 µg/mL kanamycin were added to 250 mL Erlenmeyer flasks and inoculated with a single colony transformed with a pET28a plasmid encoding the protein of interest. Flask were incubated overnight at 37 °C in a Kuhner shaker (200 rpm orbital shaking). Glycerol stocks were prepared by adding 0.5 mL of the produced seed to 0.5 mL of 50% glycerol in cryogenic vials (Corning). Expression cultures were prepared with 1% inoculation of seeds produced overnight (10 mL) into a 2 L Ultra Yield™ flask (Thomson) containing 1 L medium. Overnight Express medium was employed for the production of ajPAP, CAR33, PPK62 and proteins listed in Table S1, whereas Terrific broth was employed for other enzymes tested in this work. In the former case the medium was supplemented with 1% glycerol. Flasks were sealed with AirOTop™ seals and incubated at 37 ºC with 200 rpm shaking. After 2.5-3.5 h (OD 0.6-0.8), flasks containing Overnight Express were moved at 20 ºC and protein production continued for 20 h with 200 rpm shaking. IPTG (final 0.5 mM) was added to flasks containing Terrific broth after 2.5 h (OD 0.6-1) before moving them at 20 ºC.

Cells were spun down using an Eppendorf 5810R centrifuge at 4000 rpm for 30 minutes and stored at -80 ºC until further use.

Fermenter-scale Enzyme Production

ajPAP and CAR33 employed for scale-up experiments were produced in a 50 L fermenter (35 L working volume, 30 L/min air, 200rpm, 0.2bar).

A sterile Corning Erlenmeyer flask (500 mL) containing Luria Broth medium (100 mL, LB) and kanamycin (50 µg/mL final concentration) was inoculated with cells picked from an agar plate containing BL21(DE3) cells transformed with either ajPAP or CAR33 plasmid. The flask was incubated for 6 h at 30 °C in a Kuhner shaker with 200 rpm shaking (primary seed). The secondary seed was prepared with a 1% inoculation from the primary seed into 9 sterile glass Erlenmeyer flasks (500 mL) containing Luria Broth medium (100 mL, LB), kanamycin (50 µg/mL final concentration) and glucose (1% final concentration). The mixture was incubated at 30 °C, 200 rpm in a Kuhner shaker overnight. The production stage was initiated with a 2% inoculation using secondary seeds (700 mL) added into a 50 L fermenter containing 35 L of Terrific Broth (prepared with 12 g Tryptone, 24 g Yeast Extract, 2.2 g KH_2PO_4 and 9.4 g of K_2HPO_4 per litre of medium, 1% glycerol, 5 mL antifoam (Dow Corning 1520) and kanamycin (50 µg/mL final concentration). The culture was incubated at 200 rpm, 37 ºC for ~150 min until OD 600 nm reached 2.0. Induction was accomplished with 0.5 mM IPTG. Protein production was continued for 23 h at 22.5 ºC.

Cells were harvested by centrifugation (6,427 rcf, 4 ºC, Sorvall RC 12BP). Cell pellets were kept at -80 ºC until further use.

Production of PPK2-III Panel Enzymes

10 µL sterile milliQ water was added to the plate containing DNA synthesised by Twist Bioscience and the content mixed. BL21(DE3) and TOP10 cells (10 µL) were transferred into pre-chilled skirted Armadillo PCR Plates (ThermoFisher Scientific). 1 µL of plasmid DNA was added to each well containing competent cells. Cells were heat shocked by placing the plates in a thermal cycler at 42 ºC for 30 seconds then transferred directly to an ice bath for 2 minutes. LB medium was aliquoted into a 50 mL sterile reservoir and 100 µL added to the plates containing cells. The content of each plate was then transferred into Corning Costar 96-deep well plates pre-aliquoted with 400 uL LB and plates incubated at 37 ˚C at 300 rpm 85% humidity for 1 hour. After outgrowth, 500 µL LB containing 100 µg/mL kanamycin were added to the Costar plates containing cells which were subsequently incubated at 37 °C, 300 rpm, 85% humidity for 3.5 hours. Afterwards, 20 µL were transferred from the plate containing transformed cells to Costar 96-deep well plates pre-aliquoted with 480 uL LB containing 50 µg/mL kanamycin (two copies of such a plate were prepared from plates containing BL21 cells). The cultures were incubated overnight (18-19 hours) at 37 °C, 300 rpm, 85% humidity. The following day, glycerol stocks were prepared by adding 50 µL of seed cultures to Nunc plates (ThermoFisher Scientific) pre-aliquoted with 50% glycerol. Glycerol stocks were stored at -80 °C until further use. Expression cultures were prepared by adding 20 µL seed to to Costar 96-deep well plates pre-aliquoted with 380 µL Terrific broth containing 50 µg/mL kanamycin. Cells were grown at 37 °C, 300 rpm, 85% humidity. An expression plate was dedicated to OD measurements (100 µL culture). A total of 13 plates were grown. After 1h (OD 0.6-0.8) protein production was induced by adding IPTG (0.5 mM final concentration) and incubation continued at 20 °C, 300 rpm, 85% humidity, 20 h.

