

***In Vitro* Studies of Maleidride-Forming Enzymes**

Sen Yin,^a Steffen Friedrich,^a Vjaceslavs Hrupins^a and Russell J. Cox^{*a}

Institute for Organic Chemistry and BMWZ, Leibniz University of Hannover

Electronic Supplementary Information

1.	Gene Cloning, Expression and Protein Purification	2
2.	Enzyme Assays	12
3.	Synthetic Chemistry	26
4.	Assignment of Stereochemistry	32
5.	LCMS Analysis and Purification	34
6.	Preparation and Purification of 1	35
7.	Protein Modelling	37
8.	References	41

1. Gene Cloning, Expression and Protein Purification

1.1 General Methods

Name	Insert Fragments	Selective Marker	ID	Source
pRSETa-bfL1	bfL1	Amp	SYv20	Commercial codon-optimised expression systems
pRSETa-bfL2	bfL2	Amp	SYv21	
pRSETa-bfL3	bfL3	Amp	SYv22	
pRSETa-pvL2	pvL2	Amp	SYv23	
pRSETa-bfL5	bfL5	Amp	SYv24	
pRSETa-bfL6	bfL6	Amp	SYv25	
pRSETa-bfL9	bfL9	Amp	SYv26	
pRSETa-bfL10	bfL10	Amp	SYv27	
pET28a-bfPKS	bfPKS ACP	Amp	SH	
pET28a-bfL5	bfL5	Kan	SYv28	Manual ligation into pET28a
pET28a-bfL6	bfL6	Kan	SYv29	
pET28a-bfL9	bfL9	Kan	SYv30	
pET28a-bfL10	bfL10	Kan	SYv31	
pET28a-bfL5(Δ 18)	bfL5(Δ 1-18)	Kan	SYv32	
pET28a-bfL6(Δ 22)	bfL6(Δ 1-22)	Kan	SYv33	
pET28a-bfL10(Δ 20)	bfL10(Δ 1-20)	Kan	SYv34	Infusion cloning into pESC-URA
pESC-URA-bfL3	bfL3	Amp, Δ Ura	SYv35	
pESC-URA-pvL2	pvL2	Amp, Δ Ura	SYv36	
pESC-URA-bfL5	bfL5	Amp, Δ Ura	SYv37	
pESC-URA-bfL6	bfL6	Amp, Δ Ura	SYv38	
pESC-URA-bfL9	bfL9	Amp, Δ Ura	SYv39	
pESC-URA-bfL10	bfL10	Amp, Δ Ura	SYv40	

Table S1.1A Vectors constructed

Primer	Template	Direction	sequence 5' - 3'	Amplification	ID
28a-5F	pRSETa-bfL5	fwd	ATGGGTCGCGGATCCGAATTCATGATTGGCTATATTGTG	F28a5	SYp301
28a-5R	pRSETa-bfL5	rev	GCTCGAGTGC GGCCGAAGCTTTCAGGTGCTGCTCAGAAA	F28a5, F28asig5	SYp302
28a-6F	pRSETa-bfL6	fwd	ATGGGTCGCGGATCCGAATTCATGGTGAGCAGCAAAGTG	F28a6	SYp303
28a-6R	pRSETa-bfL6	rev	GCTCGAGTGC GGCCGAAGCTTCTAATAATGGCCATATTT	F28a6, F28asig6	SYp304
28a-9F	pRSETa-bfL9	fwd	ATGGGTCGCGGATCCGAATTCATGAGCATTTCTGGCGTAT	F28a9	SYp305
28a-9R	pRSETa-bfL9	rev	GCTCGAGTGC GGCCGAAGCTTTTAGCGGGCCACGGGCG	F28a9	SYp306
28a-10F	pRSETa-bfL10	fwd	ATGGGTCGCGGATCCGAATTCATGGTGGCGATGTTTATT	F28a10	SYp307
28a-10R	pRSETa-bfL10	rev	GCTCGAGTGC GGCCGAAGCTTTCATCAATATCCGGCGG	F28a10, F28asig10	SYp308
28asignal-5F	pRSETa-bfL5	fwd	ATGGGTCGCGGATCCGAATTCATGCAGACCCCGCGGCTAT	F28asig5	SYp309
28asignal-6F	pRSETa-bfL6	fwd	ATGGGTCGCGGATCCGAATTCATGCAGCATTTCTGCCGGAAT	F28asig6	SYp310
28asignal-10F	pRSETa-bfL10	fwd	ATGGGTCGCGGATCCGAATTCATGCATTAGCCTGGCGCGCAA	F28asig10	SYp311
urabfL3F	pRSETa-bfL3	fwd	GGAGAAAAAACC CGGATCCATGTCTACCGAAAAATACGAC	Fyeast3	SYp312
urabfL3R	pRSETa-bfL3	rev	GCTAGCCGCGGTACCAAGCTTCTACAGCGCCAGGG	Fyeast3	SYp313
urapvL2F	pRSETa-pvL2	fwd	GGAGAAAAAACC CGGATCCATGCAGGCGTATGATCA	Fyeastpv2	SYp314
urapvL2R	pRSETa-pvL2	rev	GCTAGCCGCGGTACCAAGCTTTTACAGTTTCACTTCGCTG	Fyeastpv2	SYp315
urabfL5F	pRSETa-bfL5	fwd	GGAGAAAAAACC CGGATCCATGATTGGCTATATTGTG	Fyeast5	SYp316
urabfL5R	pRSETa-bfL5	rev	GCTAGCCGCGGTACCAAGCTTTCAGGTGCTGCTCAGAAAT	Fyeast5	SYp317
urabfL6F	pRSETa-bfL6	fwd	GGAGAAAAAACC CGGATCCATGGTGGCAGCAAAGTG	Fyeast6	SYp318
urabfL6R	pRSETa-bfL6	rev	GCTAGCCGCGGTACCAAGCTTCTAATAATGGCCATATTTGCGC	Fyeast6	SYp319
urabfL9F	pRSETa-bfL9	fwd	GGAGAAAAAACC CGGATCCATGAGCATTTCTGGCGTAT	Fyeast9	SYp320
urabfL9R	pRSETa-bfL9	rev	GCTAGCCGCGGTACCAAGCTTTAGCGGGCCACGGGCG	Fyeast9	SYp321
urabfL10F	pRSETa-bfL10	fwd	GGAGAAAAAACC CGGATCCATGGTGGCGATGTTTATT	Fyeast10	SYp322
urabfL10R	pRSETa-bfL10	rev	GCTAGCCGCGGTACCAAGCTTTCATCAATATCCGGCGG	Fyeast10	SYp323

Table S1.1B PCR primers used

Isolation of Small Scale gDNA and Isolation of Small Scale DNA Plasmid from E. coli

The GenElute Plant Genomic DNA miniprep kit (Sigma) was used to extract gDNA from fungi. Isolation of gDNA was carried out according to the manufacturer's protocol. GeneJET Plasmid Miniprep Kit was used to isolate small scale DNA plasmid from *E. coli*. Isolation was performed based on manufacturer's protocol.

Isolation of DNA Plasmid from S. cerevisiae

The Zymoprep II yeast plasmid miniprep kit (Zymo research) was used to extract plasmid DNA from *S. cerevisiae*. Yeast cell walls were digested using the Zymolyase enzyme according to the manufacturer's protocol. Primary transformation plates containing multiple colonies were used as the source of yeast cells for plasmid extraction

Analytical or Preparative PCR

Analytical PCR was performed using reagent KOD Hot Start DNA polymerase throughout (Novagen). Reagent composition for 25 μ L reaction was as follow 0.5 μ L of KOD Hot Star DNA polymerase (Novagen), 2.5 μ L of enzyme buffer, 2.5 μ L of dNTPs, 1.5 μ L of $MgSO_4$, 2.5 μ L of each primer 3 μ M, 0.5 μ L of DMSO and distilled water until 25 μ L in total. The DNA template was added at 1 pg – 10 ng (low complexity) and 50 - 250 ng (high complexity). The PCR condition was initially denatured by heating at 95 $^{\circ}C$ for 2 min, followed by 35 cycles of amplification, *i.e.* denaturation at 95 $^{\circ}C$ for 30 s, 66 $^{\circ}C$ for 20 s (depend on the best annealing and long product), 70 $^{\circ}C$ for 15 s and final extension 70 $^{\circ}C$ for 10 min.

Colony PCR or Analytical PCR

Colonies of *E. coli* transformants were screened by PCR. Cells of *E. coli* were picked by sterilised toothpick and resuspended into PCR reaction. PCR reaction consists of 5 μ L of Biomix, 0.5 μ L of each reverse and forward primer, and 4 μ L of deionized water on in 10 μ L volume total. PCR program were performed as follow: 95 $^{\circ}C$ for 5 min then 35 cycles with 95 $^{\circ}C$ for 15 s, 55 $^{\circ}C$ for 30 s, 72 $^{\circ}C$ for 15-20 s/kb, finally 72 $^{\circ}C$ for 10 min.

Restriction Analysis of Plasmid DNA

Approximately 2 μ L of plasmid DNA (0.1- 0.2 μ g) was added 1 μ L of enzyme buffer, 0.25 μ L of certain restriction enzyme and sterilized water until final volume of 10 μ L. The mixture was incubated at 37 $^{\circ}C$ for an hour. DNA electrophoresis was used to check the digested DNA.

DNA Electrophoresis

Agarose gels were prepared in TAE buffer with agarose at concentrations between 0.8 – 1.5% (depend on fragment size) and with Midori green advance added for DNA visualisation. DNA was loaded on gels either using loading dye already present within a reaction (FastDigest Green Buffer (Fermentas)) or after addition of 0.1 vol 10 x loading buffer (0.25% (w/v) bromophenol blue in 10% (v/v) glycerol). Electrophoresis was performed using TAE buffer (0.04 M Tris-acetate, 0.001 M EDTA) and run at 200 V and 100 mA for 20 min. Gels were visualised by UV light using a transilluminator.

1.2 Cloning procedures

Seven genes from the byssochlamic acid pathway [the hydrolase (*bfl1*); citrate synthase (*bfl2*); citrate dehydratase (*bfl3*); two PEBPs (*bfl5* and *bfl9*) and two KSIs (*bfl6* and *bfl10*)] and one gene from the cornexstin pathway [2-methyl citrate dehydratase, *pvL2*] were obtained as codon optimised genes in the vector pRSET A (ThermoFisher).

Restriction enzyme digestion

Restriction enzyme digestions (single or double) of plasmids or PCR products were performed following the manufacturer's guidelines of the enzymes (New England Biolabs or Thermo Scientific). DNA sample were mixed with the enzyme or enzyme mixture and the appropriate buffer. The reaction mixture was incubated at 25 °C or 37 °C (depending on the restriction enzyme) for 1 to 12 h. Enzymes were inactivated by heating the sample at 65 °C or 80 °C for 20 min. Digested DNA or plasmid were purified using the NucleoSpin® Gel and PCR Clean-up kit (Macherey–Nagel) as described previously.

Dephosphorylation and ligation

Digested plasmids were dephosphorylated using shrimp alkaline phosphatase (SAP) to prevent self ligation. 1 µl of enzyme was added after digestion to the reaction mixture and incubated at 37 °C for 1 h. T4 DNA ligase (New England Biolabs) and its appropriate buffer were used for ligation reaction (final volume 10 µl) following the manufacturer's instructions. Molar ratio of 1:5 plasmid to insert was used. Ligation reactions were incubated at 4 °C overnight or at room temperature for 30 min followed by heat inactivation at 65 °C.

