SI APPENDIX

Carboxylic acid reductase is a versatile enzyme for the conversion of fatty acids into fuels and chemical commodities.

M. Kalim Akhtar [a, b], Nicholas J. Turner [b, 1], and Patrik R. Jones [a, 1]

[a] *Department of Biochemistry and Food Chemistry, University of Turku, Tykistökatu 6A, 6krs, 20520 Turku, Finland*

[b] *School of Chemistry, Manchester Institute of Biotechnology, University of Manchester,131 Princess Street, Manchester, M1 7DN*

[1] *To whom correspondence should be addressed: Patrik R. Jones, patjon@utu.fi & Nicholas J. Turner, nicholas.turner@manchester.ac.uk*

Methods

Strains and plasmids

The genes encoding for *M. marinum* CAR and *B. Subtilis* Sfp were synthesized with codon-optimization for *E. coli* (Genscript, USA), while *tesA*, *slr1192* and *ahr* were amplified from *E. coli* BL21(DE3) and *Synechocystis* sp. PCC 6803 genomes using the following PCR program: 95 °C for 2 min; 40 cycles of 95 °C for 30 s, 56 °C for 30 s and 72 °C for 1 min; and 72 °C for 20 min. Genetic inserts along with T7-based commercial vectors (Novagen) were cut with the appropriate restriction enzymes (New England Biolabs) and ligated using T4 DNA ligase (Fermentas) (Tables S3-S5). Multiple genes were assembled in artificial operons as described previously (1). Plasmids were used to transform *E. coli* BL21 (DE3) to generate the strains listed in Table S6. Table S7 summarizes the gene-products used to engineer the recombinant pathways.

Protein expression and purification

Overnight LB-grown pre-cultures of BL21(DE3) were used to inoculate 20 ml cultures of Overnight Express™ Instant TB Medium (Novagen) at 2% (v/v). Cultures were incubated overnight (18-24 h, 30 ^{o}C, 250 rpm). Cells were pelleted and resuspended in a lysis buffer containing lysozyme (2 mg/ml) and 2% (v/v) hexane, and incubated for 30 min at room temperature with gentle inversion. The insoluble debris was centrifuged (17,000*g*, 5 min) and the supernatant was applied to a microfuge spin column pre-filled with 200 µl His-Select® Nickel Affinity Gel (Sigma Aldrich). The column was washed five times with 0.5 ml 0.1 M Tris-HCl (pH 7.5, 21 ^{o}C) and the recombinant his-tagged protein eluted with 100 µl 0.4 M imidazole, prepared in 50 mM Tris-HCl buffer (pH 7.5, 21 ^{o}C). Protein recovery was estimated using the Bio-Rad Protein Assay (Biorad).

Enzyme characterization

All enzyme reactions were performed in triplicates and monitored at 340 nm for up to 15 min in 96-well microplates (Tecan M200 Affinity). Control reactions without addition of the purified enzyme were also included to take into account any background oxidation of NADPH. For the CAR assay, a 100 µl reaction volume typically contained the following components: CAR_{his} (0-10 µg/ml), 1 mM NADPH, 1 mM ATP, 10 mM $MgCl_2$ and 0.5 mM fatty acids. For K_m and V_{max} determinations, the various (co)substrates: fatty acids, NADPH and ATP were prepared at 11 different concentrations with the concentration range specified in the figure legends. Based on the initial reaction rates, the apparent K_m and V_{max} values were determined using the enzyme kinetics module of SigmaPlot (Systat Software, San Jose, CA). For the AHR assay, reactions typically contained the following components: $YjgB_{his}$ (1 or 10 µg/ml), 1 mM NADPH and 0.5 mM C_4-C_{12} aldehydes or C_4-C_{12} alcohols. For qualitative confirmation of aldehyde (in the case of CAR_{his}) and alcohol synthesis (in the case of $YjgB_{his}$), reaction mixtures (500 µl) were mixed with an organic solvent (75 µl) and the organic phase analyzed by GC-MS, as described below.

In vivo production of fatty alcohols and alkanes

Strains were cultivated either in Overnight Express™ Instant TB Medium (Novagen) or in a defined minimal medium containing: M9 salts (Sigma Aldrich); BME vitamins (Sigma Aldrich); 100 mM potassium phosphate buffer (pH 7.5, 21 ^oC); 2 mM MgSO4; 0.1 mM $CaCl_2$; micronutrient mix consisting of 10 nM $FeSO_4$, 3 µM $(NH_4)_6Mo_7O_{24}$, 0.4 mM boric acid, 30 µM $CoCl_2$, 15 µM $CuSO_4$, 80 µM $MnCl_2$ and 10 µM $ZnSO_4$; 2% (w/v) glucose; appropriate antibiotics, 50 µg/ml ampicillin and/or 50 µg/ml spectinomycin; and 50 µM IPTG. Cultures (2-5ml) were incubated for up to 48 h with shaking at 180 - 200 rpm in 50 ml sterile Falcon tubes. For total fatty alcohol and alkane quantification, 100 µl cell culture was vigorously mixed with 200 µl acetone, microfuged (17,000*g*, 5 min) and the resulting supernatant analyzed by GC-MS as described below. Glucose levels were quantified at 340nm, based on the reduction of NAD^+ catalyzed by glucose-6-phosphate dehydrogenase.

In vitro conversion of fatty acids to fatty alcohols and alkanes

In vitro alkane synthesis was carried out as described previously (2) with the addition of CAR_{his} (100 µg/ml), 1 mM NADPH, 1 mM ATP and 0.5 mM fatty acid (C_4-C_{16}); C_{14}-C_{16} fatty acid substrates were added as suspensions due to poor solubility. For *in vitro* fatty alcohol formation, ADC_{his} was replaced with Ahr_{his} (10 µg/ml) and NADH, N-phenylmethazonium methosulphate (PMS) & ferrous ammonium sulphate were omitted. All tubes were incubated in a heat block at 30 ^oC for up to 4 h without shaking. For analyte extraction, reaction mixtures were terminated either by the addition of acetone or chloroform. The organic phase/supernatant was analyzed by GC-MS as described below.

In vitro conversion of fatty acids to fatty alcohols and alkanes

For *in vitro* alkane synthesis, reactions were performed in 500 µl reaction volumes: 50 mM, Tris-HCl (pH 7.5, 21^oC), CAR_{his} (100 µg/ml), ADC_{his} (200 µg/ml), 1 mM NADPH, 1 mM ATP, 10 mM $MgCl_2$, 20 µM PMS, 1 mM NADH, 20 µM ferrous ammonium sulphate and 0.5 mM fatty acid substrates ranging from C_4 (butyric acid) to C_{16} (hexadecanoic acid). The protocol was modified according to ADC assays described previously (2). For *in vitro* fatty alcohol formation, ADC_{his} was replaced with $YjgB_{his}$ (10 µg/ml) and the following components were omitted: NADH, PMS & ferrous ammonium sulphate. Due to their

hydrophobicity, the C_{14} and C_{16} fatty acid substrates were added to the assays as suspensions. All reactions were performed in a heat block at 30 ^{o}C for up to 4 h without shaking, and terminated either by addition of acetone or chloroform. The supernatant or organic phase was analyzed by GC-MS as described below. Given that conversion was dependent on two NAPH-requiring reactions, fatty alcohol formation was also monitored by following the oxidation of NADPH at 340 nm.

In vitro conversion of TAGs to aldehydes

Reactions were performed in a 1 ml volume containing the following: 50mM Tris-HCl (pH 7.5, 21^{o}C), CAR_{his} (300 μg/ml), lipase (100 μg/ml), 1 mM NADPH, 1 mM ATP, 10 mM $MgCl_2$ and 1 mM suspensions or emulsions of the commercially purified TAGs: C_8 TAG (glyceryl trioctanoate) and C_{12} TAG (glyceryl tridodecanoate). All tubes were incubated in a heat block at 30 ^{o}C for up to 4 h. Reactions were terminated by vigorously mixing with an equal of chloroform. The lower organic phase was analyzed directly by size-exclusion HPLC as described below. For reactions containing C_{12} TAG, samples were concentrated by evaporation of solvent in a Genevac (30 ^{o}C, 10 min) prior to HPLC analysis.

Lipase-mediated *in vivo* formation of fatty alcohols

A 1 ml preinduced cell culture (OD ~10) of the strain, PC-Ahr_{his}, encoding for the CAR_{his} and Ahr_{his} enzymes was centrifuged (7000*g*, 10 min). The cell pellet was resuspended in an equal volume of 100 mM potassium phosphate buffer (pH 7.5, 21^{o}C) together with three distinct sources of TAGs: (i) harvested cells of the cell wall-less *Chlamydomonas reinhardtii* strain cc406, (ii) palm oil (Afroase) and (iii) coconut oil (Biona Organic). Cells were supplemented with 100mM glucose and incubated at 30 ^{o}C with 100 μl samples taken at 30 min, 1 h, 2 h and 5 h. Fatty alcohols were analyzed by GC-MS as described below.

Hydrocarbon analysis by GC-MS

Metabolite analysis was performed with an Agilent 7890A gas chromatograph equipped with a 5975 mass spectrometry detector. All samples (1 μl) were analyzed in splitless injection mode with the inlet temperature set at 300 °C and passed through an Equity-1 fused silica capillary column (Supelco) (30 m x 320 μm x 1 μm) at a flow rate of 1.9 ml/min, using helium as the carrier gas. The oven was initially held at 45 °C for 2.5 min and ramped up to 300 °C at a rate of 20 °C/min. Data was acquired within the 25–350 m/z range. For analyte identification, fragmentation patterns and retention times of the analytes were compared with the NIST mass spectral library and commercially available standards of fatty alcohols, acids and alkanes. A standard curve for quantification was prepared with commercial preparations of fatty alcohols and alkanes.

Hydrocarbon analysis by HPLC-size exclusion chromatography

For separation and detection of the TAGs, fatty acids and aldehydes, size-exclusion HPLC analysis was carried out using an Agilent 1200 Series HPLC module coupled to an Agilent refractive index detector. For analyte separation, samples (10 μl) were injected and passed through 3 size-exclusion columns (300 x 7.8 mm x 5 or 10 μm; Phenomenex) of different pore sizes; 50, 100 and 500 Å; and connected in series, at a flow rate 0.8 ml/min. The chromatography was performed in isocratic mode using super-purity grade chloroform (Romil) as the mobile phase with column and refraction index detector temperatures set at 40 °C. The retention times of the analytes were confirmed using commercial standards.

Stoichiometric evaluation

The stoichiometric conversion efficiency (percent conversion, mole product per mole glucose) was calculated by two independent methods, as described in detail below:

A maximum potential yield of C_{12} fatty alcohol per glucose is 0.33 and 0.288, respectively, without biomass formation. All stoichiometric evaluations were made in minimal media. Where an error is given in the text it represents the standard error (SEM, n=2-4).

The experimentally measured distribution of fatty alcohols in all TPC-Ahr strains was highly similar regardless of cultivation conditions, as summarized below:

		C12:1ol	C12ol	C14:1ol	C14ol	C16:1ol	C16ol	C18:1ol	C18ol
Complex media	TPC-Ahrhis	8	38	13	22	14	1	2	0
	TPC-Ahrhis	9	37	13	22	14	2	3	0
Minimal media	TPC2	8	41	8	26	12	3	2	0
	TPC2	8	40	9	26	12	3	2	0
	TPC-Ahr	8	47	7	24	10	2	1	0
	TPC-Ahr	8	47	7	23	10	3	1	0
	TPC-Ahrhis	10	49	8	21	10	2	1	0
	TPC-Ahrhis	10	48	8	21	10	2	1	0
	TPC-Ahrhis	10	36	11	25	12	2	1	2
	TPC-Ahrhis	8	30	11	27	17	3	1	3
	Average	8.7	41.3	9.4	23.9	12.2	2.2	1.7	0.5
	Standard error	0.3	2.1	0.8	0.7	0.8	0.2	0.2	0.4

The molar demand for glucose that is required for each of the main fatty alcohols, identified in the TPC-Ahr strain, was estimated based on the assumption that approximately 3 mole of CO_2 is released per mole of catabolized glucose through substrate oxidation. The molar demand for glucose was related to the distribution of fatty alcohols that was repeatedly observed; this is summarized in the Table below:

	A Average distribution (%)	**B** Stoichiometric glucose required (mM)	**C** Carbon distribution fatty alcohol per 100 mM glucose (%)	**D** Stoichiometric yield fatty alcohol per 100 mM glucose (%)
C12:1ol	8.7	26.1	7.8	2.6
C12ol	41.3	124	37.1	12.4
C14:1ol	9.4	33	9.9	2.8
C14ol	23.9	83.5	25	7.1
C16:1ol	12.2	48.8	14.6	3.7
C16ol	2.2	8.9	2.7	0.7
C18:1ol	1.7	7.5	2.3	0.5
C18ol	0.5	2.4	0.7	0.2
Biomass	--	--	0	--
Sum	100	335.5	100	29.9

The analysis suggests that the maximum potential molar yield of fatty alcohols with the TPC-Ahr strains is 30% (mole of fatty alcohols per mole of glucose), however, under these conditions there would be no biomass formation. Since fatty alcohol production was studied under exponential batch growth conditions, the same analysis was also repeated with the assumption that 25% of all glucose is used for biomass-formation. The maximum yield of fatty alcohols is then 22.4%.

An alternative approach to estimate the maximum potential yield of fatty alcohols is to carry out the so-called 'flux balance analysis'. The maximum potential molar yield was calculated for C_{12} alcohol with a stoichiometric model of *iJR*904 *E. coli* to which the TPC pathway was added. Within the *E. coli i*JR904 stoichiometric model (3), the following three reactions were added:

TES: C_{12}-ACP + H_2O = C_{12} fatty acid + ACP

CAR: C_{12} fatty acid + NADPH + H^+ + ATP = C_{12} aldehyde + $NADP^+$ + AMP + PPi

AHR: C_{12} aldehyde + NADPH → C_{12} alcohol + $NADP^+$

The *i*JR904 (3) and *i*AF1260 (4) models do not include any pathways for C_{14}-ACP and C_{16}-ACP formation. The calculation of the potential conversion rate for all fatty alcohols is therefore estimated according to C_{12} fatty alcohol synthesis even though the C_{14} and C_{16} alcohols are produced with and without mono-unsaturation.