Cells were spun down using an Eppendorf 5810R centrifuge at 4000 rpm for 20 minutes and stored at -80 ºC until further use.

Cell Lysis and Lyophilised Powders Preparation

Cells were weighed out (typically 20-30 g) and 4 volumes of 50 mM KPi buffer pH 7.5 (CAR33 and PPKs in Table S1) or 50 mM Tris pH 7.5 were added. Bottle containing cells were shaken at 120 rpm, 20 °C for ~20 minutes to aid resuspension. Resuspended cells were filtered through a single layer of muslin cloth into Schott bottles and lysed using a microfluidizer (Microfluidics, 1500-2000 bar operating pressure). Cell lysates were transferred into 50 mL Falcon tubes and centrifuged at 10000 rpm, 4°C, 1 h (Eppendorf 5810R). 1 mL supernatants were withdrawn for subsequent capillary electrophoresis analysis, whereas the remainder was either transferred into square petri dishes or in Lyoprotect Aluminium Trays (Teclen) equipped with Lyoprotect Membrane for volume M tray"(Teclen) and Silicone flat gasket (Teclen). Lysates were placed at -80 °C for 1-2 h. Lyophilisation was carried employing an AdVantage Pro with Intellitronics, SP Scientific (48 h program). The resulting lyophilised powders were transferred in pots and stored at -20 °C. No apparent loss in activity for ajPAP and CAR33 was observed within 8 months storage. Powders containing selected PPKs from the PPK2-III panel were stored for 2 months without apparent loss in activity.

Capillary gel electrophoresis was carried out as per manufacturer instruction (PerkinElmer LabChip GXII Touch HT Protein Characterization System) after diluting clarified cell free extracts 2-10-fold.

High-Throughput Protein Purification

Plates were removed from the -80 °C freezer and allowed to thaw. 200 µL lysis buffer (50 mM Tris pH 7.5, 300 mM NaCl, 25 mM imidazole, 10% glycerol, 0.5 mg/mL, 1 mg/mL lysozyme and 0.1 µL/mL of benzonase E1014-25 kU) were added to the plates before incubation for 2 h, 1000 rpm, room temperature (18-22 °C). Lysates were clarified by centrifugation (Eppendorf 5810R, 4000 rpm, 10 min, 4 °C) and supernatants transferred into KingFisher™ Flex 96 deep-well plates. High-throughput protein purification was carried out as per manufacturer instructions (ThermoFisher KingFisher Flex System) using a method adapted from "Purify His tag Magnetic Agarose Ni-NTA Direct"

(https://www.thermofisher.com/uk/en/home/life-science/dna-rna-purification-analysis/automated-purification-extraction/automated-protocols-software.html?open=protein) for binding across two binding plates. The method entails different transfers of Ni-NTA magnetic agarose beads (200 µL for this experiment) in equilibration (50 mM Tris pH 7.5, 300 mM NaCl, 25 mM imidazole, 10% glycerol) and wash buffers (50 mM Tris pH 7.5, 300 mM NaCl, 50 mM imidazole, 10% glycerol) before elution (50 mM Tris pH 7.5, 300 mM NaCl, 300 mM imidazole, 10% glycerol). Once the purification protocol had finished the elution plates were transferred in ice and their content transferred into a skirted Armadillo PCR Plates (ThermoFisher Scientific). Capillary electrophoresis was carried out as above without further dilution of protein samples.