1.3 Heterologous Protein Production and Purification

Each constructed plasmids was transformed into chemically competent *E. coli* BL21 (DE3) cells *via* heat shock transformation. A single colony was picked to inoculate an overnight culture (LB medium [5 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract] with kanamycin) which was used to inoculate kanamycin containing 2TY medium (Tryptone 16g/L, yeast extract 10g/L, NaCl 5g/L). Cultures were grown to an OD₆₀₀ of 0.4 - 0.6 at 37 °C before protein production was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cultures were further incubated at 16 or 25 °C for 20 h to increase the solubility of the produced protein. Cells were harvested by centrifugation (5000 x g, 4 °C, 15 min).

The harvested cell pellets were resuspended in 10 mL cold IMAC loading buffer (50 mM Tris-HCl, pH 8, 150 mM NaCl, 10 mM Imidazole). Lysis of the cells was done by sonication using a SonoPlus Ultrasonic homogenizer from Bandelin. The cells were intermittently sonicated on ice for

30 s with 30 s allowed for cooling. The total sonication time was 7 min. The cell lysate was centrifuged at 12000 rpm for 45 min at 4 °C. Samples of the supernatant and the pellet were used for SDS PAGE. The protein of interest was purified using immobilized metal ion affinity chromatography. During the whole time the cell samples were kept on ice.

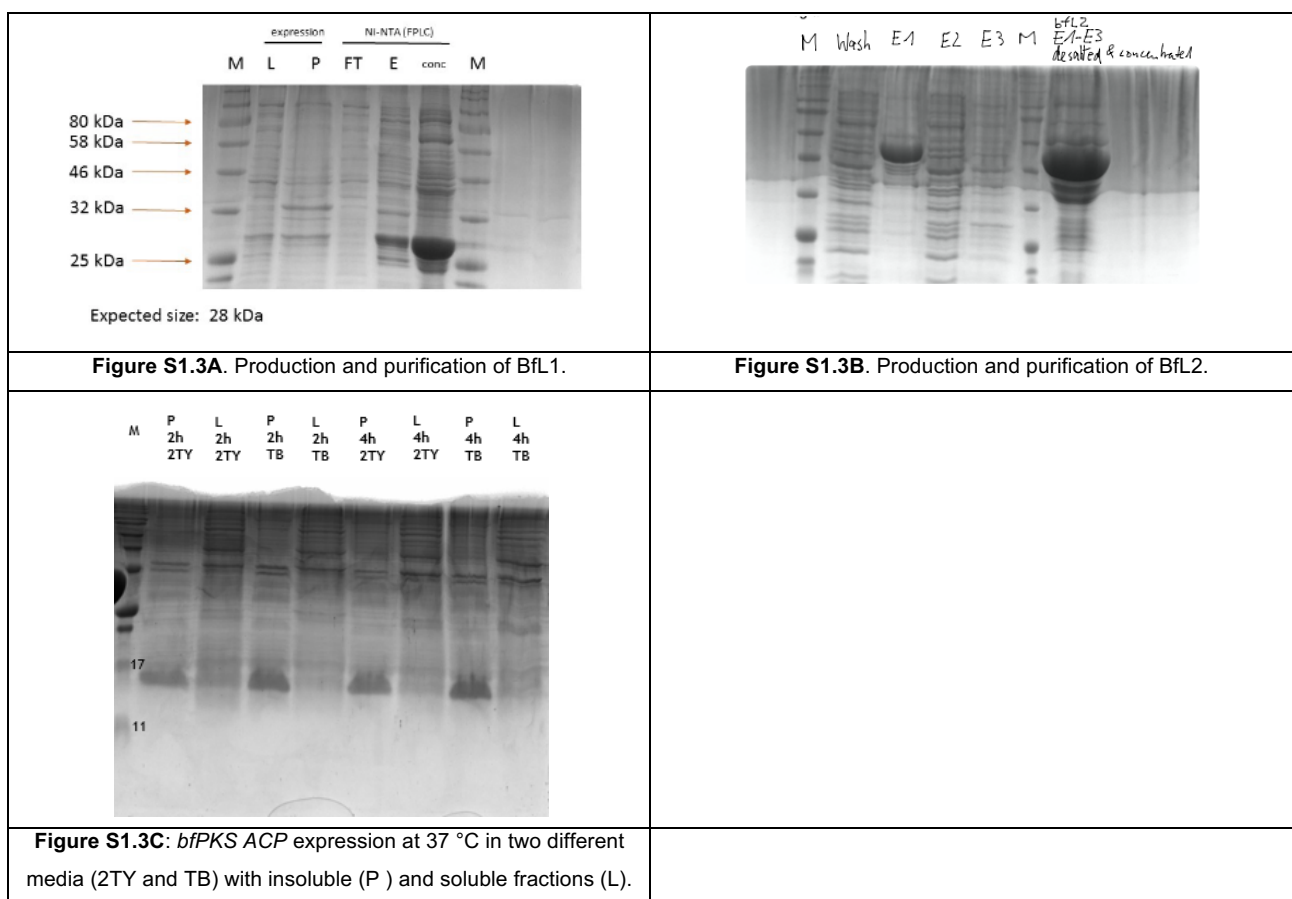
Purification of histidine-tagged recombinant proteins were done by immobilized metal ion affinity chromatography using Ni-NTA agarose (Qiagen). All buffers, washing solutions and samples were filtered through a 0.45 µm filter prior to usage. The crude lysate was passed through a Ni-NTA column (2 mL bed volume). Firstly, the column was washed with 10 mL loading buffer. Elution was performed with a stepwise gradient of imidazole (20 - 500 mM). 2 mL fractions were collected over the course of elution and analysed using SDS–PAGE. The eluate was concentrated and the binding buffer was replaced by storage buffer (50 mM potassium phosphate buffer, pH 8.0) using amicon ultra centrifugal filter units, cut off 10,000 MW, from Merck. The purified enzyme was stored at -20 °C or immediately used for enzyme assays.

bfl1 + Bfl2

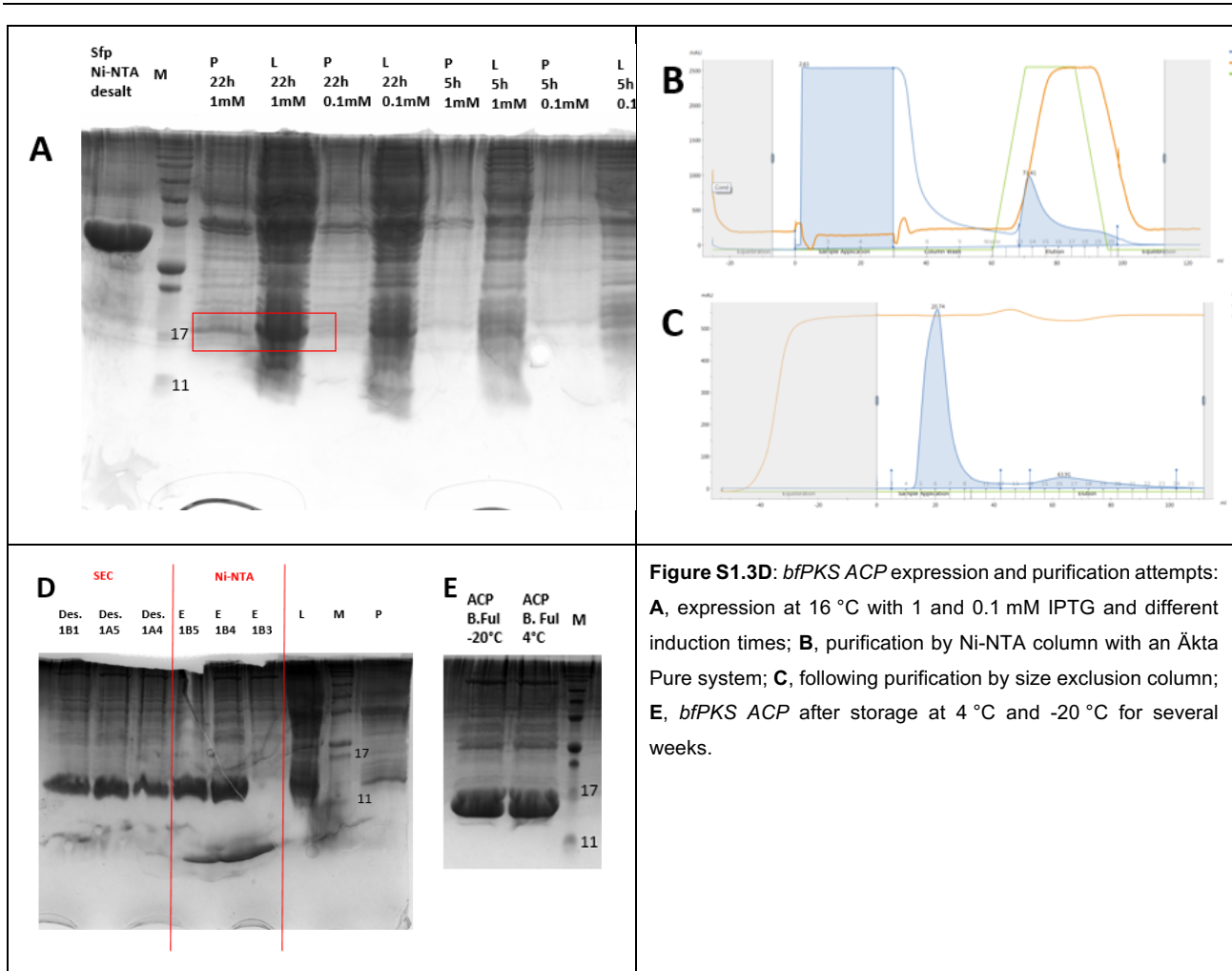
Synthetic codon-optimised *bfl1*_pRSET-A was transformed into *E. coli* BL21(DE3). A single colony was picked into 1 mL LB medium containing carbenicillin and incubated with shaking at 37 °C overnight. 100 mL LB-medium (containing carbenicillin) was inoculated with 1 mL overnight culture and incubated at 37 °C for 2 h, then cooled to 16 °C and induced with 0.1 mM IPTG. Incubated at 16 °C, 200 rpm for 16 h. Cells were collected by centrifugation and lysed in 40 mL 50 mM Naphosphate (+ 200 mM NaCl, 20% glycerol, pH 7.5 for 10 min on ice. Solids were removed by centrifugation and the soluble protein fraction purified using HisTrap 5 mL FF column (GE Life Sciences) under manufacturers instructions. *bfl2* was expressed and purified using the same procedures.

bfaCP Preparation and Acylation

The domain boundaries for the *B. fulva* PKS (*bfpKS*) ACP were determined as S2511 to E2618. A pET28a(+) vector including the desired gene (*NdeI* and *XhoI* restriction sites included) was synthesized by Baseclear (*E. coli* codon optimized) and transformed into *E. coli* BL21(DE3). First, small scale expression was performed at 37 °C for 2 and 4 h in 2TY and TB medium with 1 mM IPTG. Cell lysis (sonication) was followed by separation of the soluble and insoluble parts of the culture. SDS-PAGE analysis showed most of the protein in the insoluble fraction (Figure S1.3C).



Expression was then attempted at 16 °C for 22 h with 1 mM IPTG. Under these conditions a band in the soluble fraction at the expected size of the protein was observed (14 kDa, Figure S1.3D). In the next step a large scale expression (800 mL culture) was conducted under the same conditions. The cells were harvested, lysed and centrifuged as for BfL1 and BfL2. The soluble part was purified by Ni-NTA column (Figure S1.3D) using an Äkta Pure FPLC. Fractions of the Ni-NTA run containing the desired protein were combined and used for the size exclusion chromatography (Figure S1.3D). Combined fractions of the SEC run were concentrated and analysed by SDS-PAGE. Fractions 1A4 to 1B1 contained the expected protein (Figure S1.3D). The purified *bflPKS ACP* protein was stored at 4 °C and also at -20 °C without precipitations. -20 °C samples were used for more than 6 months without loss of activity (Figure S1.3D).