After loading the *i*JR904_GlcMM model into COBRA Toolbox 1.3.3 (5), the three reactions are added accordingly:

```
modelTes = addReaction(model,'TesA',{'ddcaACP[c]','h2o[c]','dodecanoate[c]','ACP[c]'},[-1 -1 1 1], false);

modelCAR = addReaction(modelTes,'CAR',{'dodecanoate[c]','nadph[c]','h[c]','atp[c]','dodecanal[c]','nadp[c]','amp[c]','ppi[c]'},[-1 -1 -1 -1 1 1 1 1], false);

modelAHR = addReaction(modelCAR,'AHR',{'dodecanal[c]','nadph[c]','dodecanol[c]','nadp[c]'},[-1 -1 1 1], false);
```

Then, an exchange reaction is also added for C_{12} alcohol:

```
modelTPC = addReaction(modelAHR,'dodecanolEX',{'dodecanol[c]'},[-1], true);
```

In order to obtain the maximum potential conversion rate, we changed the objective function from biomass to 'dodecanolEX', which represents the rate of accumulation of C_{12} fatty alcohol given that C_{12} fatty alcohol does not excrete from wild-type *E. coli* cultivated in minimal media, and optimize, as follows:

```
modelTPCobjC12 = changeObjective(modelTPC, 'dodecanolEX');
solution = optimizeCbModel(modelTPCobjC12,'max',false,false)
```

A solution of f: 1.7266 was obtained using glpk solver.

Since the maximum glucose uptake rate is 6 mmol per gram dry weight per h, the maximum potential molar conversion efficiency is:

1.7266/6 = 0.288 mol C_{12} fatty alcohol per mole glucose. Under such conditions, there would be no biomass formation as the robustness analysis shows above.

If the above analysis is carried out also for the cyanobacterial pathway,

AAR: C_{12}-ACP + NADPH → C_{12} aldehyde + NADP + ACP

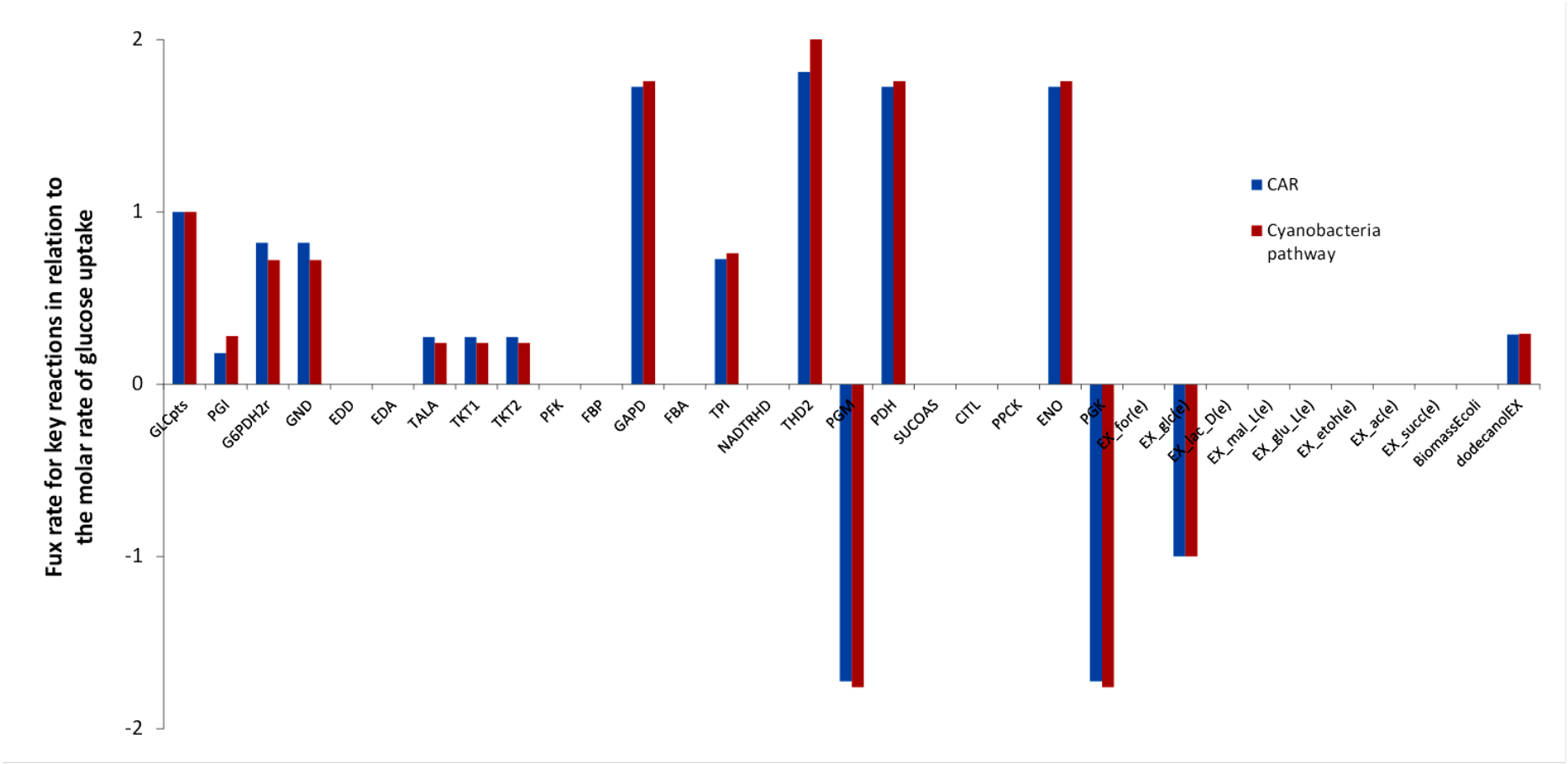
with the following reactions added to COBRA:

```
modelSchirmer = addReaction(model,'ACR',{'ddcaACP[c]','nadph[c]','dodecanal[c]','ACP[c]','nadp[c]'},[-1 -1 1 1 1], false);

modelAHR2 = addReaction(modelSchirmer,'AHR',{'dodecanal[c]','nadph[c]','dodecanol[c]','nadp[c]'},[-1 -1 1 1 1], false);

modelTPC2 = addReaction(modelAHR2,'dodecanolEX',{'dodecanol[c]'},[-1], true);

robustnessAnalysis(modelTPC2, 'dodecanolEX');
```

we obtain an f-value of 1.7600, 1.9% greater potential yield than the CAR pathway, without any biomass formation.

In order to provide insight into the relative distribution of fluxes in the optimized models, the fluxes for key enzymes in central carbon metabolism and the major excretion products were extracted from the optimized solution. The values are plotted on the next page, with enzyme names as given for the *i*JR904 *E. coli* genome-scale model (3). The main difference between the two pathways is the additional ATP requirement of the CAR-dependent pathway. Notably, the simulation is unrealistic as there is no biomass-formation. Still, the below bar graph suggests that the task of reducing $NADP^+$ is shifted slightly away from the proton gradient dependent transhydrogenase PntAB (THD2) towards the pentose phosphate pathway (as shown by the ratio of flux through PGI (phosphoglucoseisomerase) vs. G6PDH2r (glucose-6-phosphate dehydrogenase, Zwf)). As enhanced flux through the oxidative pentose phosphate pathway results in increased loss of carbon, in the form of CO_2, there is less fatty alcohol produced.

All estimates of Gibbs free energy changes under standard conditions were obtained using the eQuilibrator online platform (6).


Fux rate for key reactions in relation to the molar rate of glucose uptake
2
1
0
-1
-2
CAR
Cyanobacteria pathway
GLCpts
PGI
G6PDH2r
GND
EDD
EDA
TALA
TKT1
TKT2
PFK
FBP
GAPD
FBA
TPI
NADTRHD
THD2
PGM
PDH
SUCOAS
CITL
PPCK
ENO
PGK
EX_for(e)
EX_glc(e)
EX_lac_D(e)
EX_mal_L(e)
EX_glu_L(e)
EX_etoh(e)
EX_ac(e)
EX_succ(e)
BiomassEcoli
dodecanolEX

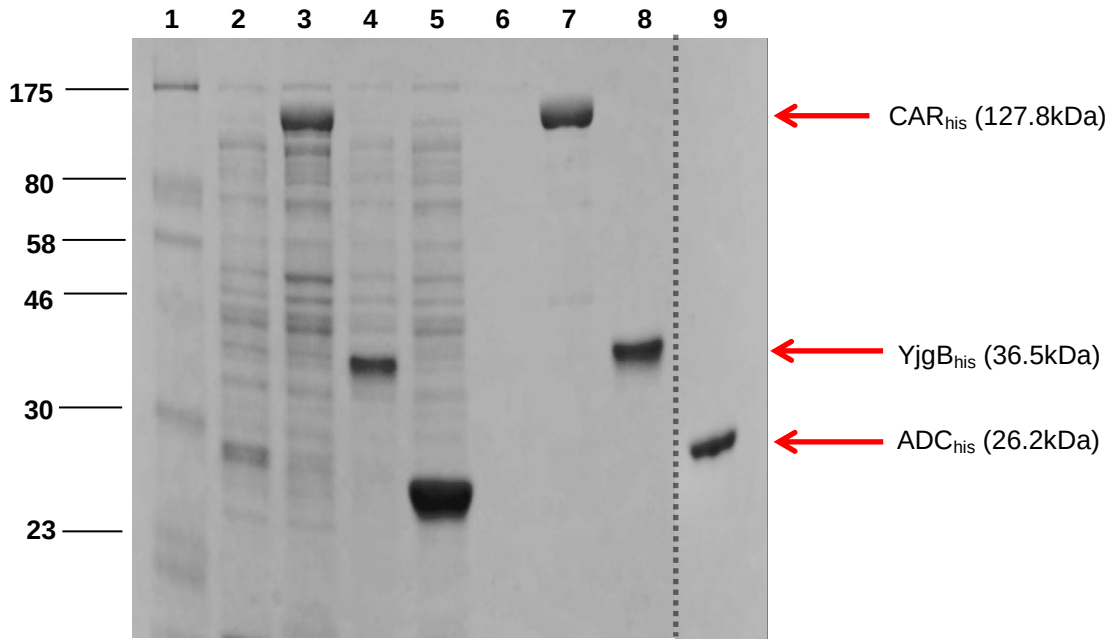


Fig. S1 SDS-PAGE of purified preparations of CAR_{his}, Ahr_{his} and ADC_{his}. Proteins were separated on 12% (w/v) acrylamide gel and stained with Coomassie Blue (7). Lane 1, molecular-mass markers (in kDa); lane 2, crude fraction of BL21(DE3); lane 3, crude fraction of PC; lane 4, crude fraction of Ahr_{his}; lane 5, crude fraction of ADC_{his}; lane 6, his-tag eluate fraction from BL21(DE3); lane 7, purified CAR_{his}; lane 8, purified crude Ahr_{his}; and lane 9, purified ADC_{his}. Purified recombinant proteins are marked with an arrow and their theoretical molecular weights indicated in brackets.

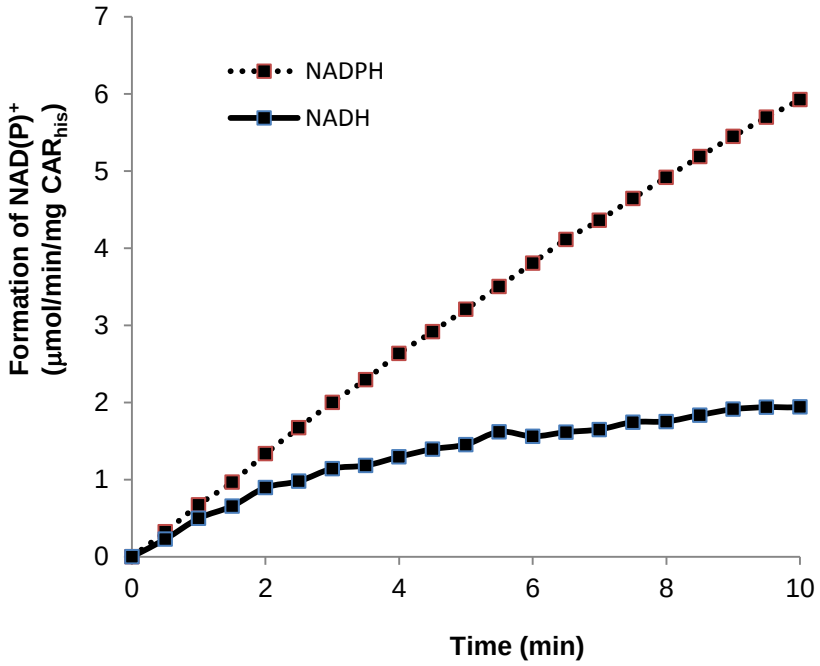


Fig. S2 NADPH specificity of CAR_{his}. Reactions were carried out in 50 mM Tris-HCl (pH 7.5, 21^{o}C) containing purified CAR_{his} (0.25 µg/ml), 2 mM benzoic acid, 10 mM $MgCl_2$ and 1mM ATP with either 1 mM NADPH or 1 mM NADH as the source of reductant. Absorbance was monitored at 340 nm at 30 s intervals over a 10 min period.