Small Scale Assay of the Initial PPK2 Panel – Experiments N17966-6 & N71966-9

Reactions (1 mL biotransformation volume, 50 mM KPi buffer pH 7.5) were carried out in Corning Costar 96-deep well plates and contained 0.1 mg/mL lyophilised cell free extract, 50 mM polyP, 100 mM $MgCl_2$ and 1 mM AMP. Negative controls were represented by lyophilised cell free extract containing penicillin acylase. Plates were incubated at 30 °C, 800 rpm using an Infors Multitron shaker. At predefined time points, 150 µL of the biotransformation mixture were transferred in a Costar 96-deep well plate pre-aliquoted with 150 µL MeCN. Quench plates were shaken (1000 rpm, 5 minutes, room temperature), spun down (Eppendorf 5810R centrifuge, 4000 rpm, 10 minutes, 4 °C) and the supernatant transferred into Nunc plates (ThermoFisher Scientific) for UPLC analysis (AXP_dev_6.M).

Small Scale AMP Loading Assay with the Initial PPK2 Panel – Experiment N71966-10

1 M AMP and 2 M polyP stocks were prepared in 50 mM KPi buffer with sonication as needed. Reactions (1 mL biotransformation volume, 50 mM KPi buffer pH 7.5) were carried out in Corning Costar 96-deep well plates and contained 1 mg/mL lyophilised cell free extract, 200 mM polyP, 100 mM $MgCl_2$ and different concentrations of AMP (20, 40 and 100 mM). Plates were incubated at 30 °C, 800 rpm using an Infors Multitron shaker. At predefined time points, 150 µL of the biotransformation mixture were transferred in a Costar 96-deep well plate pre-aliquoted with 150 µL MeCN. Quench plates were shaken (1000 rpm, 5 minutes, room temperature), spun down (Eppendorf 5810R centrifuge, 4000 rpm, 10 minutes, 4 °C) and the supernatant transferred into Nunc plates (“receiving plates”). Dilution plates were generated by transferring the content receiving plates into Nunc plates pre-aliquoted with 50 mM KPi buffer so to have ~5 mM total nucleotides in 200 µL. These plates were analysed by UPLC (AXP_dev_6.M).

Small Scale AMP Loading Assay with the Initial PPK2 Panel – Experiment N71966-12

0.5 M AMP and 5 M polyP stocks were prepared in 50 mM KPi buffer with sonication as needed. Reactions (1 mL biotransformation volume, 50 mM KPi buffer pH 7.5) were carried out in Corning Costar 96-deep well plates and contained 10 mg/mL lyophilised cell free extract, 200 mM polyP, 100 mM $MgCl_2$ and 200 mM AMP. Plates were incubated at 30 °C, 800 rpm using an Infors Multitron shaker. At predefined time points, 100 µL of the biotransformation mixture were transferred in a Costar 96-deep well plate pre-aliquoted with 100 µL MeCN (quench plates). Afterwards 5 M polyP was added to the biotransformation plate (18.5, 16.5 and 15 µL for the first, second and third polyP dose, respectively) according to the designed plate map. Work-up, dilution and analysis were carried out as in the previous section.

Small Scale CAR Experiments with Lyophilised Cell-Free Extracts and Aldehyde Reduction Controls – Experiments N72619-3/N72619-4 and N72619-5

Small-scale CAR reactions (500 µL biotransformation volume, 50 mM KPi buffer pH 7.5) were carried out in 96-well plates with glass inserts and contained 50 mg/mL CAR33, 10 mg/mL ajPAP, 1 mg/mL GDH-CDX-901, different concentrations of 4-methoxybenzoic acid and polyP (molar ratio polyP:4-methoxybenzoic acid 5:1), DMSO (8% v/v), 100 mM $MgCl_2$, 1 mM $NADP^+$, 200 mM D-Glucose, 1 mM ATP. A toluene overlay (50% v/v) was added to biphasic reactions. The concentration of each component is calculated over the aqueous layer volume for water soluble compounds and over the entire reaction volume (including toluene) for 4-methoxybenzoic acid. The target substrate/polyP concentration was achieved through batch additions of the substrate along with polyP in 30 minutes intervals. Plates were incubated for 20 h at 30 °C, 800 rpm using an Infors Multitron shaker. After 20 h, reactions were transferred to Corning Costar 96-deep well plates and quenched with 2 volumes of 50% MeOH in IPA. Quench plates were shaken (1000 rpm, 5 minutes, room temperature), spun down (Eppendorf 5810R centrifuge, 4000 rpm, 10 minutes, 4 °C) and the supernatant transferred into Nunc plates (ThermoFisher Scientific) for UPLC analysis (CAR_dev_50%_220.M).