Holo-Bfpks ACP was synthesised by treatment of *apo-ACP* with Sfp. 50 mM Tris buffer (pH 8.8), 10 mM MgCl₂, 1 mg/ml *apo-ACP*, 0.01 mg/ml Sfp and 1 mM CoA were incubated for 30 min at 30 °C. The protein was analysed by mass spectrometry. For the loading reaction of Hexanoyl-CoA and Hex-2-enoyl CoA to his₆-*bfpks ACP*, the following components were mixed: 50 mM Tris buffer (pH 8.8), 10 mM MgCl₂, 0.01 mg/mL Sfp, 1.25 mg/ml *apo bfpks ACP* and 25 μM acyl-CoA. The mixture was incubated for 30 min at 30 °C. The protein was analysed by mass spectrometry.

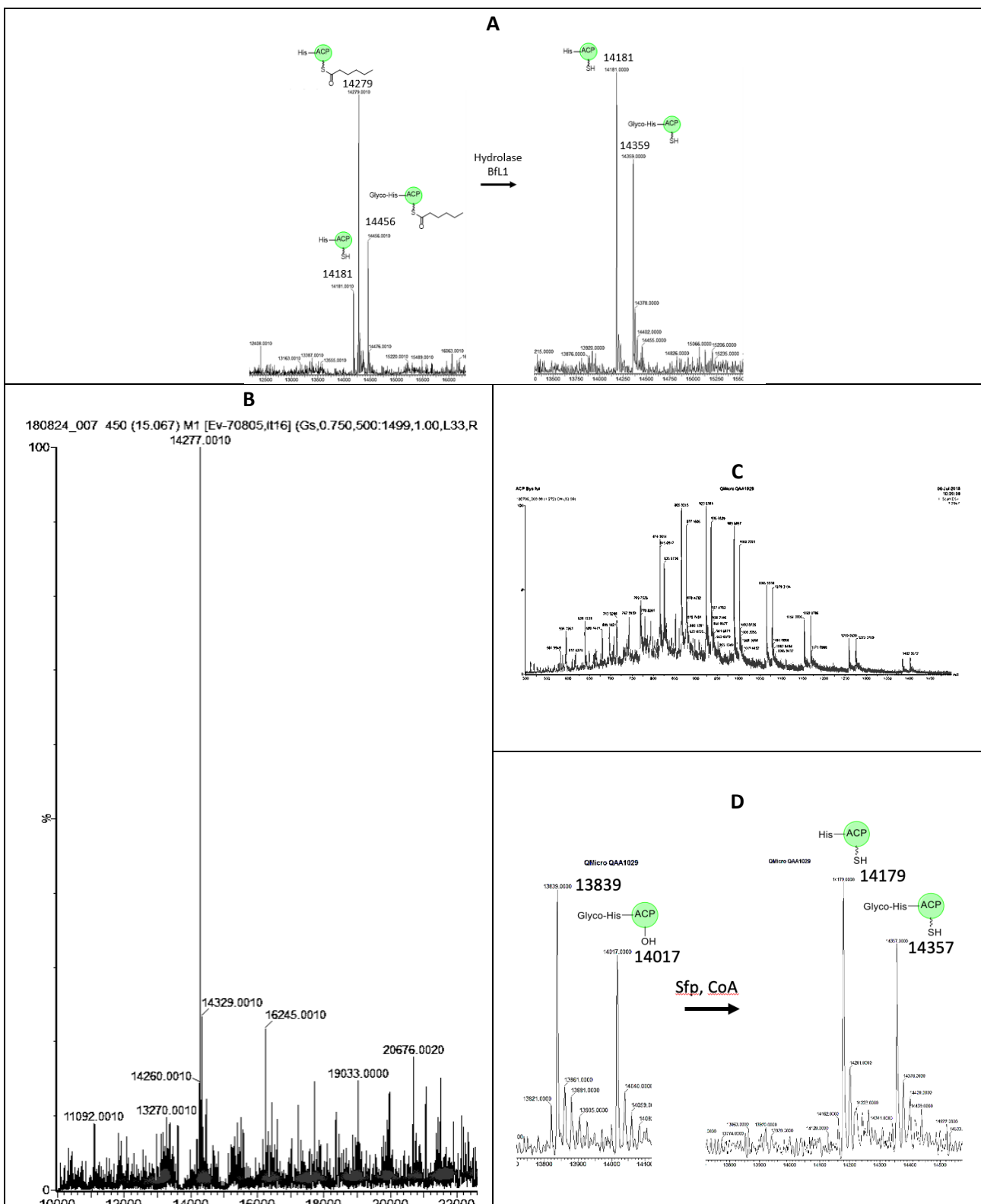


Figure S1.3E. **A**, Hexanoyl bfpKS ACP **15a** created with Sfp, apo-ACP and hexanoyl CoA **15d** and its conversion to holo-ACP by Bfl1 Hydrolase. Note the presence of a glycosylated species sometimes formed during fermentation. Cleavage of the his-tag also removes the glycosyl unit; **B**, Hexanoyl bfpKS ACP **14a** created with Sfp, apo-ACP and hexanoyl CoA **14d**; **C** Mixture of apo-his-bfACP and glyco-apo-his₆-bfACP raw MS data; **D**, synthesis of holo-bfPKS ACP by incubation of apo-ACP (mixture of glyco-forms) with CoASH and Sfp.

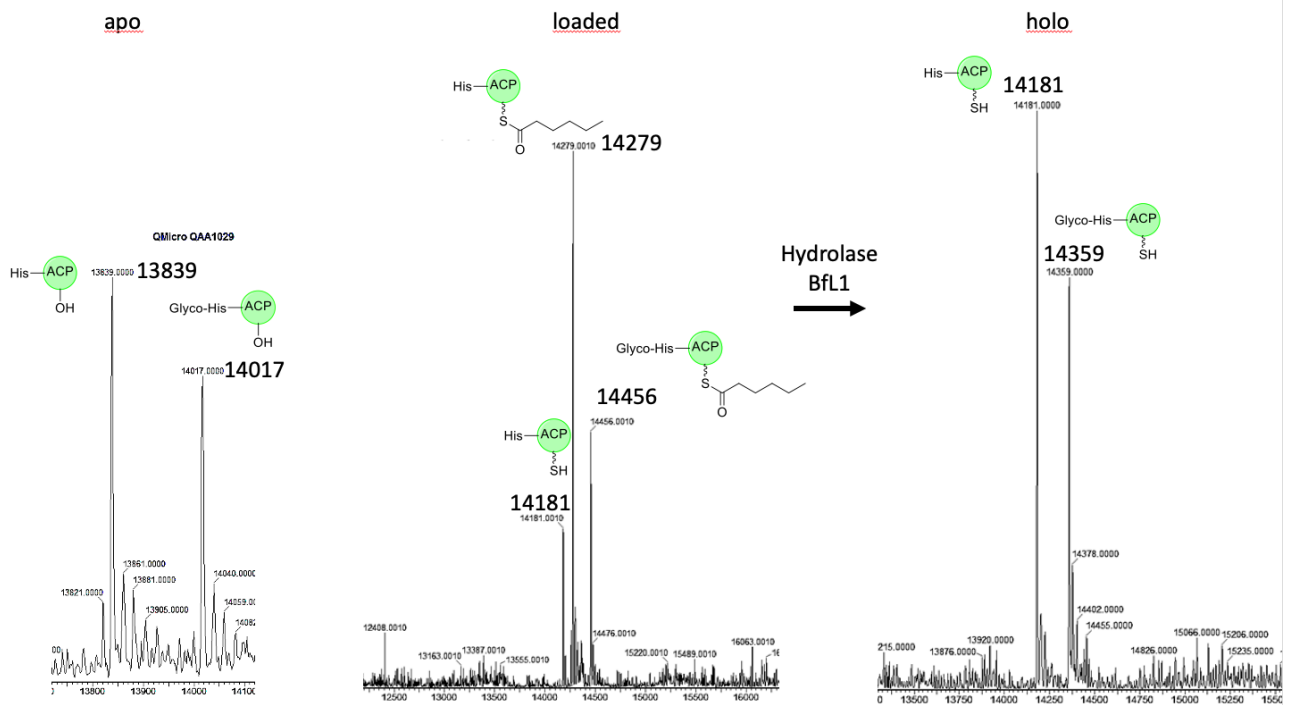


Figure S1.3X. Conversion of *apo*-BfACP to hexanoyl BfACP and then to *holo*-BfACP (mixture of glyco-forms).

Bfl3, *PvL2*, *Bfl5* *Bfl6* and *Bfl10* - yeast expression

Genes *bfl3*, *pvL2* and *bfl5*, 6 and 10 were cloned into multiple cloning site 2 of pESC-URA between *Bam*H I and *Hind* III restriction sites (**Fig. S1.3F**). The In-fusion cloning method was used for the construction of the plasmid. The whole ORF fragments of *bfl3*, *pvL2* and *bfl5*, 6 and 10 were cloned into pESC-URA and confirmed by colony PCR. Plasmid DNA was extracted from positive colonies for yeast transformation.

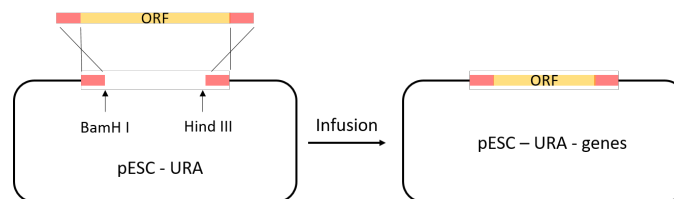


Fig. S1.3F Yeast expression plasmid construction.

Induction and Expression

Individual pESC-URA plasmids containing the genes (*bfl3*, *pvL2* and *bfl5*, 6 and 10) and the empty vector were transformed into yeast using a LiOAc procedure. The transformation mixture was plated on SM-URA agar plates (0.17 % (w/v) Yeast nitrogen base, 0.5 % (w/v) (NH₄)₂SO₄, 2 % (w/v) D(+)-Glucose Monohydrate, 0.077 % (w/v) Complete supplement mixture minus Uracil, 1.5 % (w/v) Agar). Eight colonies from each transformation were selected for further growth and PCR confirmation.

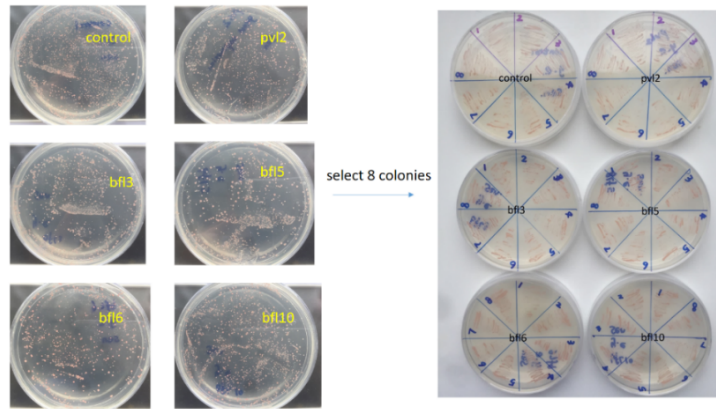


Fig. S1.3G Yeast transformation results of pESC-URA-*bfl3*, *pvL2* and *bfl5,6,10*.

The yeast cells were collected from the second plates and tested by colony PCR. The positive yeast colonies went into the next inducing and expression step. Because the expression vector contains the *S. cerevisiae GAL1* promoter, it can be induced by galactose. By using media lacking uracil and containing galactose, the yeast was induced for 16 hours at 30 °C. Glass beads were then used to break the cells. The crude cell extract was then checked by SDS-PAGE.

PvL2 and Bfl3 have some weak but visible expression on SDS-PAGE compared to the empty vector transformant. The molecular weight of PvL2 and Bfl3 are 55 kDa, and this was consistent with the SDS-PAGE results (**Fig. S1.3H**). However, Bfl5, 6 and 10 had no visible protein on the SDS-PAGE. Bfl5, 6 and 10 could directly be used for the *in vitro* assay.