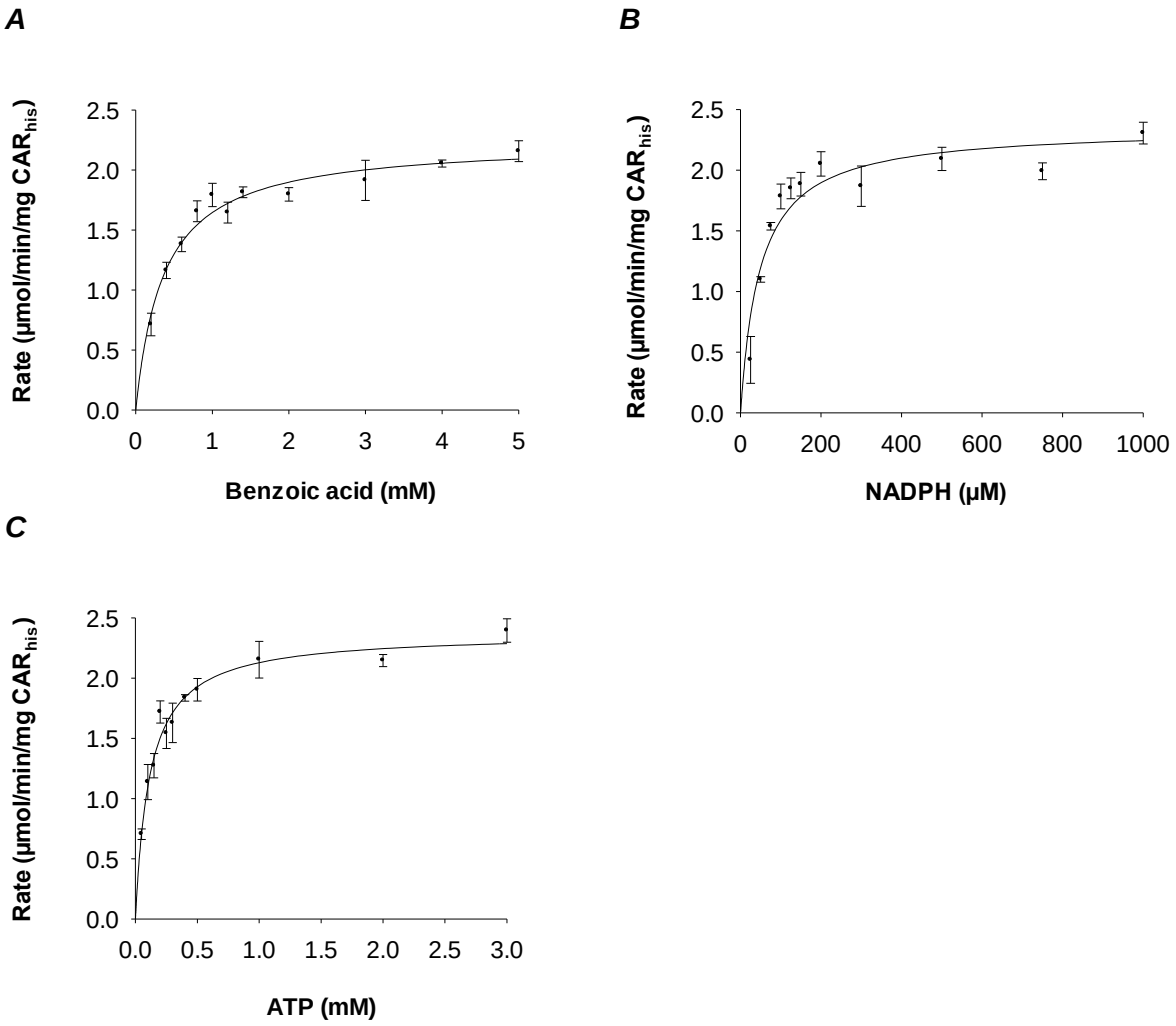


Fig. S3 Kinetic analysis of CAR_{his} in the presence of benzoic acid. Reactions were carried out in the presence of varying concentrations of **(*A*)** benzoic acid (0.2-0.5 mM), **(*B*)** NADPH (250-1000 µM) and **(*C*)** ATP (0.05-3 mM). The K_m and V_{max} were determined from non-linear regression plots, using the enzyme kinetics module of Sigmaplot (Systat Software, San Jose, CA). Results are means ± SEM of triplicate experiments.

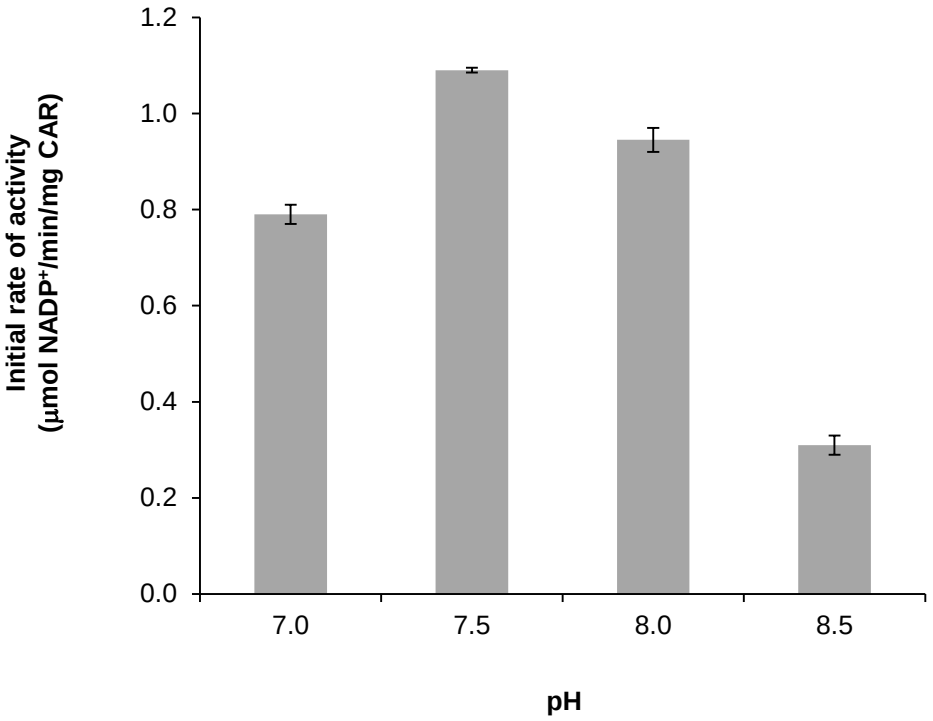


Fig. S4 Effect of pH on CAR_{his} activity. Reactions were carried out in 50 mM Tris-HCl buffer (21^{o}C) containing purified CAR_{his}, 2 mM benzoic acid, 1 mM ATP and 1 mM NADPH within the pH range of 7 to 8.5. The initial rate of CAR activity was determined from the rate of NADPH oxidation at 340 nm. Results are means ± SEM of duplicate experiments.

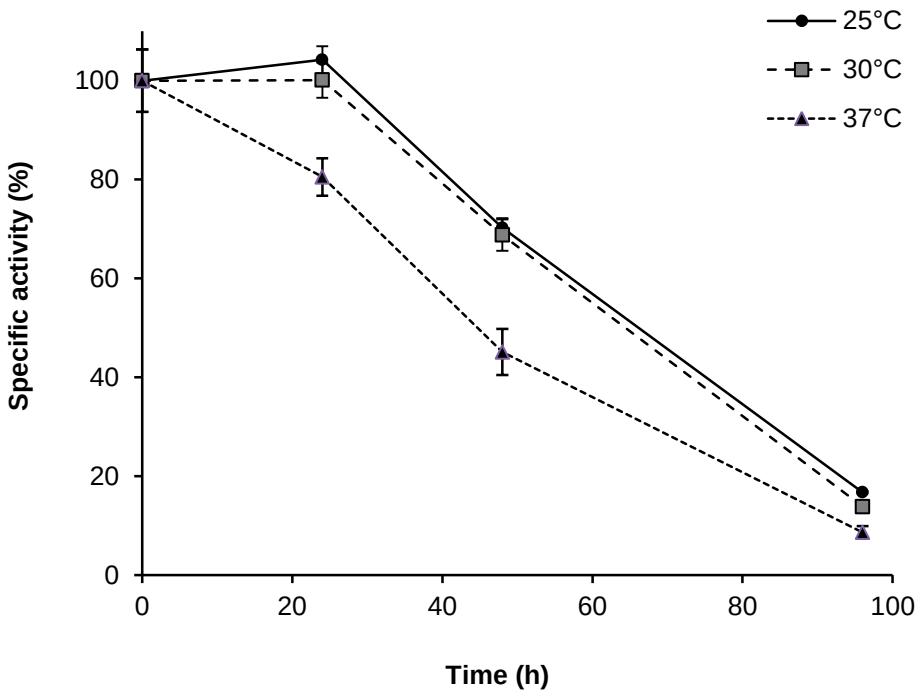
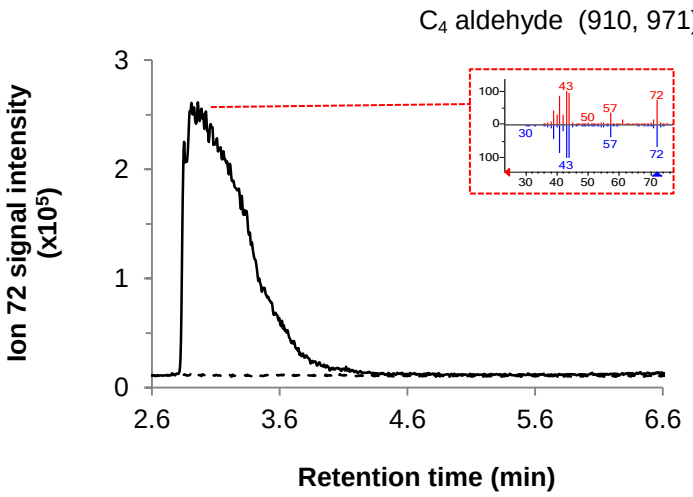
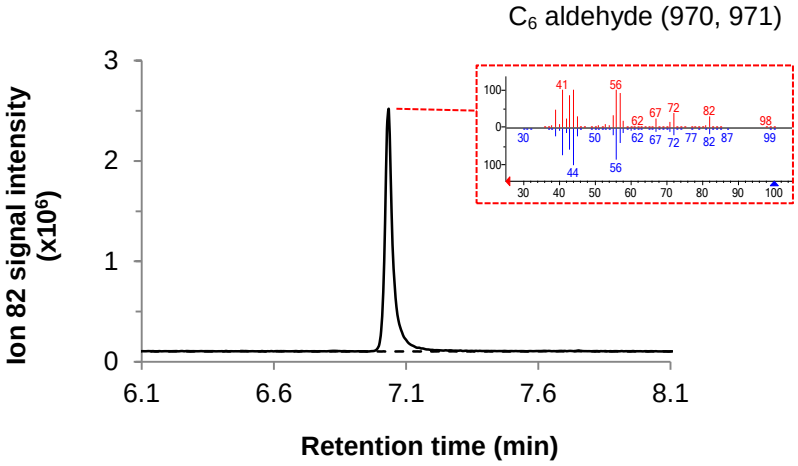


Fig. S5 Thermostability of CAR_{his}. In the absence of reducing agents, CAR_{his} was incubated without shaking at three different temperatures; 25^{o}C, 30^{o}C and 37 ^{o}C; for up to 96 h in 50 mM Tris-HCl buffer (pH 7.5, 21^{o}C). Samples were taken at specific time intervals and the initial rate of CAR activity determined in the presence of 2 mM benzoic acid, 1 mM ATP, 10 mM $MgCl_2$ and 1 mM NADPH. Results are means ± SEM. of duplicate experiments.

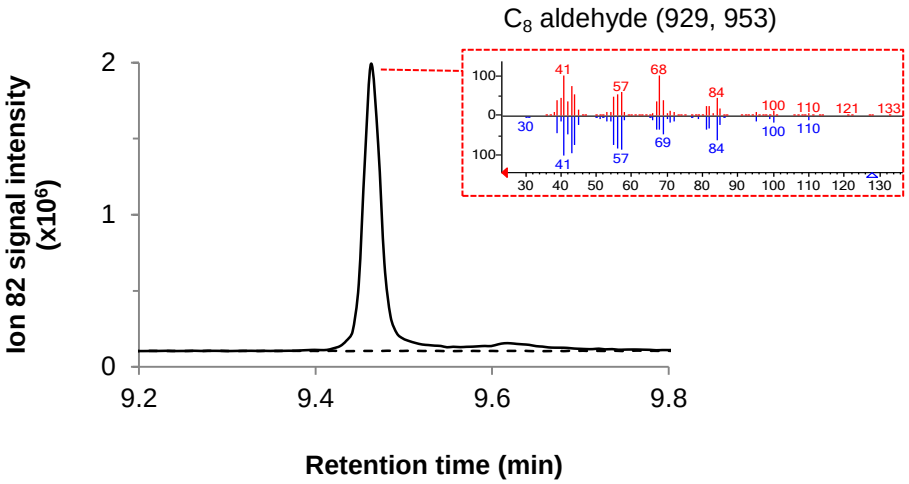
A


C4 aldehyde (910, 971)
Ion 72 signal intensity (x10^5)
Retention time (min)

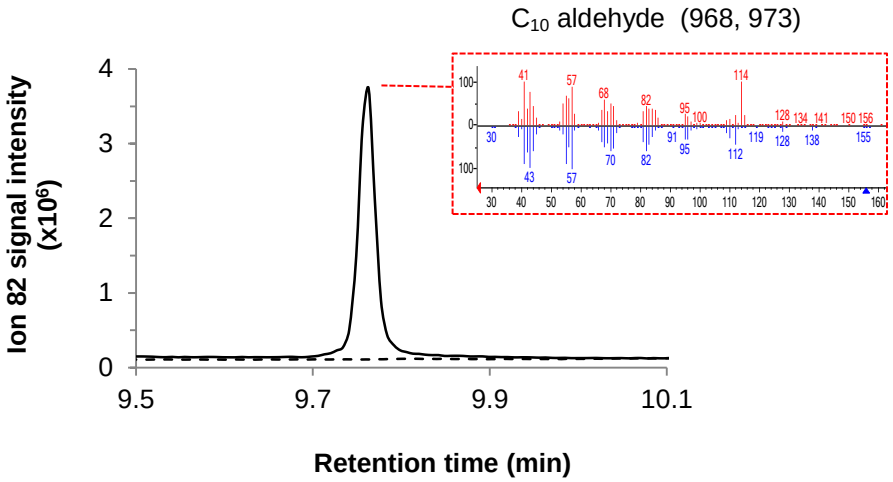
B


C6 aldehyde (970, 971)
Ion 82 signal intensity (x10^6)
Retention time (min)

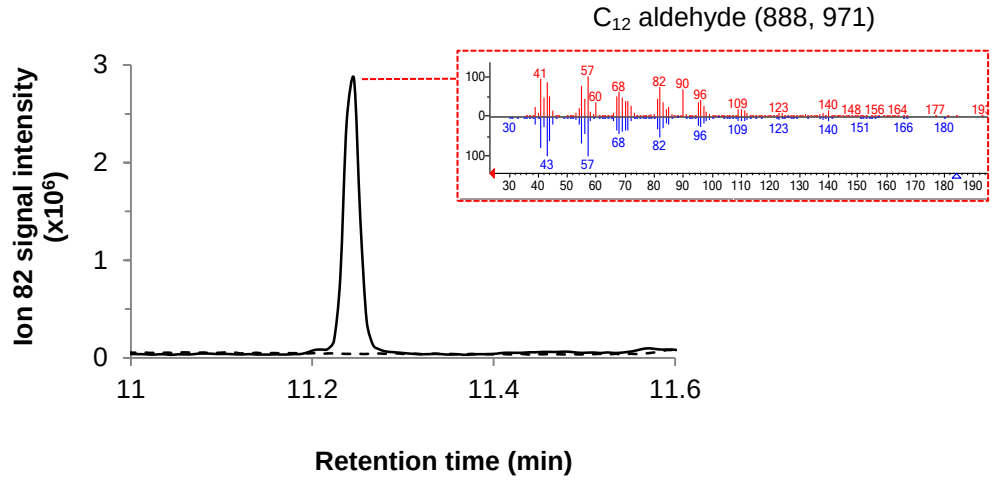
C


C8 aldehyde (929, 953)
Ion 82 signal intensity (x10^6)
Retention time (min)