Aldehyde reduction control reactions (200 µL biotransformation volume, 50 mM KPi pH 7.5) were carried out in Corning Costar 96-deep well plates and contained 1 mg/mL GDH-CDX-901, 10 mM 4-methoxy benzaldehyde, 100 mM polyP, 1 mM ATP, 1 mM $NADP^+$, 200 mM D-glucose, 100 mM $MgCl_2$, 8% DMSO and different loadings of either ajPAP or CAR33. Plates were incubated for 20 h at 30 °C, 800 rpm using an Infors Multitron shaker. Reaction quench and analysis were carried out as above.

General Procedure for 30 mL-scale CAR Reactions

Reactions were run on an EasyMax102 system equipped with 50 mL reaction vessels. pH was controlled with a SevenExcellence pH Meter and adjusted to pH 7.5 with 1 M NaOH and an SP-50 dosing unit. Temperature was 30 °C and stirring 400 rpm. The biotransformations contained 10 mg/mL CAR33, 1 or 10 mg/mL ATP recycling enzyme, 1 mg/mL GDH-CDX-901, 75 or 150 mM 4-methoxybenzoic acid, 7% v/v DMSO, different amounts of polyP, 40 mM $MgCl_2$, 1 mM $NADP^+$, 300 mM D-Glucose, 1 mM ATP, toluene (50% v/v). 4-methoxybenzoic acid and polyP were added over 5 h for fed-batch reactions from 2 M or 5.8 M stocks, respectively. Final concentrations take dosing into account (final volume ~35 mL). The concentration of each component is calculated over the entire volume at the end of the reaction. At predefined time intervals, reaction aliquots were withdrawn with a plastic Pasteur pipette whose entire content was transferred into 2 mL glass vials (~0.5 mL biotransformation mixture as judged by graduations on the vial) before quenching with 1 mL of 50% MeOH in IPA. Typically, 5 time points were collected in duplicates. After 5 h, the contents of the vials were transferred in Corning Costar 96-deep well plates which were processed as indicated above. The content of the generated Nunc plates was subsequently diluted 5 or 10-fold in MeOH before analysis for 75 mM or 150 mM substrate loading reactions, respectively.

After 20-22 h, the entire content of the 50 mL vessel was added with 2 volumes of 50% MeOH in IPA, transferred in 50 mL Falcon tubes, spun down (Eppendorf 5810R centrifuge, 4000 rpm, 10 minutes, 4 °C) and the supernatant diluted 5/10 fold in MeOH and analysed.

Experiments IDs (see “General” section) are:

- N72619-50: experiments with 1 and 10 mg/mL PPK12 at 150 mM 4-methoxybenzoic acid loading
- N72619-52: experiment with 1 mg/mL PPK12 at 75 mM 4-methoxybenzoic acid loading
- N72619-53: experiments with 1 and 10 mg/mL ajPAP
- N72619-54: experiments without toluene and PPK12
- N72619-55: experiments at 150 mM polyP loading and PPK12

- N72619-57: experiments with batch 4-methoxybenzoic acid or 500 mM polyP loading both carried out with PPK12
- N72619-58: experiments with batch 4-methoxybenzoic acid and batch polyP with PPK12

Screening of Purified PPK2-III Enzymes

Reactions (0.4 mL biotransformation volume, 50 mM Tris buffer pH 7.5) were carried out in Corning Costar 96-deep well plates and contained 30 µL purified protein (maximum protein concentration was 0.76 mg/mL as determined by capillary electrophoresis), 5 mM AMP, 50 mM polyP (based on phosphate units) and 50 mM $MgCl_2$. At predefined time points, 50 µL biotransformation mixture were added into a Greiner 384-well plate pre-aliquoted with 50 µL MeOH. Quench plates were shaken (1000 rpm, 5 minutes, room temperature), spun down (Eppendorf 5810R centrifuge, 4000 rpm, 10 minutes, 4 °C) and the supernatant transferred into a Greiner 384-well plate for UPLC analysis (AXP_dev_6.M). Reactions were performed at 30 °C, 800 rpm.

In subsequent experiments with diluted purified proteins, selected purified proteins were diluted 100-fold before addition to the reaction mixture.

ATP formation rates were calculated for data obtained within 15/30 minutes from the start of the reaction with enzyme concentrations leading to a linear increase in ATP concentration over time.