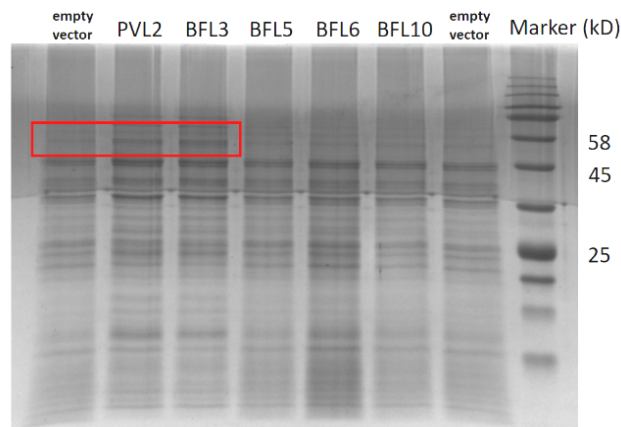


Fig. S1.3H SDS-PAGE results of yeast expression of *pvL2/ bfl3* and *bfl5, 6 and 10*. The control is empty vector transformant. The red frame highlights the possible expression.

1.4 Yeast transformation and expression

Yeast transformation

Transformation of *S. cerevisiae* was conducted based on lithium acetate/single-stranded carrier DNA/PEG protocol. A single colony of *S. cerevisiae* was inoculated into a 10 mL YPAD (1 % (w/v) Yeast extract, 2 % (w/v) Tryptone, 2 % (w/v) D(+)-Glucose Monohydrate, 0.03 % (w/v) Adenine)

starter culture and grown overnight at 28 °C with shaking at 200 rpm. The starter culture was then added to 40 mL of YPAD in a 250 ml flask and incubated at 28 °C with shaking at 200 rpm for 5 hours, after which the culture was centrifuged at 3000 × g for 5 min and the supernatant discarded. The cells were washed with 25 ml sterile H₂O and the centrifugation repeated; the pellet was then resuspended in 1 mL 0.1 M LiOAc and transferred to a 1.5 ml microfuge tube. The cells were then pelleted at 14500 × g for 15 sec and the supernatant discarded, after which the cells resuspended with 400 μL 0.1 M LiOAc and transferred 50 μL to a new 1.5 mL microfuge then pelleted again at 14500 × g for 15 sec and the supernatant discarded. Then, it was added 240 μL of PEG solution (50% (w/v) PEG 3350), 36 μL 1 M LiOAc, 50 μL SS-DNA (denaturated salmon testis DNA, 2 mg/mL in TE buffer) and up to 34 μL of DNA then mixed by vortexing and incubated at 30 °C for 30 min and 42 °C for 30 min. The cells were pelleted at 6500 rpm for 15 sec then gently resuspended in 1 mL of sterilized H₂O. For each transformant 200 μL aliquots were spread on SM-URA plates and incubated at 28 °C for 3-4 days until colonies appeared.

Yeast expression

pESC-Ura vector was transformed to w303B *S. cerevisiae*. The transformation reaction onto two SM-URA plates with dextrose (D-glucose) as the carbon source. Incubate the plates at 30 °C for 2-3 days. When the colony showed on the plates, 8 colonies were picked for yeast colony PCR. The positive colony was then streaked out on an SM-URA plate, and incubate at 30 °C overnight. The transformed colony was then inoculated into 50 mL of glucose-ura medium (0.67% (w/v) yeast nitrogen base, 0.13% (w/v) of complet supplement mixture minus uracil, 2% (w/v) glucose monohydrate) at 30 °C overnight. The yeast cells were harvested and inoculated into 100 mL of galactose-ure medium (0.67% (w/v) yeast nitrogen base, 0.13% (w/v) of complet supplement mixture minus uracil, 2% (w/v) galactose monohydrate) for 16 h at 30 °C. The cells were harvested by centrifugation at 1000 × g for 5 min. The cells were ready for further biochemical analysis.

2. Enzyme Assays

2.1 Hydrolase

small molecules

Assays were conducted in 100 μ L total volume of sodium phosphate buffer (pH 7.0, 50 mM) plus glycerol (10%) containing substrate (6 mM) and BfL1 enzyme (*ca* 0.5 mg/mL) at 30 °C for 2h. CH₃CN (100 μ L) was added and protein removed by centrifugation. The assay mixture (20 μ L) was examined directly by LCMS.

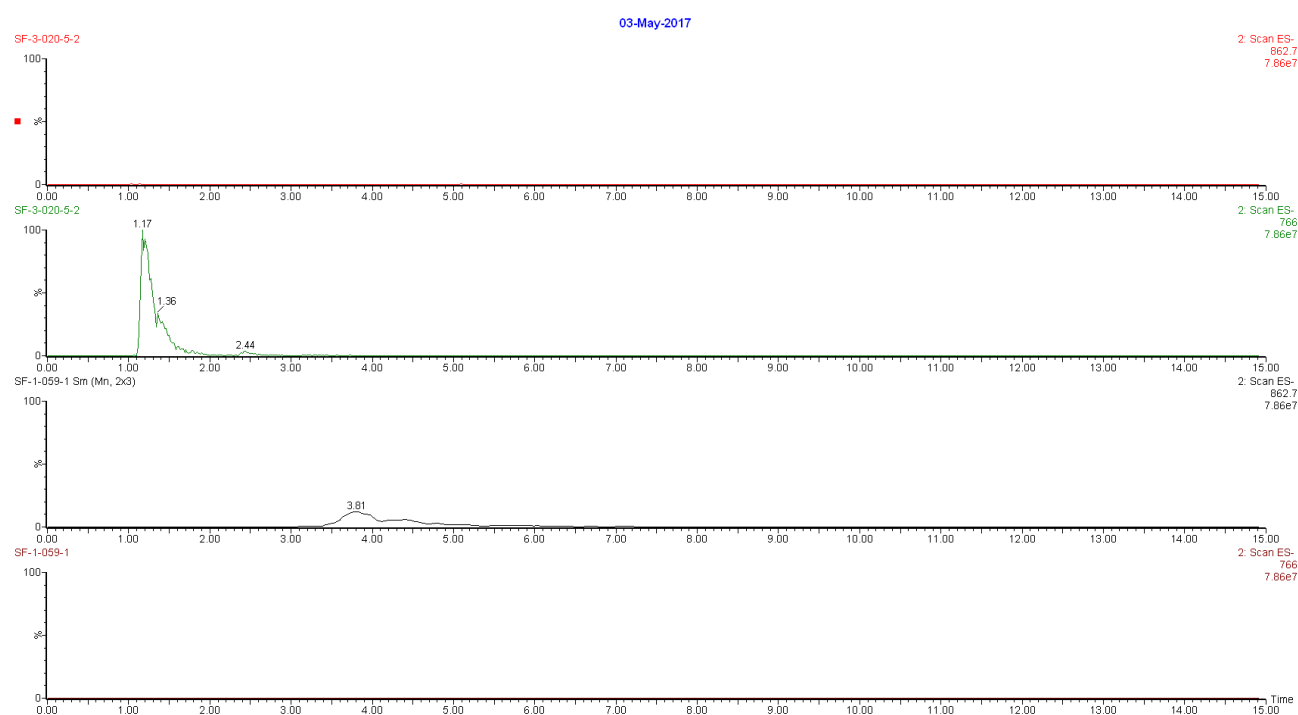


Figure S2.1A. Typical Hydrolase data, extracted ion chromatograms, ES⁻. From top: *m/z* 867.2 corresponding to hexenoyl CoA **14d** after hydrolase treatment; *m/z* 766 corresponding to CoASH after hydrolase treatment; *m/z* 867.2 corresponding to hexenoyl CoA **14d** before hydrolase treatment; *m/z* 766 corresponding to CoASH before hydrolase treatment.

Acyl ACPs

See section 1.3 for details of production and hydrolysis of acyl-BfACP.

2.2 Citrate Synthase

1 mM hexaketide substrates (**14a-d** and **15a-d**) and 1 mM oxaloacetate were incubated with purified CS protein BfL2 at 30 °C for 2 hours in 100 μ L PBS buffer. The reaction was stopped by addition of 100 μ L acetonitrile. Protein was removed by centrifugation and the supernatant was analysed by direct LCMS. According to the LCMS analysis of the reactions, BfL2 only showed activity on 2 – hexan/hexen-oyl – CoA, and formed the expected citrate product (**Fig. S2.2A** and **Fig. S2.2B**). Hexanoyl – CoA **15d** has *m/z* [M - H]⁻ 864. After the BfL2 reaction, a new compound was observed

by ELSD. The new peak had m/z $[M - H]^-$ 247 which was consistent with the citrate product **30a**. The boiled enzyme control showed no activity on the substrate (**Fig. S2.2A**).

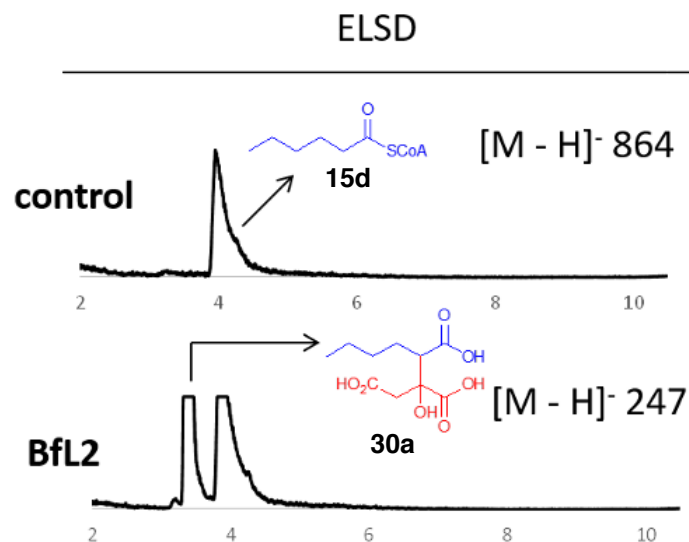


Fig. S2.2A LCMS analysis of BfL2 reaction on hexanoyl – CoA substrate.

Hexenoil – CoA **14d** was also used as substrate for BfL2 reaction. Hexenoil – CoA **14d7** ($[M - H]^-$ 862) was catalyzed by BfL2 to form a new peak which has m/z $[M - H]^-$ 245. The molecular weight of this compound was consistent with the citrate product **16**. The boiled enzyme control showed no activity on the substrate (**Fig. S2.2B**).