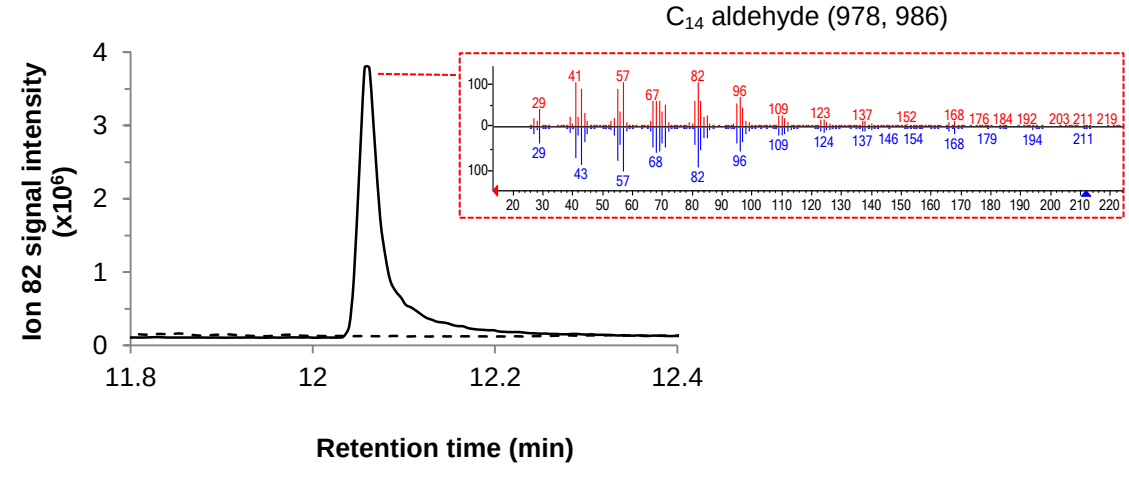
D


C10 aldehyde (968, 973)
Ion 82 signal intensity (x10^6)
Retention time (min)

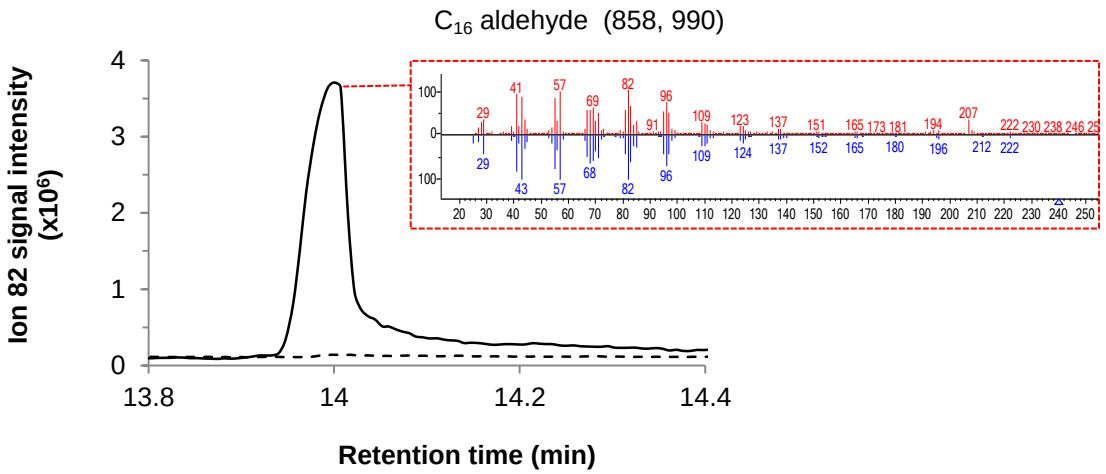
E


C12 aldehyde (888, 971)
Ion 82 signal intensity (x10^6)
Retention time (min)

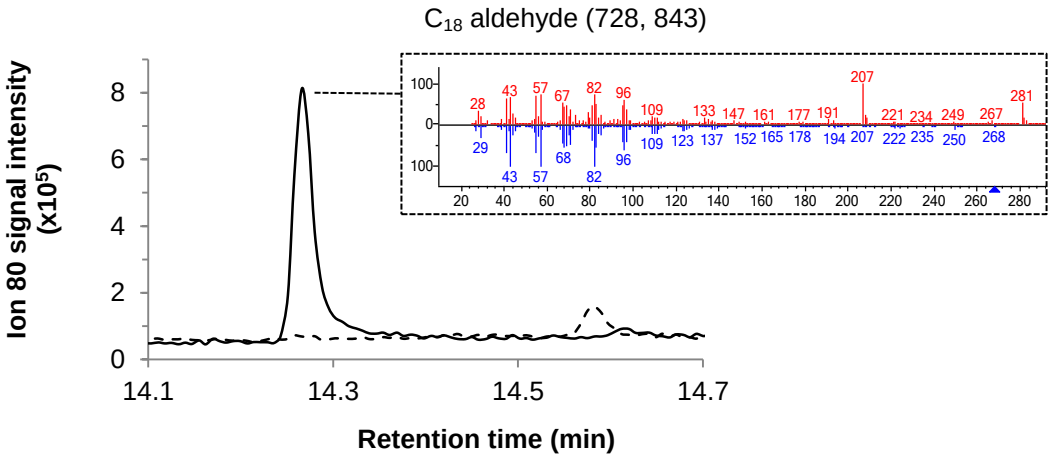
F


C14 aldehyde (978, 986)
Ion 82 signal intensity (x10^6)
Retention time (min)

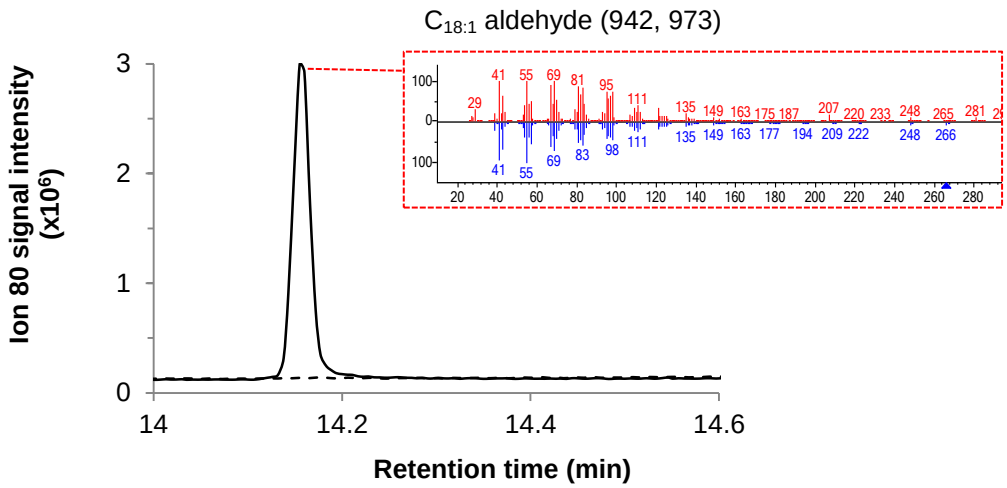
G


C16 aldehyde (858, 990)
Ion 82 signal intensity (x10⁶)
Retention time (min)

H


C18 aldehyde (728, 843)
Ion 80 signal intensity (x10⁵)
Retention time (min)

I


C18:1 aldehyde (942, 973)
Ion 80 signal intensity (x10⁶)
Retention time (min)

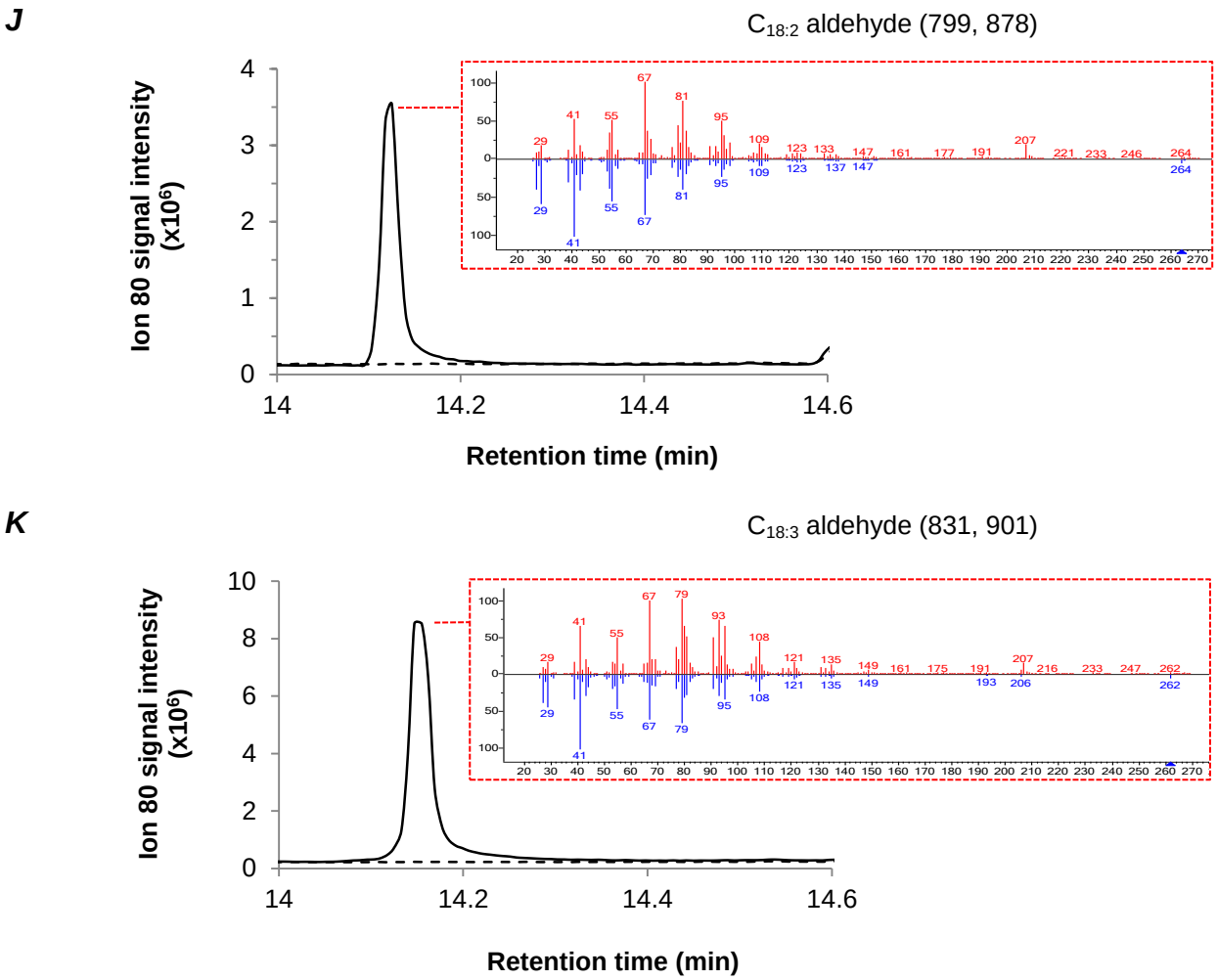


Fig. S6 GC-MS confirmation of aldehyde synthesis by CAR_{his}. Reactions were carried out in the presence of CAR_{his} (100-300 µg/ml), 1 mM NADPH, 1 mM ATP, 10 mM $MgCl_2$ and 0.5mM C_4-C_{18} fatty acids **(*A*-*K*)**. Reactions were incubated at 30 ^oC for up to 4 h or in the case of the longer chain substrates ($\geq C_{16}$) up to 12 h. Reactions were mixed with chloroform and the organic phase analyzed by GC-MS (see Methods). For clarity, only the peak representing the aldehyde is shown. Dashed lines represent chromatograms without the addition of CAR_{his}. Aldehyde formation was confirmed by comparing the mass spectrum of the analyte (in red) against reference standards (in blue) from the NIST spectral database. The match factors and reverse match factors quantitatively describe the spectral match between a sample spectrum and the library spectrum, and are indicated in brackets. The reverse match factor is obtained by excluding peaks that are present in the sample spectrum but not the library spectrum. Both values are derived from a modified cosine of the angle between the spectra, otherwise known as the normalized dot product. A value of 900 or greater signifies an excellent match; 800–900, a good match; and 700–800, a fair match.