Experiments IDs (see “General” section) are:

- N71966-24: experiments with undiluted purified proteins
- N71966-25: experiments with diluted purified proteins
- N71966-16: experiments with purified ajPAP

Screening of Cell-Free Extracts Containing Selected PPK2-III Enzymes

PolyP Tolerance (Experiment N71966-29)

Reactions (0.4 mL biotransformation volume, 50 mM Tris buffer pH 7.5) were carried out in Corning Costar 96-deep well plates and contained 50 µg/mL lyophilised cell-free extract, 5 mM AMP, 100-400 mM polyP (based on phosphate units) and 50 mM $MgCl_2$. At predefined time points, 50 µL biotransformation mixture were added into a Greiner 384-well plate pre-aliquoted with 50 µL MeOH. Reactions were performed at 30 °C, 800 rpm. Quench plates were shaken (1000 rpm, 5 minutes, room temperature), spun down (Eppendorf 5810R centrifuge, 4000 rpm, 10 minutes, 4 °C) and the supernatant transferred into a Greiner 384-well plate for UPLC analysis (AXP_dev_6.M).

Stability (Experiment N71966-30)

Solutions containing 10 mg/mL lyophilised cell-free extract were prepared. 900 µL buffer (50 mM Tris pH 7.5, 0.1 M Tris pH 9 or 0.1 M sodium citrate pH 5) were added to 2 mL glass vials, whereas 700 µL or 600 µL 50 mM Tris pH 7.5 were added for 20 and 30% DMSO incubations respectively. 675 µL of 50 mM Tris pH 7.5 were added to vials for toluene incubations (50% v/v). 200 µL or 300 µL DMSO were added to vials for 20 or 30% DMSO incubations, respectively.100 µL of the initially prepared enzyme solutions were added to each vial and 75 uL were added to vials where toluene incubations were carried out (1 mg/mL final enzyme concentration). Vials were incubated at 30 °C using an Eppendorf ThermoMixer® C equipped with a thermoblock suitable for vials (12 mm vessels thermoblock) and a Thermotop. 100 µL sample were withdrawn after 2, 4, 24 and 48h and transferred into a Corning 96-well round bottom plate. Samples were immediately transferred at -80 °C until reactions were carried out. After 48h, reactions were carried out following conditions indicated for the polyP tolerance assay (100 mM polyP employed). After 15 minutes, reactions were quenched and analysed as indicated above.

Procedure for the Multigram-scale CAR Reaction with ajPAP – Experiment N72619-48

Reactions were run on an EasyMax402 system equipped with a 500 mL reaction vessel. pH was controlled with a SevenExcellence pH Meter and adjusted to pH 7.5 with 1 M NaOH and an SP-50 dosing unit. Temperature was 30 °C and stirring 400 rpm. The biotransformations contained 2 mg/mL CAR33, 2 mg/mL ajPAP, 0.2 mg/mL GDH-CDX-901, 50 mM $MgCl_2$, 1 mM $NADP^+$, 300 mM D-Glucose, 1 mM ATP, 25 mM polyP and toluene (50% v/v), at the beginning of the reaction (250 mL initial volume). The concentration of each component is calculated over the entire biotransformation volume. 37 mmol polyP were added over the first 5 h. 4-methoxybenzoic acid was added over the first 5 h from a 2 M stock in DMSO (final 7% v/v) and its final concentration was 70 mM.

After 5 h from the start of the biotransformation, ajPAP (530 mg), CAR33 (530 mg), GDH (53 mg), $MgCl_2$ (842 mg), $NADP^+$ (209 mg), D-Glucose (4.2 g), ATP (150 mg) and 2.9 M polyP (15 mL) were added. Dosing was repeated after 22 h.

Time points were collected and analysed as for reactions carried out on a 30 mL-scale.

After 44 h, the reaction was acidified to pH 2 with 37% HCl (15.5 mL). Precipitated material was filtered under vacuum (starting at 395 mbar) through Celite® 545 (Sigma-Aldrich, 16 g total) into two 500 mL flasks. Filters were washes with toluene (200 mL). The filtrate was transferred into a 1 L separating funnel and the aqueous layer separated. Unreacted starting material was removed with a saturated solution of sodium bicarbonate (50 mL). Brine (2x40 mL followed by another 40 mL addition) was added to try to break the resulting emulsion which was subsequently removed by another filtration step. Another wash with saturated sodium bicarbonate was carried out (100 mL) and the combined aqueous layers extracted with toluene (300 mL). The combined organic phases were dried over $MgSO_4$ and toluene evaporated under vacuum (55 mbar, 40 °C). The obtained yellow liquid was placed in a vacuum over for 14 h at 30 °C to remove residual toluene and afford 3 g of isolated products (70.3% isolated yield).