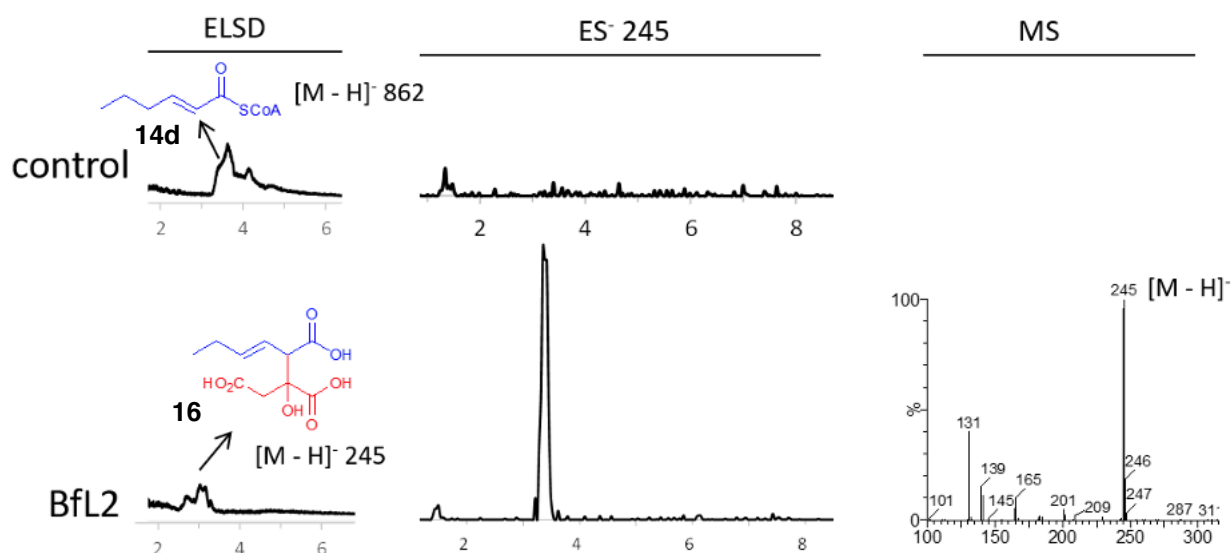
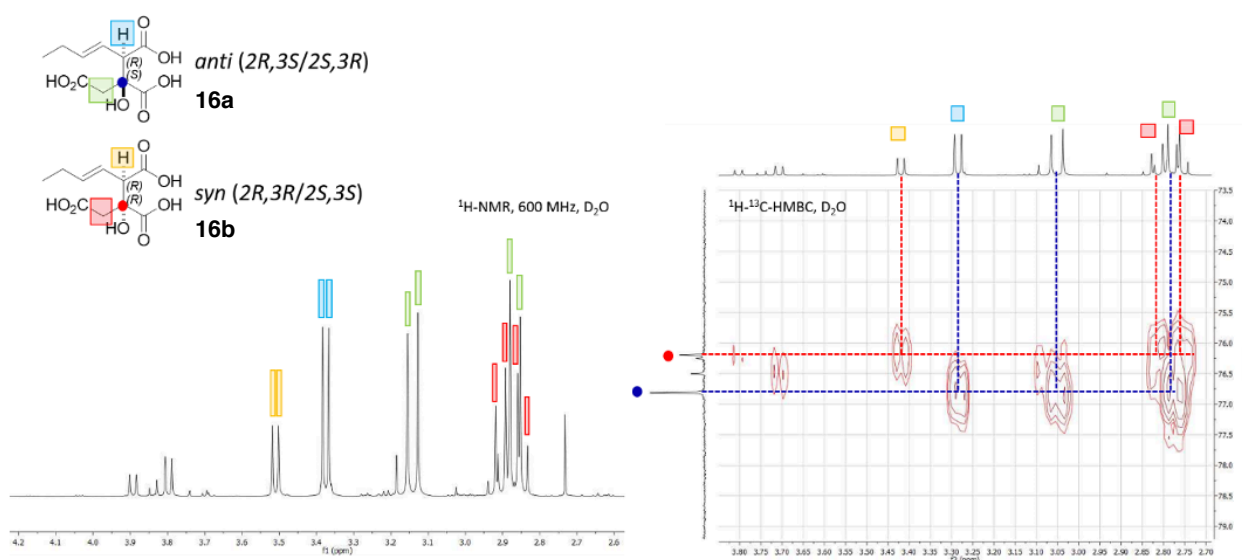
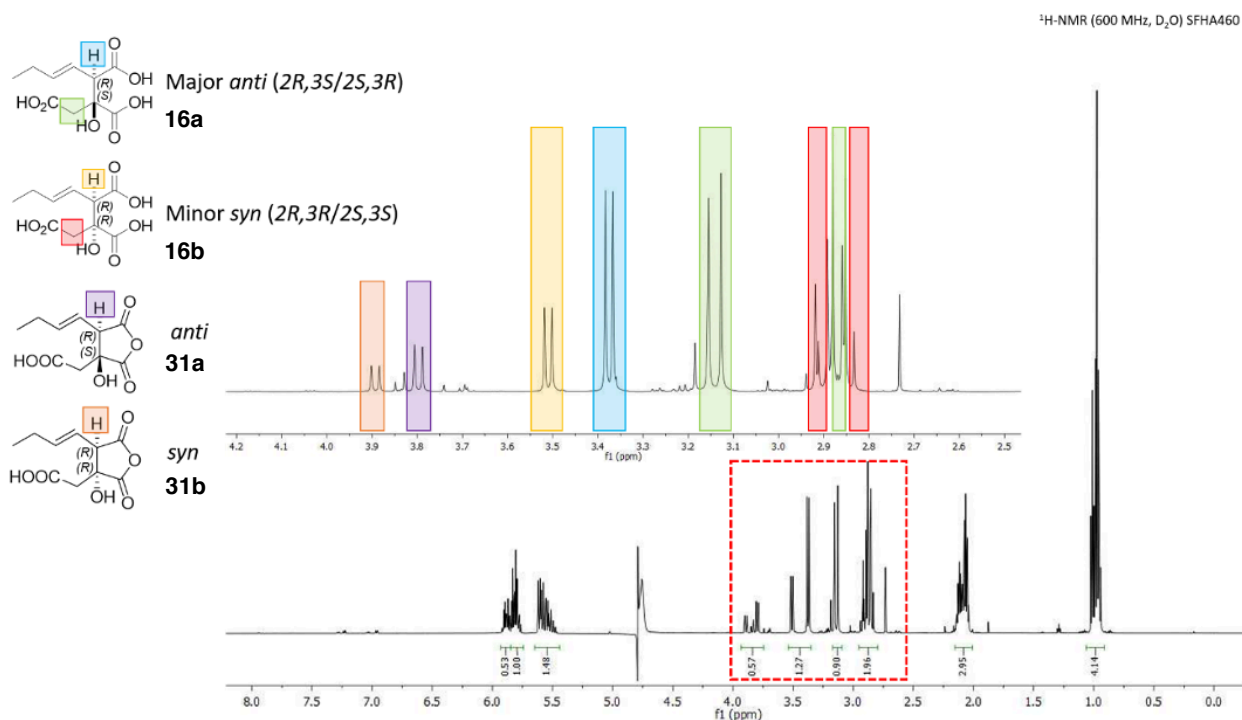


Fig. S2.2B LCMS analysis of BfL2 reaction on hexenoil – CoA substrate.

Synthetic alkyl citrate **16ab** is a mixture of different stereoisomers and ring-open and ring-closed forms. Comparing with the literature, the proton signals were annotated as follows: the proton indicated in blue was from the major *anti*- diastereomer **16a** ($2R, 3S / 2S, 3R$); and the proton indicated in yellow was the minor *syn*- diastereomer **16b** ($2R, 3R / 2S, 3S$).^{1,2,3} From HMBC data, we could see the difference between two diastereoisomers (**Fig. S2.2D**). It appeared that compound **16** seems to

cyclises by itself, the proton signals in purple are from the cyclised isomers *anti* **31a** and *syn* **31b**. None of these compounds have any significant UV absorption when the mixture was examined by LCMS.



The mixture of *syn* and *anti* diastereomers of the synthetic citrate **16** was analysed by LCMS. According to the analytical LCMS of **16**, the retention time of **16a** (major compound) is 3.3 min, and **16b** is 3.5 min. Neither of these two isomers have specific UV absorption. Both of these two isomers have [M - H]⁻ 245 (Fig. S2.2E).

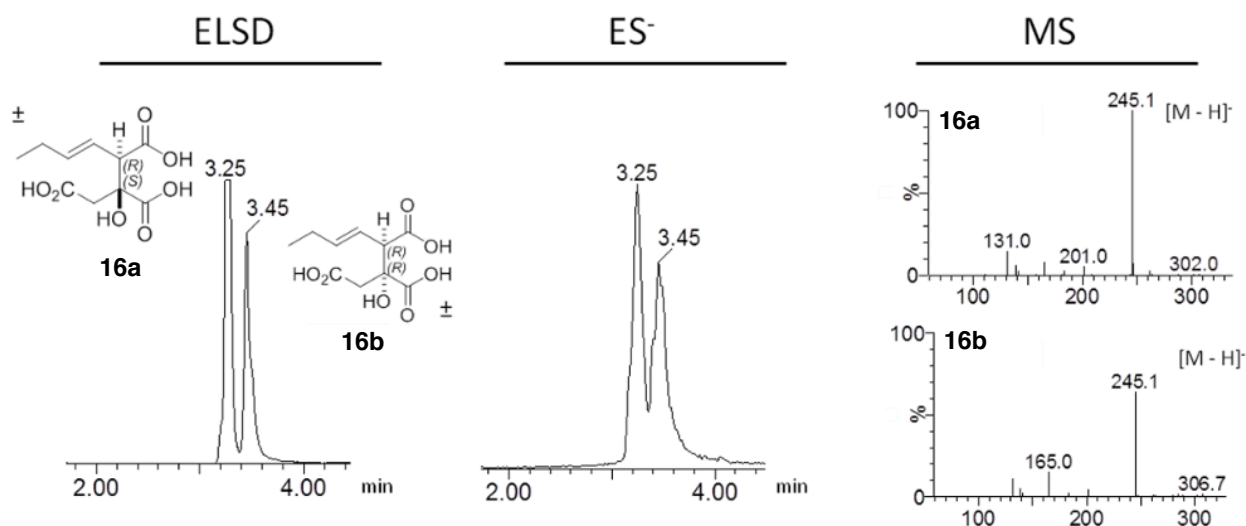


Fig. S2.2E LCMS analysis of citrate substrate

Extracted ion chromatogram analysis for the 245 ion showed two peaks. However, the BfL2 reaction trace only showed one peak which had the same retention time as the *anti* - diastereomer **16a** (Fig. S2.2F).

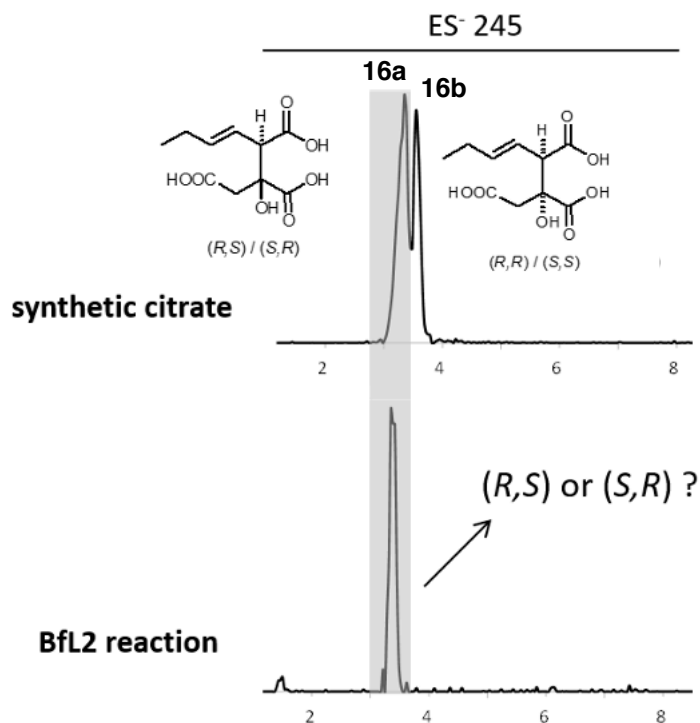


Fig. S2.2F LCMS analysis of BfL2 reaction and synthetic citrate substrate.

2.3 Dehydratase

Citrate mixture **16ab** (1 mM) was incubated with cell free extract of the yeast strain expressing PvL2 or BfL3 at 30 °C for 2 hours in 100 μ L PBS buffer. The reaction was stopped by addition of 100 μ L CH₃CN. Protein was removed by centrifugation and the supernatant was analysed directly by LCMS.