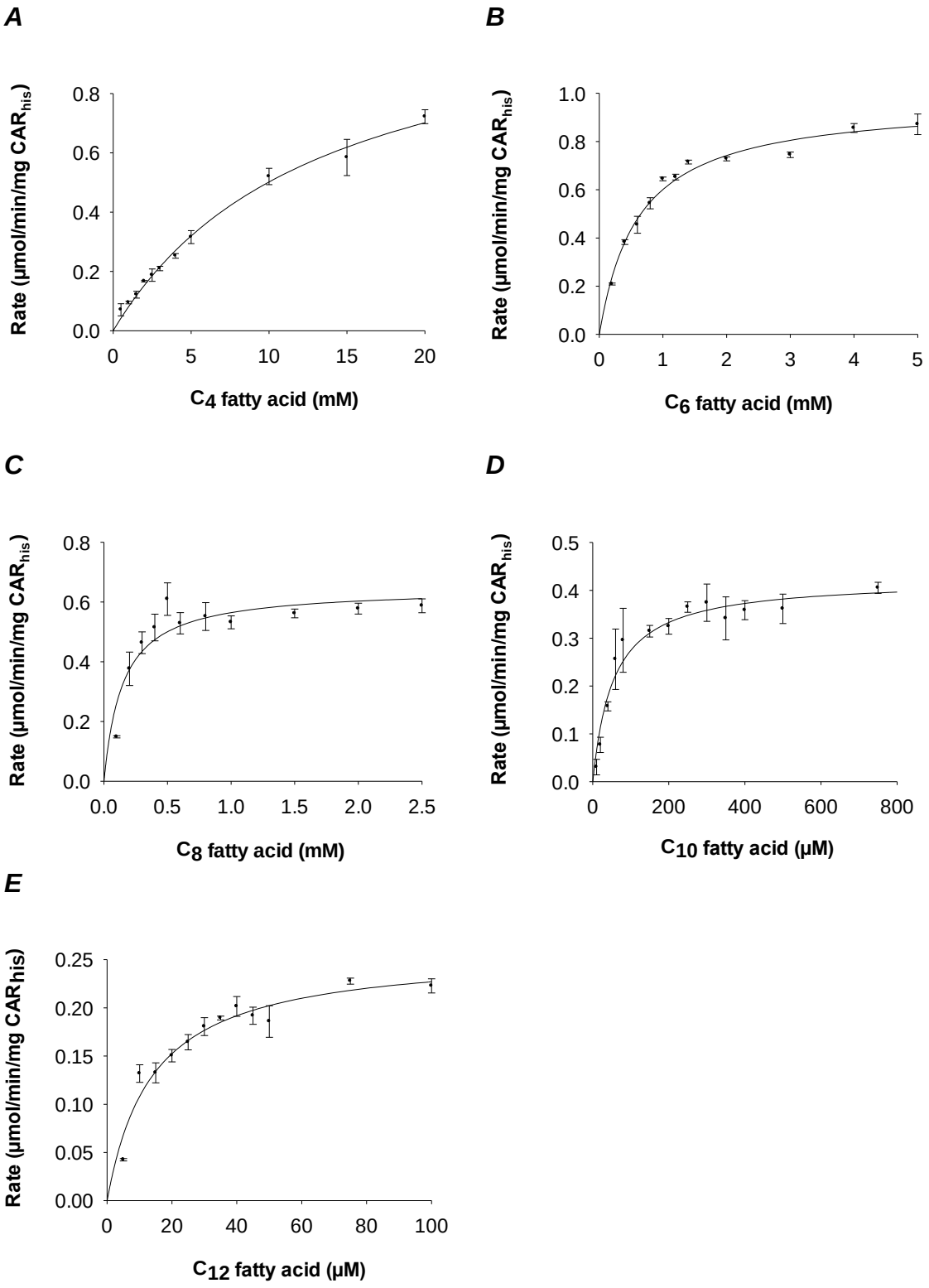
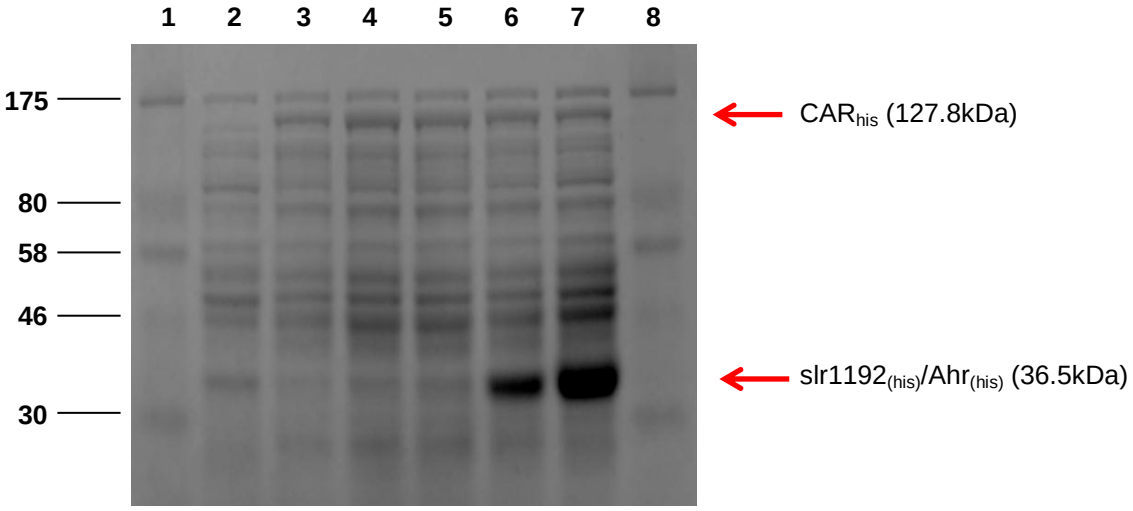

A
Rate (µmol/min/mg CARhis)
0.0
0.2
0.4
0.6
0.8
0
5
10
15
20
C4 fatty acid (mM)
B
Rate (µmol/min/mg CARhis)
0.0
0.2
0.4
0.6
0.8
1.0
0
1
2
3
4
5
C6 fatty acid (mM)
C
Rate (µmol/min/mg CARhis)
0.0
0.2
0.4
0.6
0.8
0.0
0.5
1.0
1.5
2.0
2.5
C8 fatty acid (mM)
D
Rate (µmol/min/mg CARhis)
0.0
0.1
0.2
0.3
0.4
0.5
0
200
400
600
800
C10 fatty acid (µM)
E
Rate (µmol/min/mg CARhis)
0.00
0.05
0.10
0.15
0.20
0.25
0
20
40
60
80
100
C12 fatty acid (µM)

Fig. S7. Kinetic analysis of CAR_{his} based on C_4 to C_{12} saturated fatty acid substrates. Reactions were carried out in 50mM Tris-HCl buffer containing CAR (0.5 to 4 μg/ml), 1 mM NADPH and fatty acid substrates ranging from 0 to 20 mM. For the C_4 substrate, the buffering capacity was increased to 100mM. Reactions were monitored at 340nm at 30 s intervals. The apparent K_m and V_{max} values were determined from non-linear regression plots using the enzyme kinetics module of SigmaPlot (Systat Software, San Jose, CA). The saturated fatty acids evaluated were: **(*A*)** C_4 fatty acid (0.5-20 mM), **(*B*)** C_6 fatty acid (0.2-5 mM), **(*C*)** C_8 fatty acid (0.1-2.5 mM), **(*D*)** C_{10} fatty acid (10-750 μM) and **(*E*)** C_{12} fatty acid (5-100 μM). Results are means ± SEM of triplicate experiments.

A


1
2
3
4
5
6
7
8
175
80
58
46
30
CAR$_{his}$ (127.8kDa)
slr1192$_{(his)}$/Ahr$_{(his)}$ (36.5kDa)

B

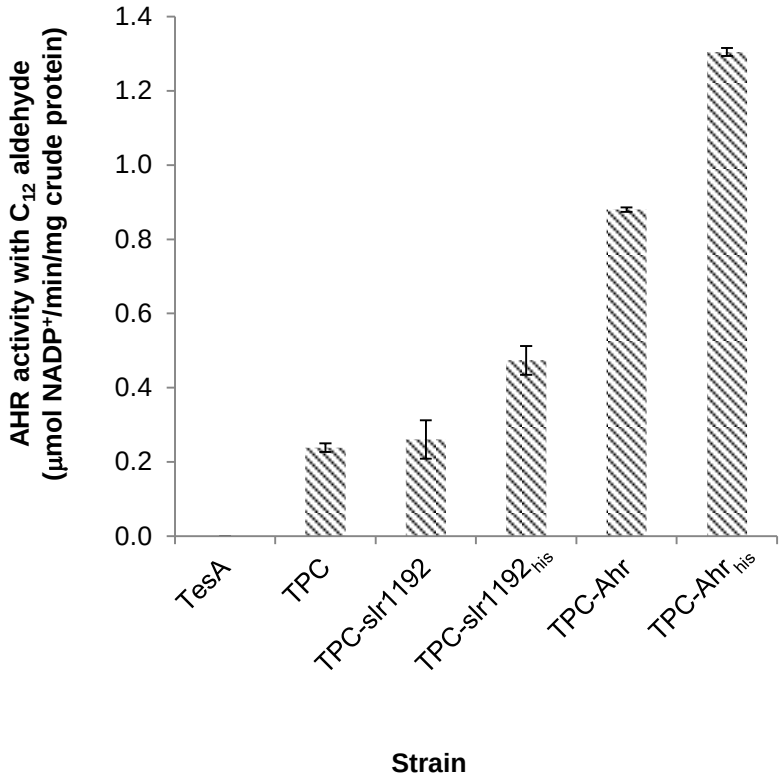

AHR activity with C12 aldehyde
(µmol NADP+/min/mg crude protein)
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
TesA
TPC
TPC-slr1192
TPC-slr1192his
TPC-Ahr
TPC-Ahrhis
Strain

Fig. S8. Activity and expression levels of AHR in crude soluble fractions. **(*A*)** The crude fractions were analyzed by SDS-PAGE for expression levels of AHR. Lane 1, molecular-mass markers (in kDa); lane 2, crude fraction of TesA; lane 3, crude fraction of TPC; lane 4, crude fraction of TPC-slr1192; lane 5, crude fraction of TPC-slr1192$_{his}$; lane 6, crude fraction of TPC-Ahr; lane 7, crude fraction of TPC-Ahr$_{his}$; and lane 8, molecular-mass markers (in kDa). Proteins bands representing the CAR and AHR enzymes are indicated by red arrows. **(*B*)** AHR activity was determined in crude fractions of all fatty alcohol-producing strains using C_{12} aldehyde as the substrate. The TesA strain was included as a negative control. Results are means ± SEM of duplicate experiments.

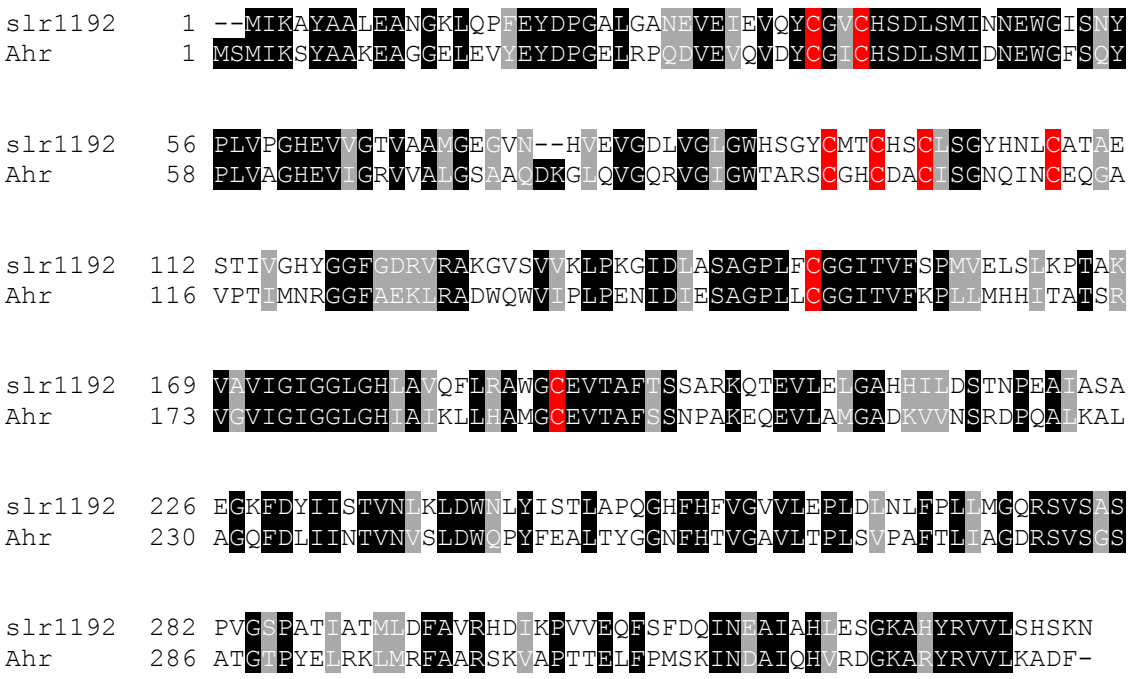


Fig. S9. Amino acid sequence alignment of *E. coli* Ahr and *Synechocystis* sp. PCC 6803 slr1192. Sequences were aligned using ClustalW 1.8 (http://www.ebi.ac.uk/Tools/msa/) (8) and formatted using Boxshade 3.21 (http://www.ch.embnet.org/software/BOX_form.html). Dark-shaded boxes depict identical amino acids, while lighter-shaded boxes signify amino acid residues with similar properties. Conserved amino acid residues for zinc binding are highlighted in red.

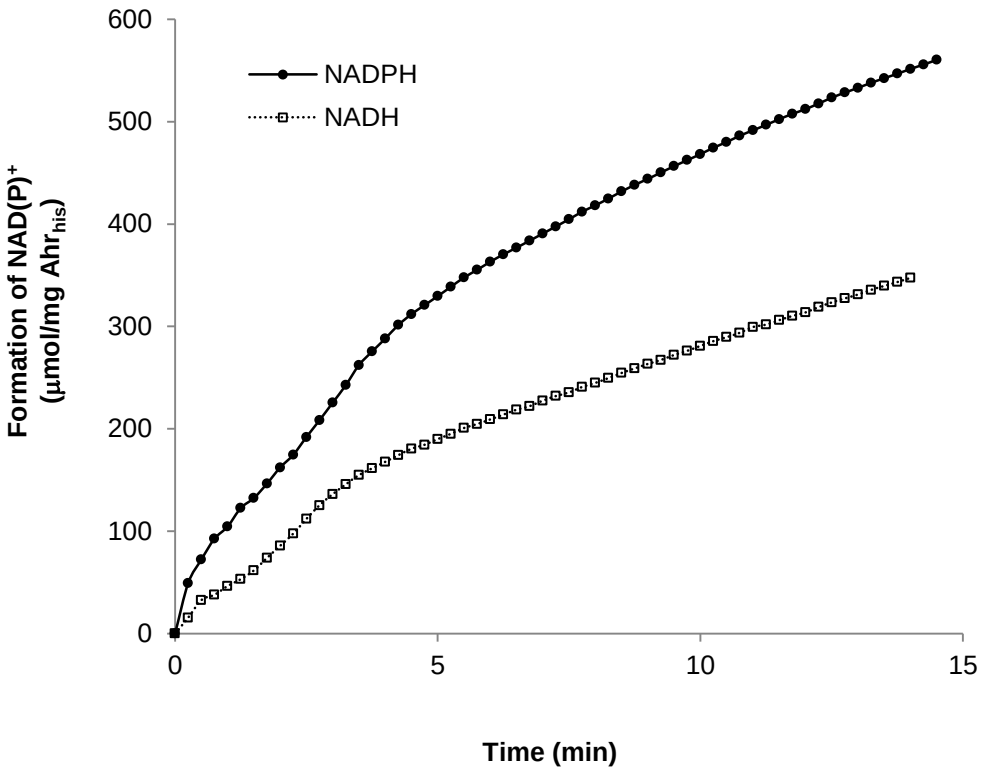
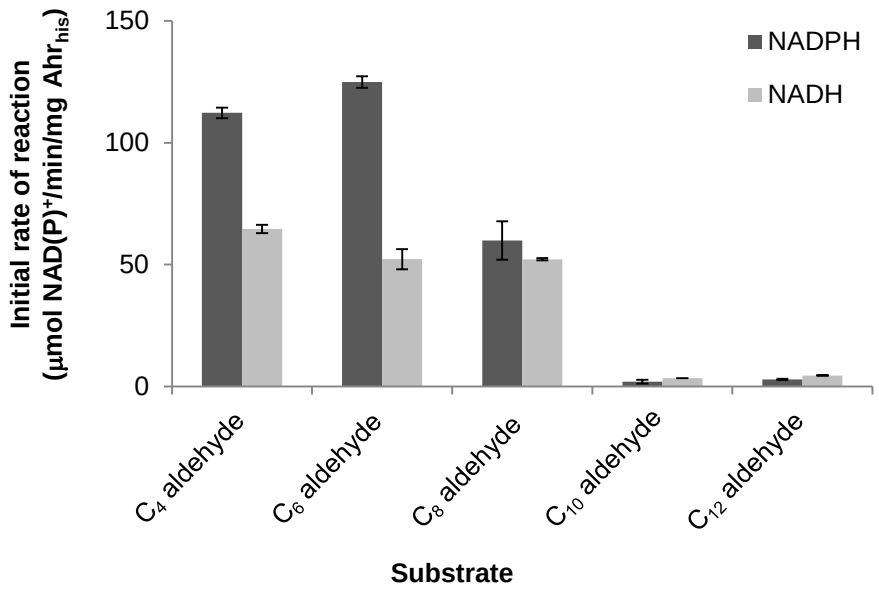