Procedure for the Multigram-scale CAR Reaction with PPK12 – Experiment N72619-60

Reactions were run on an EasyMax402 system equipped with a 500 mL reaction vessel. pH was controlled with a SevenExcellence pH Meter and adjusted to pH 7.5 with 1 M NaOH and an SP-50 dosing unit. Temperature was 30 °C and stirring 400 rpm. The biotransformations contained 7.6 mg/mL CAR33, 0.76 mg/mL PPK12, 0.76 mg/mL GDH-CDX-901, 37.5 mM $MgCl_2$, 0.75 mM $NADP^+$, 225 mM D-Glucose, 0.75 mM ATP, 235 mM polyP, 65 mM 4-methoxybenzoic acid, 4.3% v/v DMSO and toluene (150 mL, 37.5% v/v). Reported concentrations refer to the final reaction volume (400 mL final volume).

Time points were collected and analysed as for reactions carried out on with ajPAP.

After 22 h, the reaction was acidified to pH 2 with 37% HCl (12.2 mL). Precipitated material was filtered under vacuum (starting at 395 mbar) through Celite® 545 (Sigma-Aldrich, 22 g total) into two 500 mL flasks. Filters were washes with toluene (100 mL). The filtrate was transferred into a 1 L separating funnel and the aqueous layer separated. The organic layer was washed with a saturated solution of sodium bicarbonate (30 mL). Brine washes (3x30 mL) broke the resulting emulsion. Combined aqueous layers were extracted with toluene (40 mL). The combined organic phases were dried over $MgSO_4$ and evaporation of toluene under vacuum (55 mbar, 40 °C) yielded 2.6 g products (yellow liquid, 73.4% isolated yield).

References

1. K. Motomura, R. Hirota, M. Okada, T. Ikeda, T. Ishida and A. Kuroda, *Appl Environ Microbiol*, 2014, **80**, 2602-2608.
2. S. M. Resnick and A. J. Zehnder, *Appl Environ Microbiol*, 2000, **66**, 2045-2051.
3. B. Nocek, S. Kochinyan, M. Proudfoot, G. Brown, E. Evdokimova, J. Osipiuk, A. M. Edwards, A. Savchenko, A. Joachimiak and A. F. Yakunin, *Proc Natl Acad Sci U S A*, 2008, **105**, 17730-17735.
4. B. P. Nocek, A. N. Khusnutdinova, M. Ruszkowski, R. Flick, M. Burda, K. Batyrova, G. Brown, A. Mucha, A. Joachimiak, Ł. Berlicki and A. F. Yakunin, *ACS Catalysis*, 2018, **8**, 10746-10760.
5. S. Mordhorst, J. Singh, M. K. F. Mohr, R. Hinkelmann, M. Keppler, H. J. Jessen and J. N. Andexer, *Chembiochem*, 2019, **20**, 1019-1022.
6. A. E. Parnell, S. Mordhorst, F. Kemper, M. Giurrandino, J. P. Prince, N. J. Schwarzer, A. Hofer, D. Wohlwend, H. J. Jessen, S. Gerhardt, O. Einsle, P. C. F. Oyston, J. N. Andexer and P. L. Roach, *Proc Natl Acad Sci U S A*, 2018, **115**, 3350-3355.
7. K. Katoh, J. Rozewicki and K. D. Yamada, *Brief Bioinform*, 2019, **20**, 1160-1166.
8. X. Robert and P. Gouet, *Nucleic Acids Res*, 2014, **42**, W320-324.
9. F. Chacón-Huete, C. Messina, F. Chen, L. Cuccia, X. Ottenwaelder and P. Forgione, *Green Chemistry*, 2018, **20**, 5261-5265.
10. L.-Y. Jiang, J.-J. Ming, L.-Y. Wang, Y.-Y. Jiang, L.-H. Ren, Z.-C. Wang and W.-C. Cheng, *Green Chemistry*, 2020, **22**, 1156-1163.
11. Y. Huang, B. Niu, Y. Gao, L. Fu and W. Li, *Bioinformatics*, 2010, **26**, 680-682.
12. M. V. Han and C. M. Zmasek, *BMC Bioinformatics*, 2009, **10**, 356.