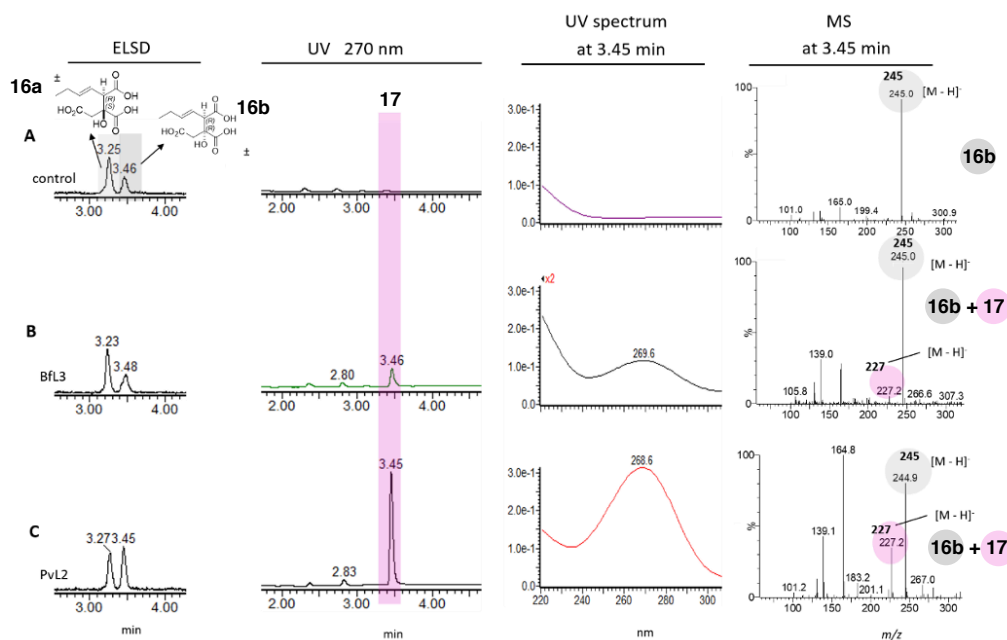


Fig S2.3A LCMS analysis of PvL2 and BfL3 reactions *in vitro*.

Both PvL2 and BfL3 showed activity on the mixture of citrate diastereomers (Fig. S2.3A). From the ELSD detector, it was obvious that the peak area at 3.5 min (from PvL2 reaction) increased compared with the empty vector control (Fig. S2.3A A/C), whereas, BfL3 didn't show much difference with the control (Fig. S2.3A B). But when it comes to the UV detector, both PvL2 and BfL3 formed a new peak with λ_{\max} 269 nm (Fig. S2.3A B and C). The empty vector control had no UV absorption in the same position (Fig. S2.3A A).

From the previous study of Dr. Agnieszka Szwalbe, the maleic anhydride monomer **1** (Figure S2.3B) is known to decarboxylate spontaneously.⁴ At the same time, **1** and its decarboxylated form **2** also equilibrate with ring-opened forms of **17** and **18**, both of them had similar UV spectra (λ_{\max} = 269 nm) and the expected masses (m/z 227.5 and 183.5 [$M + H_2O - H$]⁻ respectively). The features of the compound **18** were consistent with the new compound synthesised from the PvL2 and BfL3 *in vitro* reactions (Fig. S2.3C B and C). However, this new product co-elutes with the minor *syn* diastereomer of **16b** at 3.45 min.

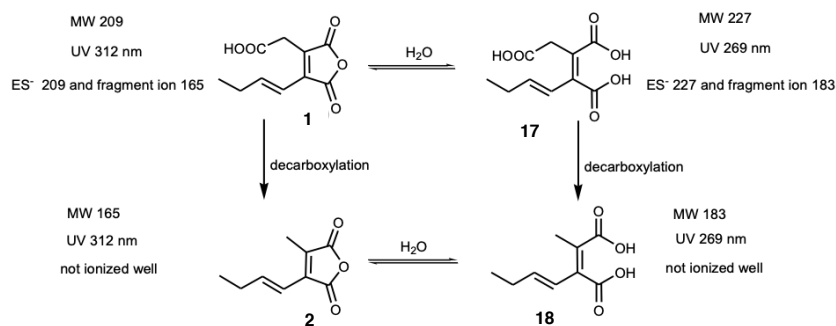
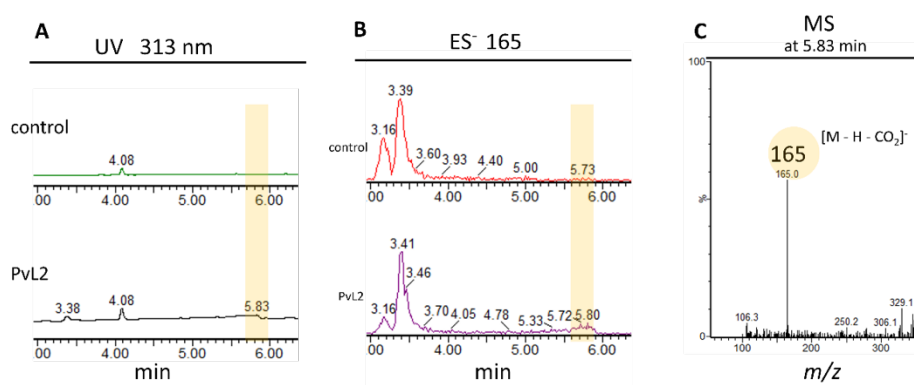


Figure S2.3B Decarboxylation and ring opening of maleic anhydride monomer.

In addition, when we observed the extracted UV chromatogram at 313 nm (the maximum absorption wavelength of maleic anhydride monomer), a small peak showed at 5.83 min (**Fig. S2.3C A**). The ES⁻ of this peak was 165 which corresponds to the fragment $[M - \text{COOH}]^-$ from maleic anhydride monomer **1** (**Fig. S2.3C B and C**). However, the empty vector control had no peak in this region. The UV spectrum of this peak was too weak to detect.



Scheme S2.3C Chromatography of PvL2 reaction: **A**, under UV 313nm; **B**, ES⁻ 165; **C**, MS spectrum of the 5.8 min peak.

Stereoselectivity of 2-methyl citrate dehydratase

Initial *in vitro* assays showed that both PvL2 and BfL3 were active, but the results did not reveal the stereoselectivity of the enzymes. Therefore, a purification method was set up for separating the isomeric **16a anti** (2*R*, 3*S* / 2*S*, 3*R*) and **16b syn** (2*S*, 3*S* / 2*R*, 3*R*) diastereomers by prep-LCMS (**Fig S2.3D**). Both **16a** (2*R*, 3*S* / 2*S*, 3*R*) and **16b** (2*S*, 3*S* / 2*R*, 3*R*) had $[M - \text{H}]^-$ 245 (**Fig. S2.3D A**), and these two compounds were well-separated on a C₁₈ reverse-phase prep-column. The two citrate isomers do not have UV absorption (**Fig. S2.3D C**) but can be observed by ELSD (**Fig. S2.3D B**).

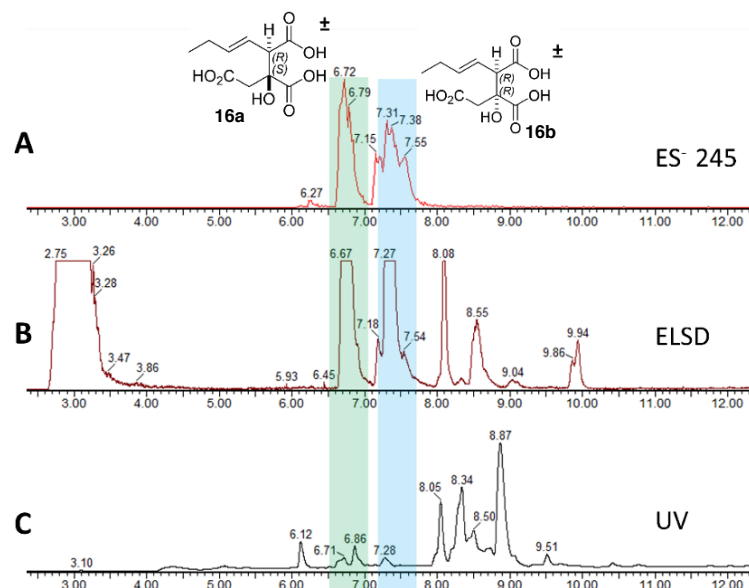


Fig. S2.3D LCMS purification result of citrate substrates, and the shaded peaks were purified: **A**, ES⁻ 245 ion extraction; **B**, ELSD; **C**, UV chromatogram.

The pure racemic isomers of **16a** (*2R*, *3S* / *2S*, *3R*) and **16b** (*2S*, *3S* / *2R*, *3R*) were used individually for the DH assays *in vitro*. After 3 hours of incubation with substrate and enzymes (PvL2 or BfL3), the *anti*-isomer of **16a** (*2R*, *3S* / *2S*, *3R*) was clearly converted to a new product (the pink shade **Fig S2.3E A and B**) compared with the control (**Fig S2.3E C**). However, *syn* diastereomer **16b** (*2S*, *3S* / *2R*, *3R*) had no reaction (**Fig S2.3E D, E and F**). Both PvL2 and BfL3 were selective for the **16a** (*2R*, *3S* / *2S*, *3R*) *anti* diastereomer.

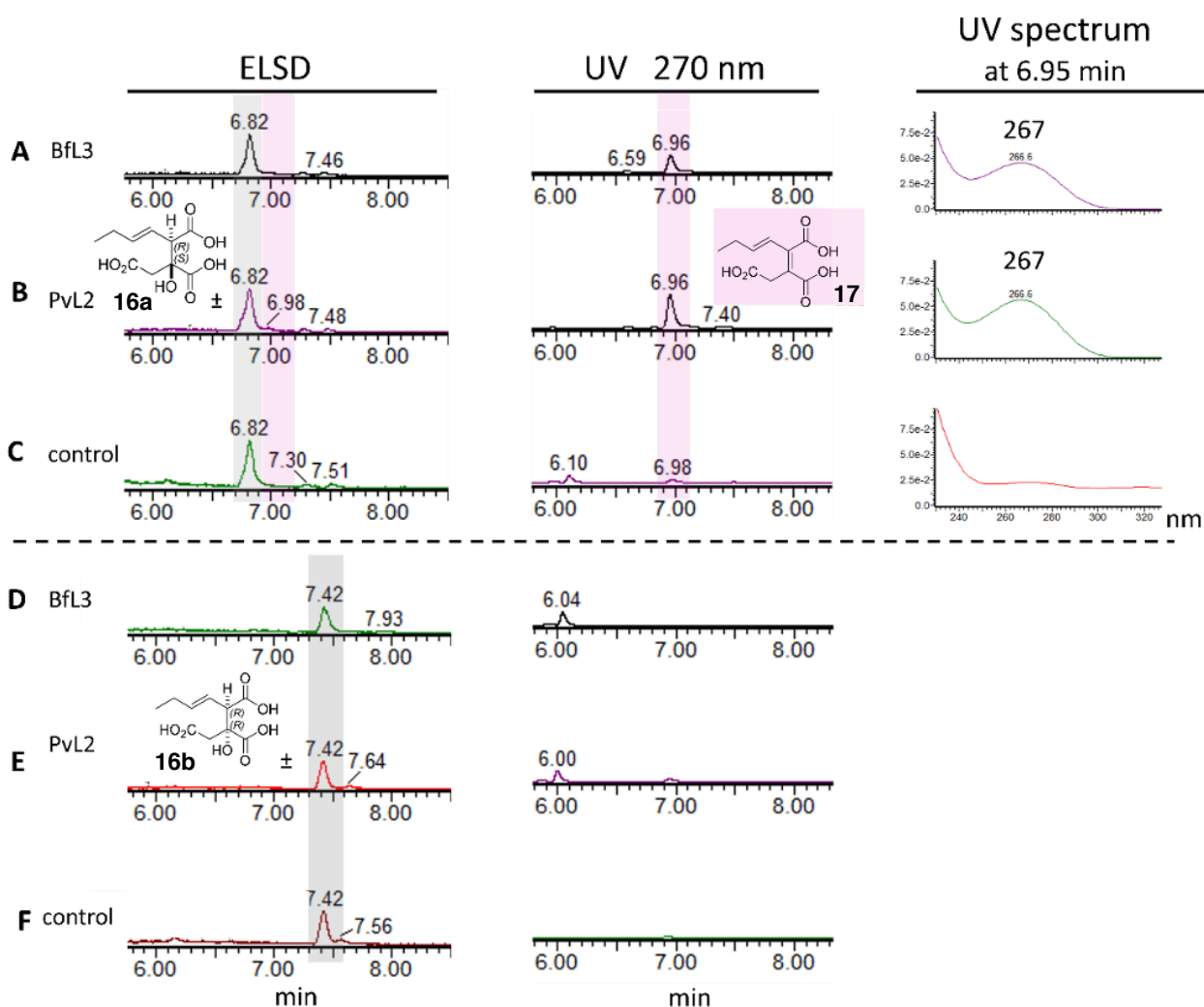


Fig S2.3E LCMS analysis of the DH reactions: **A**, **B** and **C** treated with (2*R*, 3*S* / 2*S*, 3*R*) *anti* substrate using enzyme BfL3, PvL2 and empty vector control respectively. **D**, **E** and **F** treated with (2*R*, 3*R* / 2*S*, 3*S*) *syn* substrate using enzyme BfL3, PvL2 and empty vector control respectively.