Fig. S10. NADPH specificity of Ahr_{his}. Reactions were carried out in 50 mM Tris-HCl (pH 7.5, 21^{o}C) buffer containing Ahr_{his} (1 µg /ml), 0.5 mM NADPH or NADH and 0.5 mM C_6 aldehyde. Absorbance was monitored at 340 nm at 15 s intervals over a 15 min period.

A



B

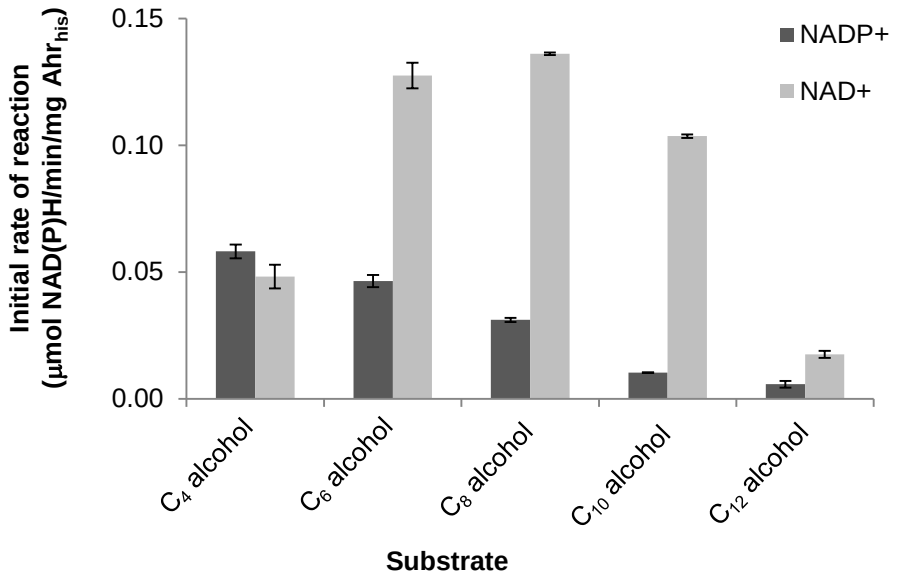
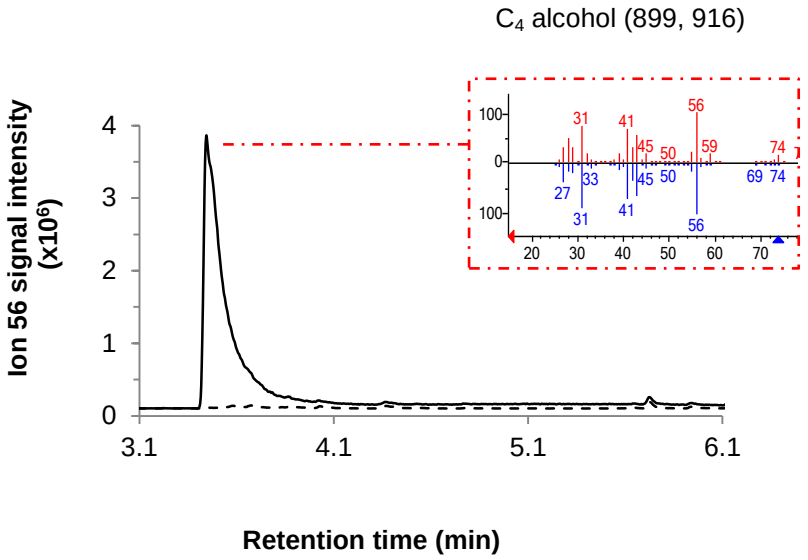
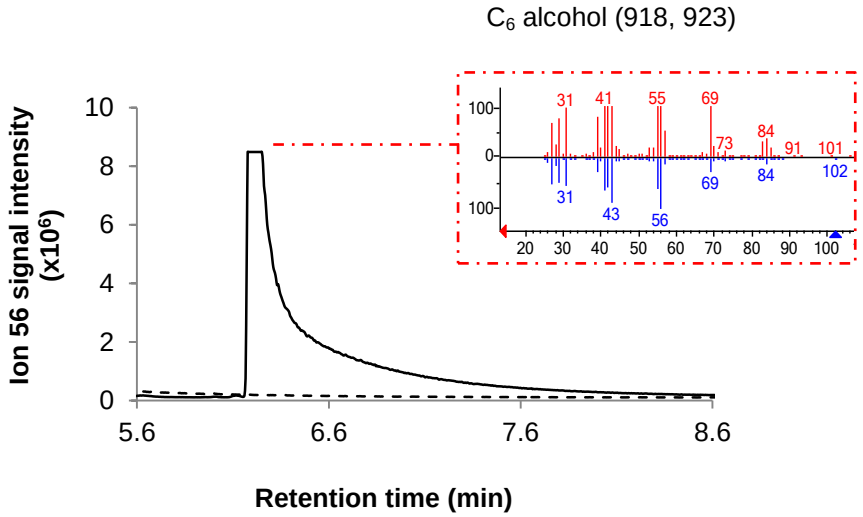


Fig. S11. Substrate specificity of Ahr_{his}. Reactions were carried out in 50mM Tris-HCl buffer containing Ahr_{his} (1 µg or 10 µg /ml) , 0.5 mM NAD(P)H or $NAD(P)^+$ with either **(*A*)** 0.5mM aldehydes (C_4-C_{12}) or **(*B*)** 0.5 mM alcohols (C_4-C_{12}). Initial rates for both reactions were determined from the rate of NAD(P)H oxidation or $NAD(P)^+$ reduction. Results are means ± SEM of duplicate experiments.

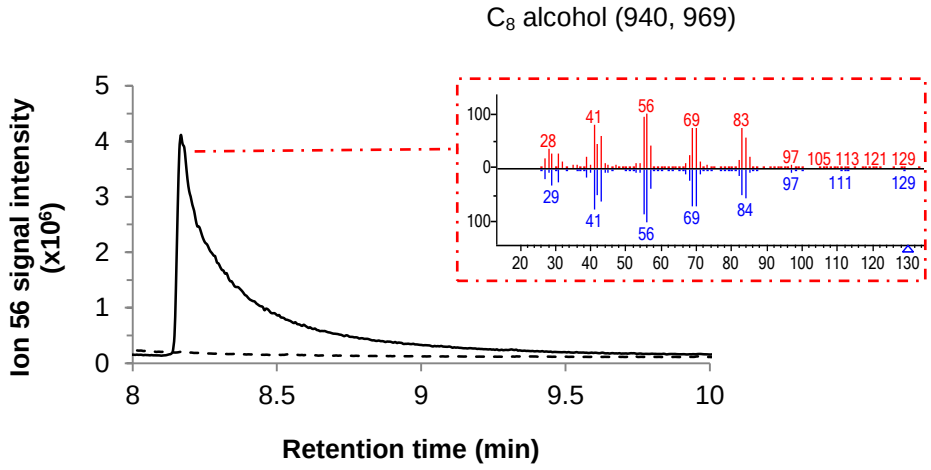
A


C4 alcohol (899, 916)
Ion 56 signal intensity (x10^6)
Retention time (min)

B


C6 alcohol (918, 923)
Ion 56 signal intensity (x10^6)
Retention time (min)

C


C8 alcohol (940, 969)
Ion 56 signal intensity (x10^6)
Retention time (min)

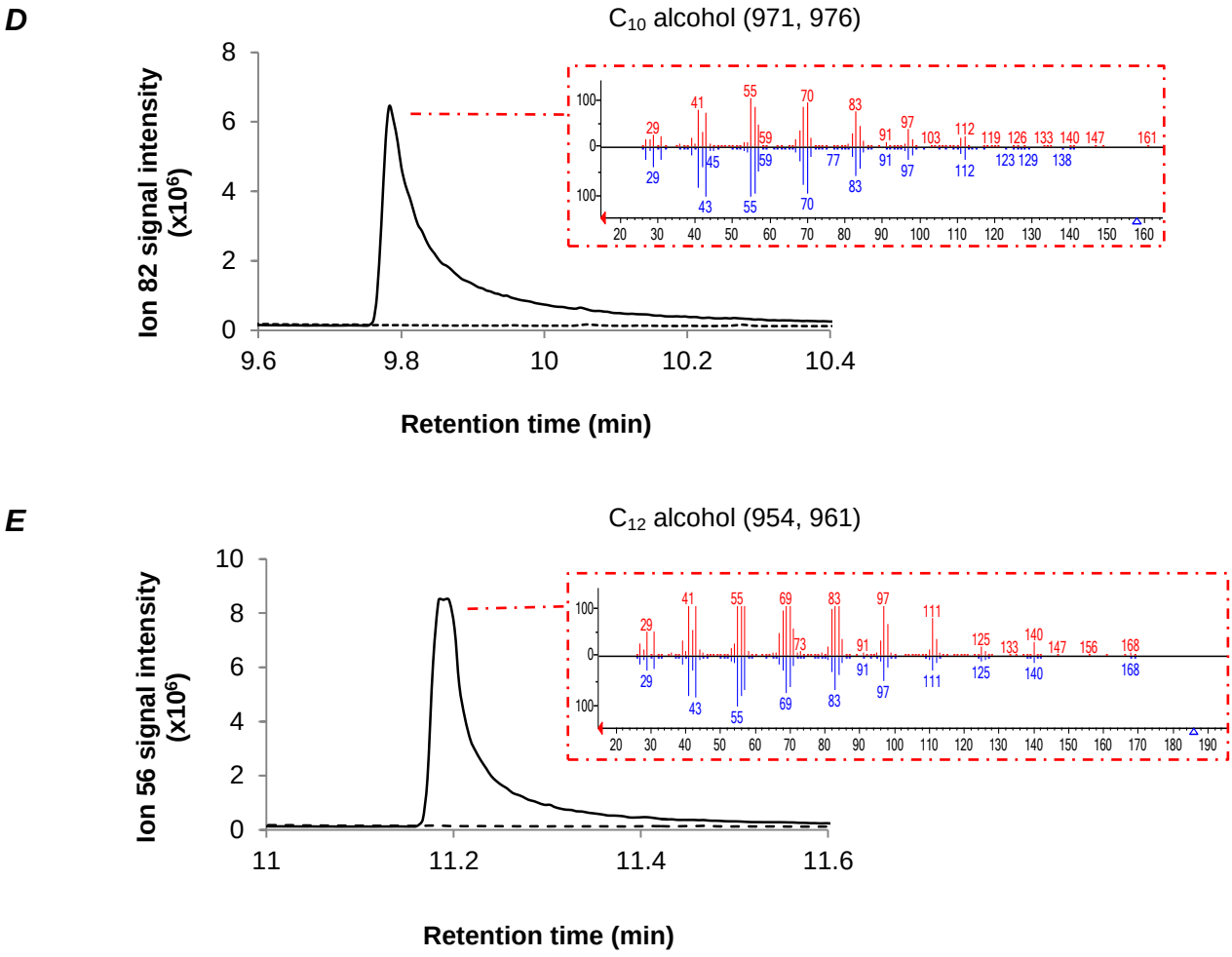


Fig. S12. GC-MS verification of alcohol synthesis by Ahr_{his}. Reactions were carried out in 50mM Tris-HCl buffer containing Ahr_{his} (10-30 µg/ml), 1 mM NADPH and 0.5mM C_4-C_{12} aldehydes **(*A*-*E*)**. Reactions were incubated at 30 ^{o}C for up to 2 h and terminated by addition of chloroform. The organic phase analyzed by GC-MS (see Methods). For clarity, only the peak representing the alcohol is shown. Dashed lines represent chromatograms without the addition of the enzyme. Alcohol formation was confirmed by comparing the mass spectrum of the analyte (in red) against reference standards (in blue) from the NIST spectral database. The match factors and reverse match factors are indicated in brackets. Refer to **Fig. S6** for explanation of (reverse) match factors.