The citrate substrate **16** has no conjugated double bond and therefore no significant UV absorption. The product **17** has a conjugated system, so the maximum absorption wavelength has a blue shift to 267 nm (**Fig. S2.3F B**). The molecular weight changed from $[M - H]^-$ 245 to $[M - H]^-$ 227. The fragment ion 165 $[M - H_2O - COOH]^-$ is due to decarboxylation (**Fig. S2.3F A**).

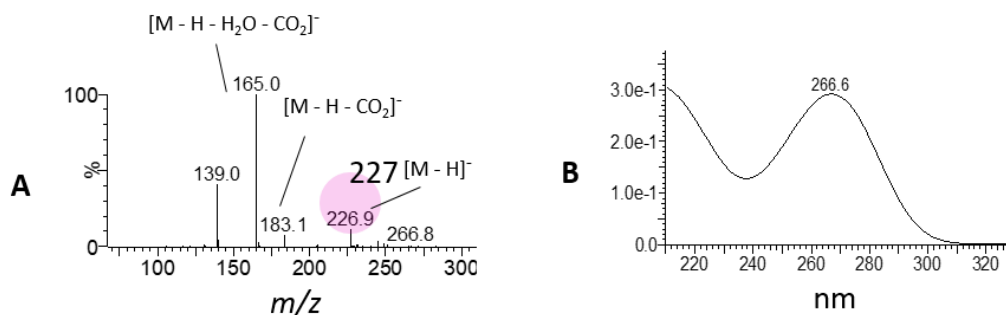


Fig. S2.3F UV and MS spectrum of product: **A**, UV spectrum; **B**, MS spectrum.

The ES⁻ extracted ion chromatogram (EIC) at *m/z* 245 and 227 were examined, the retention time of *anti* citrate substrate **16a** (2*R*, 3*S* / 2*S*, 3*R*, grey shaded) was at 6.8 min (**Fig. S2.3G A and C**), while the double bond product **17** (pink shaded) was at 7.0 min (**Fig. S2.3G D**). In this chromatographic method, *anti* citrate substrate **16a** (2*R*, 3*S* / 2*S*, 3*R*) had about 0.1 min time shift comparing with product **17**. As for the reaction of empty vector control, there was no product peak shown on chromatogram (**Fig. S2.3G B**).

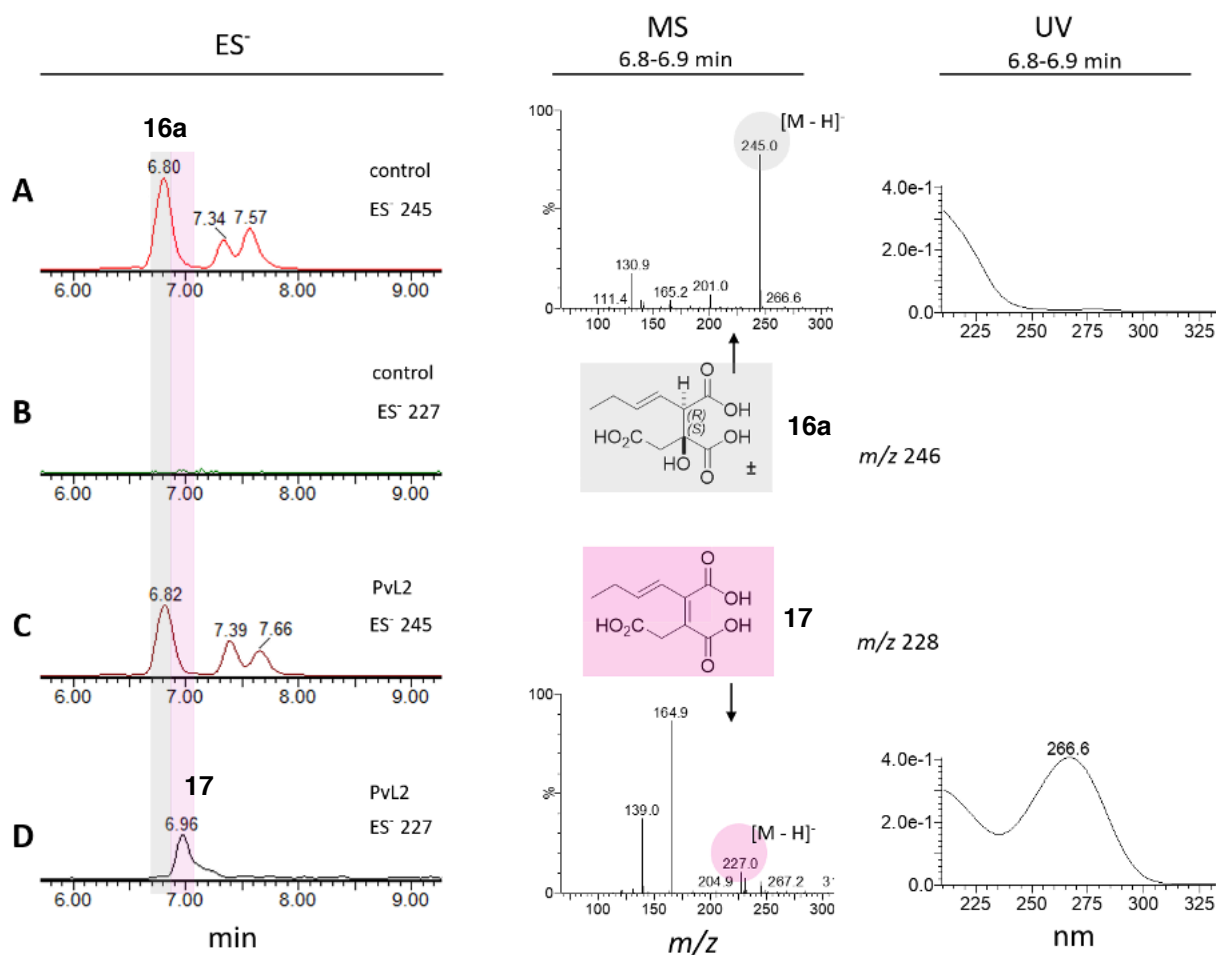


Fig. S2.3G ES⁻ spectrum of substrate and product: **A** and **B**, empty vector control in ES⁻ 245 and 227; **C** and **D**, PVL2 reaction in ES⁻ 245 and 227.

It was interesting that when we observed the extracted UV chromatogram at 313 nm (the maximum absorption wavelength of maleic anhydride monomer), a small peak showed at 9.7 min (**Fig. S2.3H B**). The ES⁻ of this peak was 165 which corresponds to the fragment [M-COOH]⁻ from maleic anhydride monomer **1** (**Fig. S2.3H B**). The UV spectrum of this peak had a maximum absorption at 312 nm which was consistent with anhydride monomer **1** (**Fig. 3.2.7.5 B**). However, the empty vector control had no peak in this region (**Fig. S2.3H A**). It means that the DH enzyme can dehydrate citrate and at the same time the product **17** can cyclise to **1** in the reaction conditions.

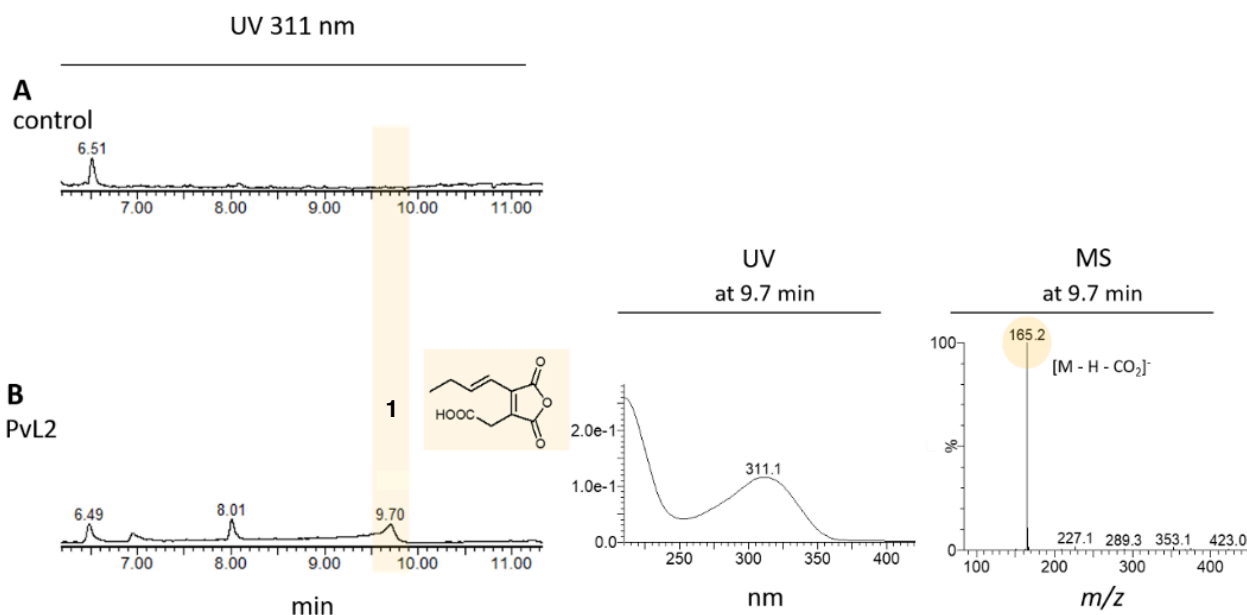


Fig. S2.3H Chromatography of PVL2 and empty vector control reaction under UV 313nm, and MS and UV spectrum of the maleic anhydride monomer: **A**, control; **B**, PVL2.

2.4 KI/PEBP

The pESC-URA vector with whole ORF of *bfl5/6/9* and *10* were transformed into yeast host *Saccharomyces cerevisiae* w303B, and checked by yeast colony PCR (**Fig. S2.4A**). Then BfL5-2, BfL6-1, BfL9-2, BfL10-2 were chosen for protein expression.