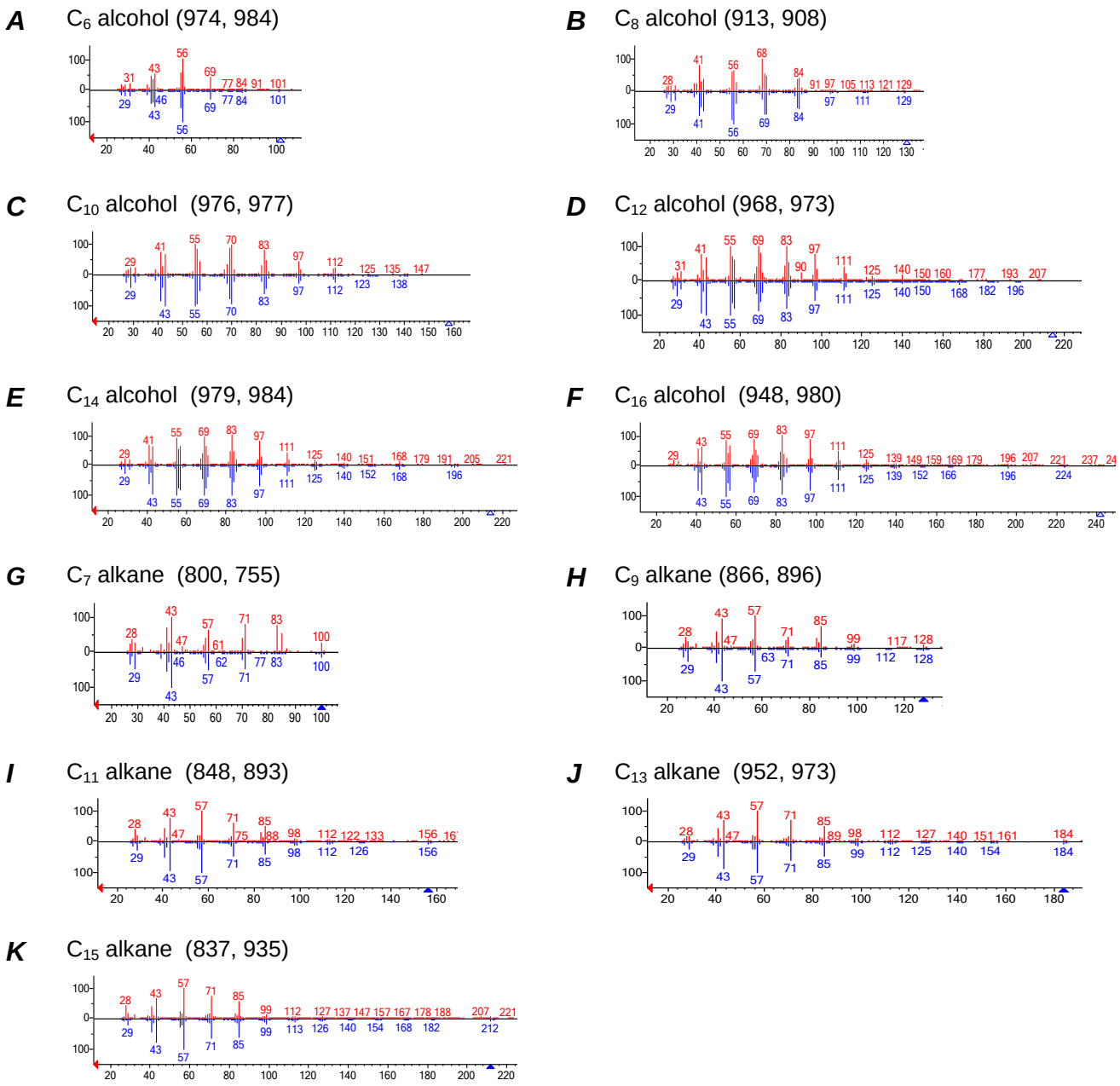
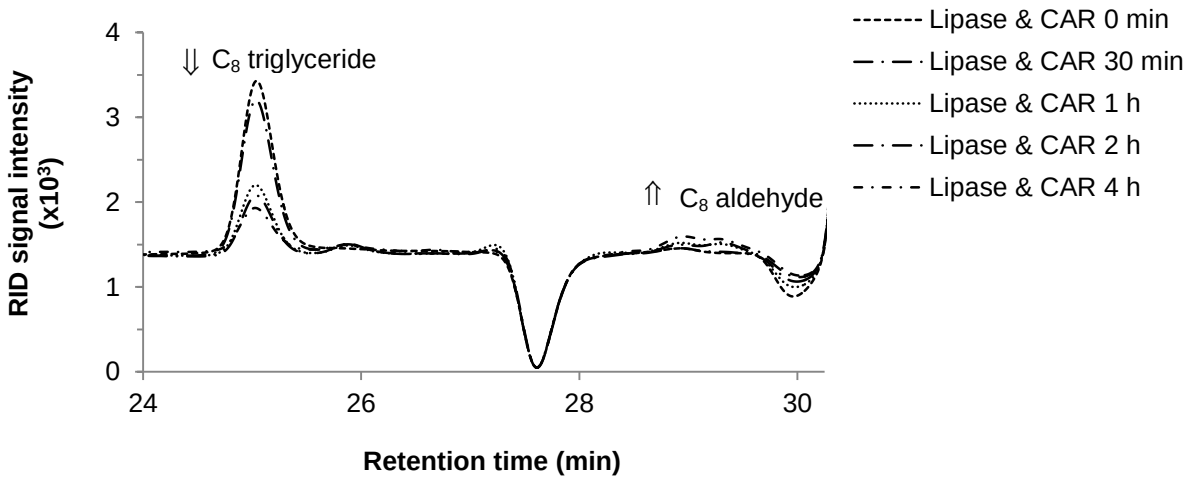
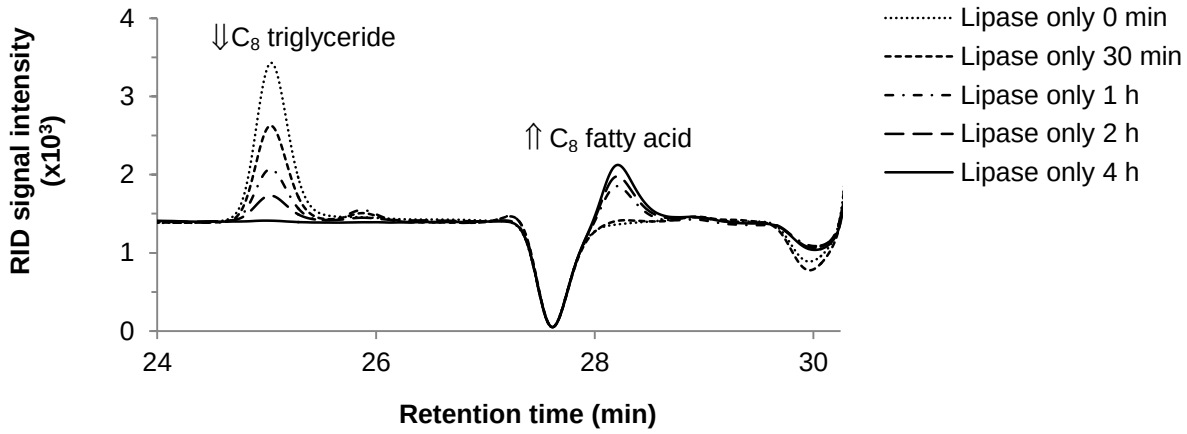


Fig. S13. GC-MS confirmation of enzyme-mediated formation of alcohols and alkanes. Alcohol **(*A*-*F*)** and alkane formation **(*G*-*K*)** were verified by comparing the mass spectrum of the analyte (in red) against the reference standards (in blue) from the NIST spectral database. The match factors and reverse match factors are indicated in brackets. Refer to **Fig. S6** for explanation of (reverse) match factors.

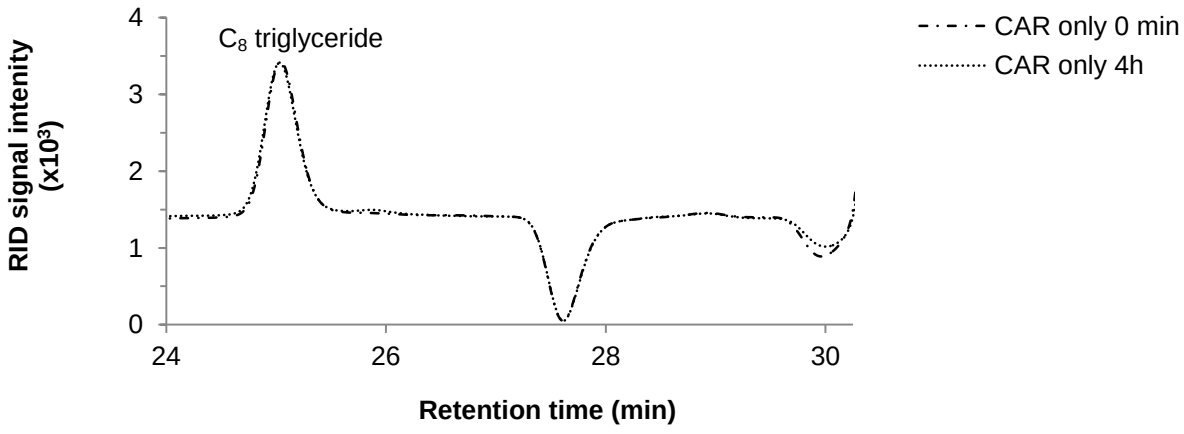
A


⇓ C_8 triglyceride
⇑ C_8 aldehyde
RID signal intensity (x10³)
Retention time (min)
Lipase & CAR 0 min
Lipase & CAR 30 min
Lipase & CAR 1 h
Lipase & CAR 2 h
Lipase & CAR 4 h

B


⇓C_8 triglyceride
⇑ C_8 fatty acid
RID signal intensity (x10³)
Retention time (min)
Lipase only 0 min
Lipase only 30 min
Lipase only 1 h
Lipase only 2 h
Lipase only 4 h

C


C_8 triglyceride
RID signal intenity (x10³)
Retention time (min)
CAR only 0 min
CAR only 4h

D

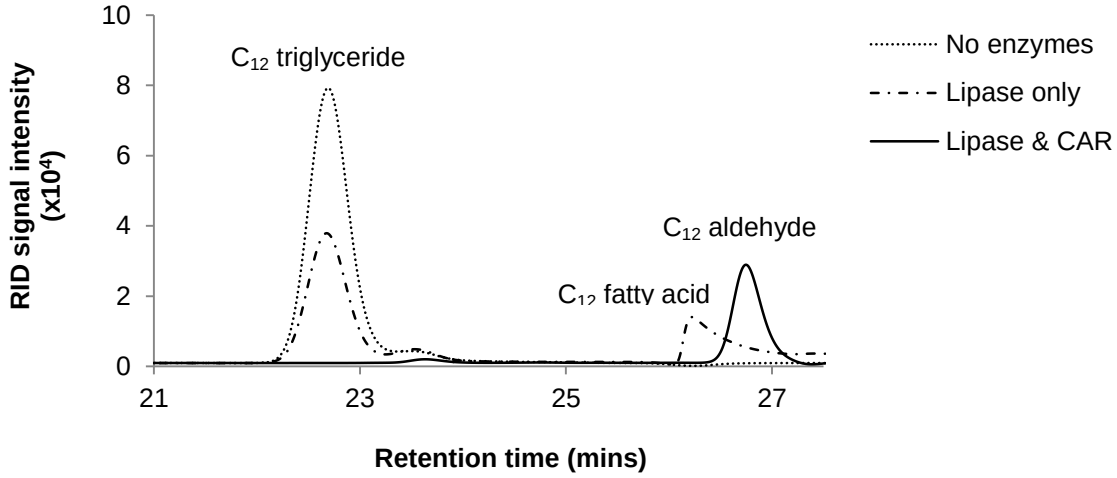


Fig. S14. Conversion of TAGs to fatty acids and aldehydes. Reactions were performed in 50 mM Tris-HCl (pH 7.5, 21^{o}C) containing 2 mM NADPH, 2 mM ATP, 10 mM $MgCl_2$ and 0.5 mM C_8-TAG (glyceryl trioctanoate) in the presence of **(*A*)** lipase and CAR enzymes **(*B*)** lipase only (0.1 mg/ml) and **(*C*)** CAR only (0.3 mg/ml). A further reaction using **(*D*)** C_{12}-TAG (glyceryl trilaurate) as the substrate was also carried out.

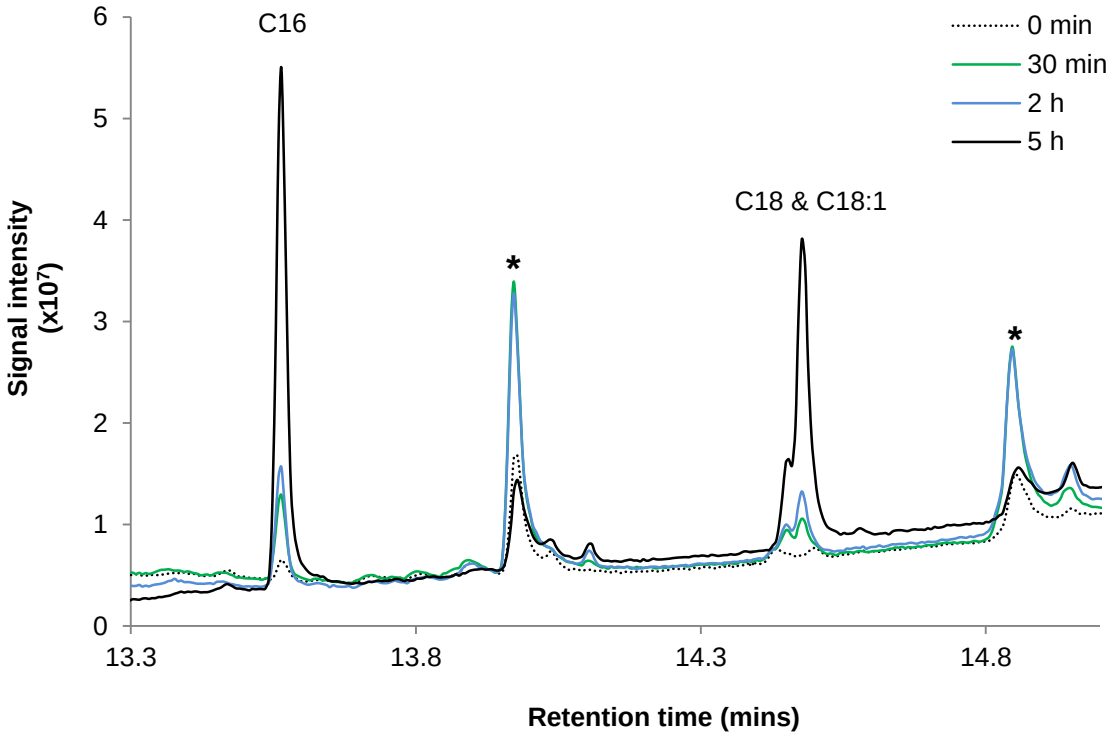


Fig. S15. *In vivo* synthesis of fatty alcohols using palm oil as substrate. Harvested cells from a preinduced culture of PC-Ahr$_{his}$, were resuspended in 50 mM potassium phosphate buffered medium, supplemented with glucose and lipase (1 mg/ml) and incubated at 30 ^{o}C with shaking at 150 rpm for up to 5 h in the presence of palm oil. Asterisk denotes fatty acid.

Table S1: Key features of the fatty acyl-ACP/TAG to fatty alcohol-producing systems. Enzyme abbreviations according to the legend of Figure 1.

Fatty alcohol producing system	Direct co-substrate requirement	Typical reaction(s) catalyzed*	ΔG (kJ mol^{-1})	*In vivo* range of fatty alcohols**						Yield of fatty alcohols ($mg\ L^{-1}$)	Reference
				C8	C10	C12	C14	C16	C18		
Lipase or TesA	ATP	C_8-TAG + 3 H_2O → 3 C_8 fatty acid + glycerol	-32.1								
+ Sfp	2 NADPH	C_{12}-ACP + H_2O → C_{12} fatty acid + ACP	-21.5								
+ CAR		C_{12} fatty acid + NADPH + H^+ + ATP → $NADP^+$ + AMP + PPi + C_{12} aldehyde	+4.4							~360	This work
+ Ahr		C_{12} aldehyde + NADPH + H^+ → C_{12} alcohol + $NADP^+$	-18.8								
		Pathway summation of ΔG	**-46.5 or-35.9**								
FatB or TesA	ATP	C_{12}-ACP + H_2O → C_{12} fatty acid + ACP	-21.5								
+ FadD	2 NADPH	C_{12} fatty acid + CoA + ATP → C_{12}-CoA + AMP + PPi	-10.5								
+ Acr1	CoA	C_{12}-CoA + NADPH + H^+ →C_{12} aldehyde + $NADP^+$ + CoA	+14.9							~60	9
		C_{12} aldehyde + NADPH + H^+ → C_{12} alcohol + $NADP^+$	-18.8								
		Pathway summation of ΔG	**-35.9**								
AAR	2 NADPH	C_{12}-ACP + NADPH + H^+ → C_{12} aldehyde + $NADP^+$ + ACP	+14.9								
+ native aldehyde reductase activity		C_{12} aldehyde + NADPH + H^+ → C_{12} alcohol + $NADP^+$	-18.8								
		Pathway summation of ΔG	**-3.9**							~140	10

*Calculation of ΔG values under standard conditions performed with eQuilibrator (4)

**Green box=detectable levels of fatty alcohols, grey box = undetected fatty alcohols

Table S2: Substrate specificity and turnovers of the metabolic enzymes required for fatty-acyl ACP to fatty alcohol conversion

Enzyme	Accession number	*In vitro* substrate specificity of metabolic enzymes*									Apparent or observed maximal K_{cat} (turnover/min)	Reference
		C6	C8	C10	C12	C14	C16	C18	C20	C22		
TesA	C6EKX0										~5 (for C18:1)	11
CAR	B2HN69										~27 (for C18:1)	This work
Ahr	C6EC82										~109 (for C12)	This work
ADC	Q7V6D4										~0.1 (for C18:1)	This work, 10
AAR	Q54765										~0.01 (for C18:1)	10
FadD	C5W549										~16 (for C18:1)	12
Acr1**	P94129										~0.003 (for C16)	13
ACR2***	A1U3L3										~3 (for C18:1) (acyl coA to aldehyde) ~550 (for C12) (aldehyde to alcohol)	14

*Green box=observed activity, blue box= undetected activity, grey box = untested substrates
**Partially purified enzyme.
***Purified ACR2 from *Marinobacter aquaeolei* VT8

Table S3. List of the primers and templates used for gene(s) amplification.

Gene(s)	PCR template	Primers employed[a]
tesA	*E. coli* BL21 (DE3)	FP: 5'- aa*ccatgg*cggacacgttattgattctgggtgatagcc -3' RP: 5'- aaaagcttttatgagtcatgatttactaaaggctgcaactgcttcg -3'
slr1192	*Synechocytis* sp. PCC 6803	FP: 5'- attaatcatatgattaaagcctacgctgccctggaag -3' RP: 5'- aacctaggtatggctgagcactacccgataatgggc -3'
ahr	*E. coli* BL21 (DE3)	FP: 5'- attaat*ccatgg*tctagataattaatggatccaggaggaaacatatgtcgatgataaaaagctatgccgcaaaag -3' RP: 5'- attaatcctaggaagcttctcgagtcaaaaatcggctttcaacaccacgcgg -3'
ahr_{his}	*E. coli* BL21 (DE3)	FP: 5'- attaat*ccatgg*gccaccaccaccaccaccacggcagccatatgtcgatgataaaaagctatgccgcaaaagaag -3' RP: 5'- attaatcctaggaagcttctcgagtcaaaaatcggctttcaacaccacgcgg -3'
adc_{his}	pET28-ADC_{his}	FP: 5'- attaat*ccatgg*cacaccaccaccaccaccaccaa -3' RP: 5'- attaat*cctagg*ttacaccatccgtgccgcagcc -3'
sfp + car_{his}	pET-TPC	FP: 5'- attaat*ccatgg*gcaagatctacggcatatacatggaccgcc -3' RP: 5'- aataat*cctagg*gaattcttacagcaggcccag -3'

[a] Restriction enzyme sites are underlined and italicized

Abbreviations used: FP, forward primer; RP, reverse primer

Table S4. Summary of the vector backbone and gene(s) inserts used for plasmid construction.

	Ligated fragments[a]	
Plasmid	**Vector backbone**	**Gene(s) insert**
pET-TPC	pET-BPC (*Nco*I/*Hind*III)	*tesA* (*Nco*I/*Hind*III)
pET-TP	pET-TPC (*EcoR*I)	n/a
pET-PC	pET-TP (*Nco*I/*EcoR*I)	*sfp* + car_{his} (*Nco*I/*EcoR*I)
pET-TesA	pET-TP (*BamH*I)	n/a
pCDF-slr1192	pCDF-Ahr (*Nde*I/*Avr*II)	*slr1192* (*Nde*I/*Avr*II)
pCDF-$slr1192_{his}$	pCDF- Ahr_{his} (*Nde*I/*Avr*II)	*slr1192* (*Nde*I/*Avr*II)
pCDF-Ahr	pCDF-Duet1 (*Nco*I/*Avr*II)	*ahr* (*Nco*I/*Avr*II)
pCDF- Ahr_{his}	pCDF-Duet1 (*Nco*I/*Avr*II)	ahr_{his} (*Nco*I/*Avr*II)
pCDF- ADC_{his}	pCDF-Duet1 (*Nco*I/*Avr*II)	adc_{his} (*Nco*I/*Avr*II)

[a] Restriction enzyme cuts are specified in brackets.

Table S5. List of the plasmids used in the study including their properties.

Plasmid	Recombinant proteins encoded	Source
pETDuet-1	None	Novagen
pCDFDuet-1	None	Novagen
pET28-ADC_{his}	ADC_{his}	2
pCDF-ADC_{his}	ADC_{his}	This study
pET-TesA	TesA	This study
pET-TP	TesA & Sfp	This study
pET-PC	Sfp & Car_{his}	This study
pET-TPC	TesA, Sfp & Car_{his}	This study
pCDF-slr1192	slr1192	This study
pCDF-$slr1192_{his}$	$slr1192_{his}$	This study
pCDF-Ahr	Ahr	This study
pCDF-Ahr_{his}	Ahr_{his}	This study

Table S6. List of *E. coli* strains engineered in the study.

Strain name	Host strain	Plasmids/genotype	Recombinant proteins induced	Source or reference
BL21(DE3)	----------	F–*ompT hsdS*^{B}(rB–mB–) *gal dcm* (DE3)	----------	Novagen
Control	BL21(DE3)	pETDuet-1, pCDFDuet-1	None	This study
slr1192	BL21(DE3)	pCDF-slr1192	slr1192	This study
slr1192$_{his}$	BL21(DE3)	pCDF-slr1192$_{his}$	slr1192$_{his}$	This study
Ahr	BL21(DE3)	pCDF-Ahr	Ahr	This study
Ahr$_{his}$	BL21(DE3)	pCDF- Ahr$_{his}$	Ahr$_{his}$	This study
TesA	BL21(DE3)	pET-TesA, pCDF-Empty	TesA	This study
ADC$_{his}$	BL21(DE3)	pET28-ADC$_{his}$	ADC$_{his}$	This study
TP	BL21(DE3)	pET-TP, pCDF-Empty	TesA & Sfp	This study
PC	BL21(DE3)	pET-PC, pCDF-Empty	Sfp & CAR$_{his}$	This study
PC-Ahr$_{his}$	BL21(DE3)	pET-PC, pCDF- Ahr$_{his}$	Sfp, CAR$_{his}$ & Ahr$_{his}$	This study
TPC	BL21(DE3)	pET-TPC, pCDF-Empty	TesA, Sfp & CAR$_{his}$	This study
TPC-slr1192	BL21(DE3)	pET-TPC, pCDF-slr1192	TesA, Sfp, CAR$_{his}$ & slr1192	This study
TPC-slr1192$_{his}$	BL21(DE3)	pET-TPC, pCDF-slr1192$_{his}$	TesA, Sfp, CAR$_{his}$ & slr1192$_{his}$	This study
TPC-Ahr	BL21(DE3)	pET-TPC, pCDF-Ahr	TesA, Sfp, CAR$_{his}$ & Ahr	This study
TPC-Ahr$_{his}$	BL21(DE3)	pET-TPC, pCDF-Ahr$_{his}$	TesA, Sfp, CAR$_{his}$ & Ahr$_{his}$	This study
TPC-ADC$_{his}$	BL21(DE3)	pET-TPC, pCDF-ADC$_{his}$	TesA, Sfp, CAR$_{his}$ & ADC$_{his}$	This study

Table S7. List of the enzymes used to engineer the metabolic pathways

Enzyme	Organism source	Enzyme encoded for	Enzyme modifications	Enzyme accession number	Reference
TesA	*E. coli*	Acyl-ACP/CoA thioesterase	Signal sequence removed	C6EKX0	15
Sfp	*B. subtilis*	Phosphopantetheinyl transferase	None	P39135	16
CAR_{his}	*M. marinum*	Carboxylic acid reductase	N-terminal 6x his tag	B2HN69	This study
slr1192	*Synechocystis* sp PCC 6803	Aldehyde reductase	None	P74721	17
$slr1192_{his}$	*Synechocystis* sp PCC 6803	Aldehyde reductase	N-terminal 6x his tag	As above	17
ADC_{his}	*P. marinus*	Aldehyde decarbonylase	N-terminal 6x his tag	NP_895059.1	2
Ahr	*E. coli*	Aldehyde reductase	None	C6EC82	This study
Ahr_{his}	*E. coli*	Aldehyde reductase	N-terminal 6x his tag	As above	This study

References

1. Akhtar MK, Jones PR (2008) Engineering of a synthetic hydF-hydE-hydG-hydA operon for biohydrogen production. *Anal Biochem* 373(1):170-172.

2. Das D, Eser BE, Han J, Sciore A, & Marsh EN (2011) Oxygen-independent decarbonylation of aldehydes by cyanobacterial aldehyde decarbonylase: a new reaction of diiron enzymes. *Angew Chem Int Ed Engl* 50(31):7148-7152.

3. Reed JL, Vo TD, Schilling CH, Palsson BO (2003) An expanded genome-scale model of *Escherichia coli* K-12 (iJR904 GSM/GPR). *Genome Biol* 4(9):R54.

4. Feist AM, *et al.* (2007) A genome-scale metabolic reconstruction for *Escherichia coli* K-12 MG1655 that accounts for 1260 ORFs and thermodynamic information. *Mol Syst Biol* 3:121.

5. Becker SA, *et al.* (2007) Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox. *Nat Protoc* 2(3):727-738.

6. Flamholz A, Noor E, Bar-Even A, Milo R (2012) eQuilibrator-the biochemical thermodynamics calculator. *Nucleic Acids Res* 40(Database issue):D770-775.

7. Laemmli, UK (1970) Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature* 227(5259):680–685.

8. Goujon M, McWilliam H, Li W, Valentin F, Squizzato S, Paern J, Lopez R (2010) A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res* 38 (Web Server issue):W695-9.

9. Steen EJ, Kang Y, Bokinsky G, Hu Z, Schirmer A, McClure A, del Cardayre SB, Keasling JD (2010) Microbial production of fatty-acid-derived fuels and chemicals from plant biomass. *Nature* 463(7280):559-562.

10. Schirmer A, Rude MA, Li X, Popova E, del Cardayre SB (2010) Microbial biosynthesis of alkanes. *Science* 329(5991):559-562.

11. Spencer AK, Greenspan AD, Cronan JE Jr (1978) Thioesterases I and II of *Escherichia coli*. Hydrolysis of native acyl-acyl carrier protein thioesters. *J Biol Chem* 253(17):5922-5926.

12. Kameda K, Nunn WD (1981) Purification and characterization of acyl coenzyme A synthetase from *Escherichia coli*. *J Biol Chem* 256(11):5702-5707.

13. Reiser S, Somerville C (1997) Isolation of mutants of *Acinetobacter calcoaceticus* deficient in wax ester synthesis and complementation of one mutation with a gene encoding a fatty acyl coenzyme A reductase. *J Bacteriol* 179(9):2969-2975.

14. Willis RM, Wahlen BD, Seefeldt LC, Barney BM (2011) Characterization of a fatty acyl-CoA reductase from *Marinobacter aquaeolei* VT8: a bacterial enzyme catalyzing the reduction of fatty acyl-CoA to fatty alcohol. *Biochemistry* 50(48):10550-8.

15. Cho H, Cronan JE Jr (1995) Defective export of a periplasmic enzyme disrupts regulation of fatty acid synthesis. *J Biol Chem* 270(9):4216-4219.

16. Quadri LE, Weinreb PH, Lei M, Nakano MM, Zuber P, Walsh CT (1998) Characterization of Sfp, a *Bacillus subtilis* phosphopantetheinyl transferase for peptidyl carrier protein domains in peptide synthetases. *Biochemistry* 37(6):1585-1595.

17. Vidal R, López-Maury L, Guerrero MG, Florencio FJ (2009) Characterization of an alcohol dehydrogenase from the *Cyanobacterium Synechocystis* sp. strain PCC 6803 that responds to environmental stress conditions via the Hik34-Rre1 two-component system. *J Bacteriol* 191(13):4383-4391.