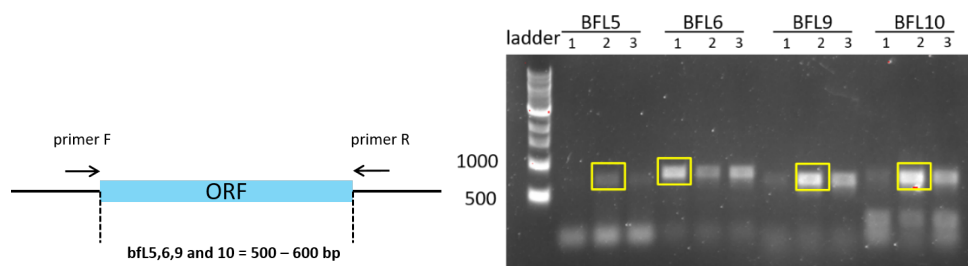


Fig. S2.4A Yeast colony PCR.

Cell-free extracts of BfL5, 6, 9 and 10 yeast transformants were prepared by using a glass bead vortex method. Maleic anhydride monomer suspension **1** (15 μ L) and each cell extraction BfL5/6/9 and 10 (20 μ L) then mixed. After reaction at 30 $^{\circ}$ C for 8 h, the reaction was stopped by adding 100 μ L of acetonitrile. The protein was precipitated by centrifugation and the supernatant was analysed by LCMS.

The result showed that the substrate **1** was converted to three products. These were decarboxylated product **2** and ring-opened compound **17**. A new peak was observed in the chromatogram (**Fig. S2.4B A**). However, the empty vector control did not produce the new peak (**Fig. S2.4B E**). Analysis of the MS spectrum and UV spectrum of peak **1** showed that the compound (red shade) was identical with byssochlamic acid **1**. It had molecular ion of m/z 331 (**Fig. S2.4C**) and

maximum UV absorption at 248 nm which was the same as byssochlamic acid **1** (Fig. S2.4C). The retention time of peak **1** was also identical to the product from the *A. oryzae* transformants (Fig. S2.4B G). The results indicated that the mixed protein of two KI (BfL6 and 10) and two PEBP (BfL5 and 9) could dimerise the maleic anhydride monomer. However, there is no evidence of heptadride **10** formation according to the LCMS results.

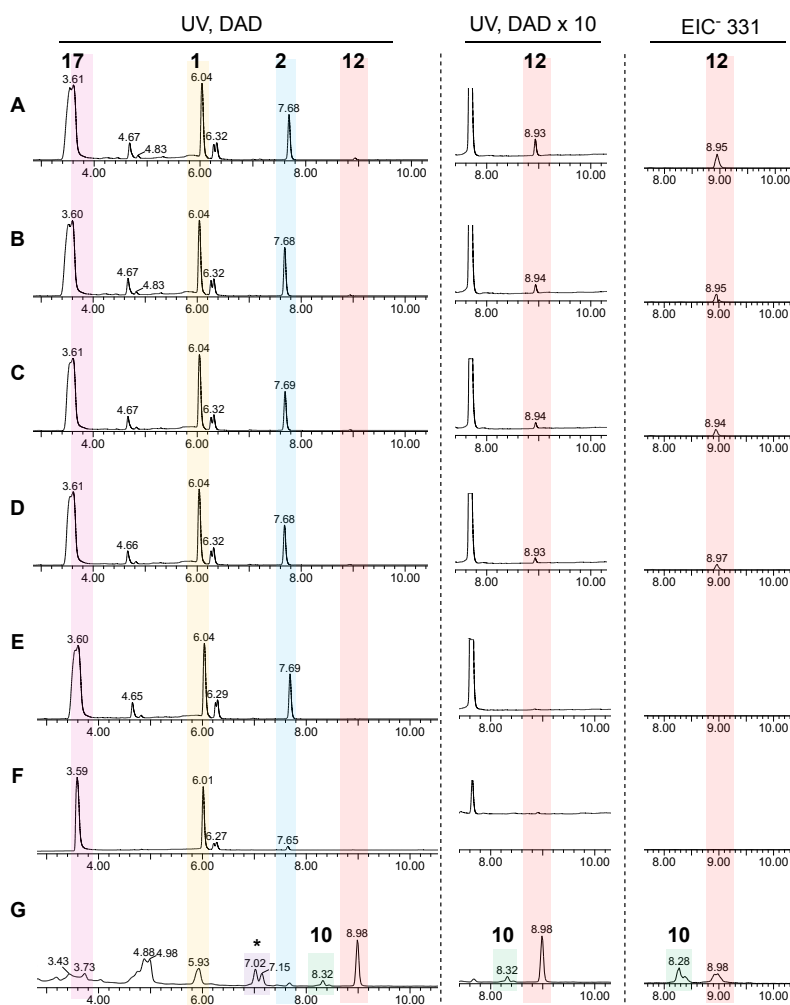


Fig. S2.4B LC-MS analysis the *in vitro* test using yeast expression protein KI and PEBP: **A**, two KSI and two PEBP; **B**, two KI; **C**, BfL6; **D**, BfL10; **E**, empty vector control; **F**, only substrate; **G**, producing *A. oryzae* transformant.

At the same time, BFL6 and BFL10 were also tested without PEBP enzymes. A weak byssochlamic acid **1** peak was also detected (Fig. S2.4C B). Two KSI proteins together could catalyse the dimerization step without the help of the PEBP proteins which was consistent with the previous *in vivo* study in *A. oryzae*.⁵ In addition, BfL6 and BfL10 were tested separately in two different reactions. Both BfL6 (Fig. S2.4B C) and BfL10 (Fig. S2.4B D) had a weak activity of dimerization to give a low concentration of byssochlamic acid **12**. When we look into the ES⁻ 331 of LCMS data, it was clearer to see the production of byssochlamic acid **12** product from the KI enzymes (in red shade, Fig. S2.4B A/B/C/D and G)

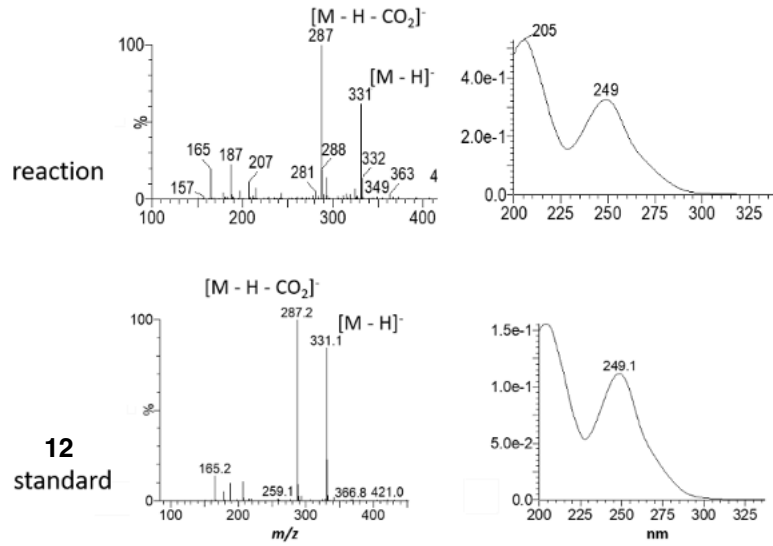


Fig. S2.4C MS spectrum and UV spectrum of peak 12.

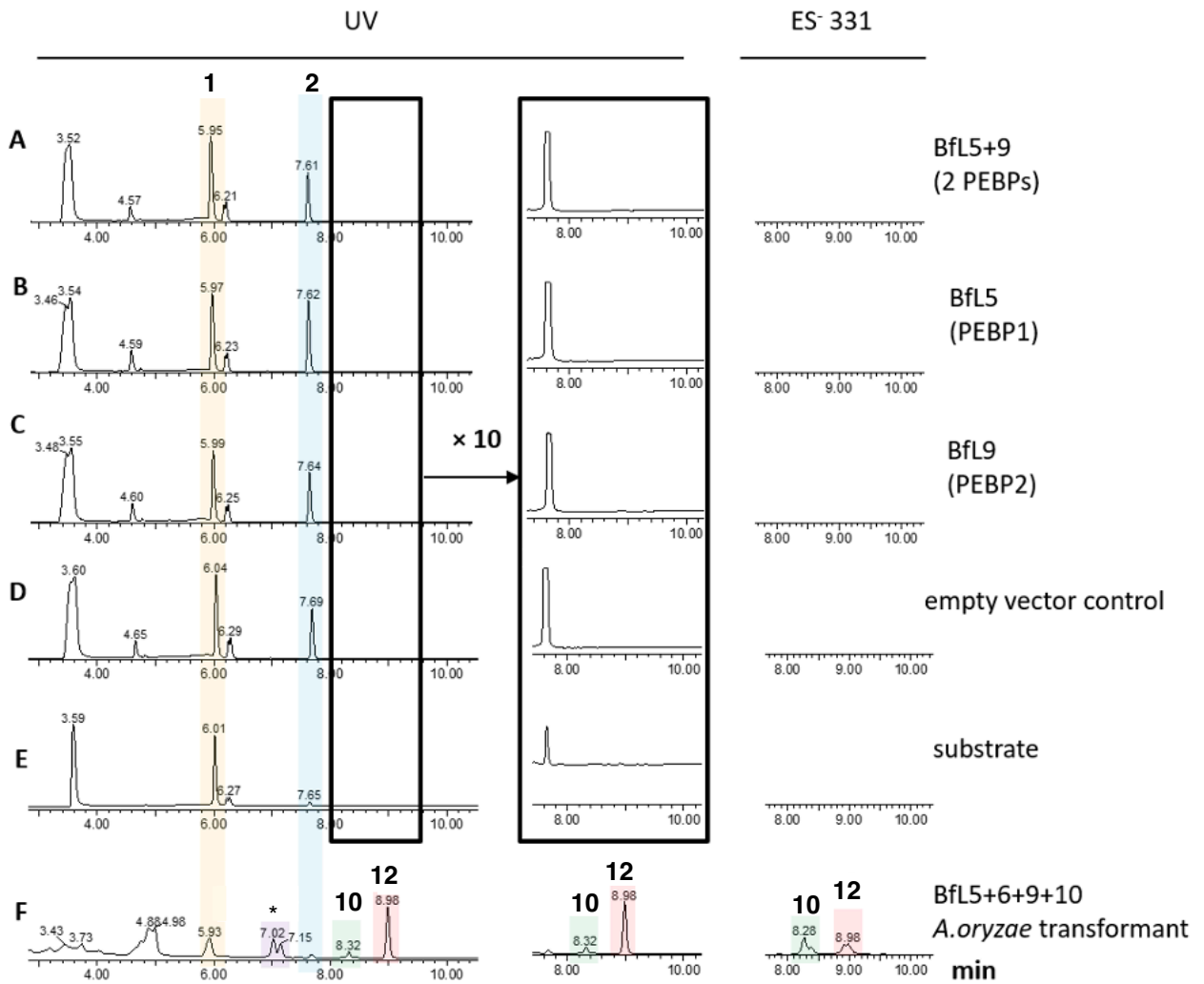


Fig. S 2.4D LCMS analysis the *in vitro* test using yeast expression protein PEBPs alone.

* = diacid forms of byssochlamic acid 12.

Finally, PEBP enzymes BFL5 and BFL9 were tested in this study to check if they were related to the dimerization reactions. The result indicated that no matter whether both BFL5 and BFL9 (**Fig. S2.4D A**) were tested together or they worked separately (**Fig. S2.4D B and C**), there was no byssochlamic acid **1** formation compared with empty vector control (**Fig. S2.4D D**) and *A. oryzae* transformant (**Fig. S2.4D F**).

In vitro study of *E. coli* expressed KI and PEBP enzymes

As it was displayed previously, KI and PEBP enzymes were also expressed in *E. coli*. Due to N terminal signal peptide, it was difficult to express BfL5, BfL6, and BfL10 directly using whole ORF amino acid sequence. The signal peptide sequence was removed by PCR when the expression vector was constructed. To improve the solubility, chaperone proteins were co-expressed with KI and PEBP enzymes. Two KI enzymes were used in this study.

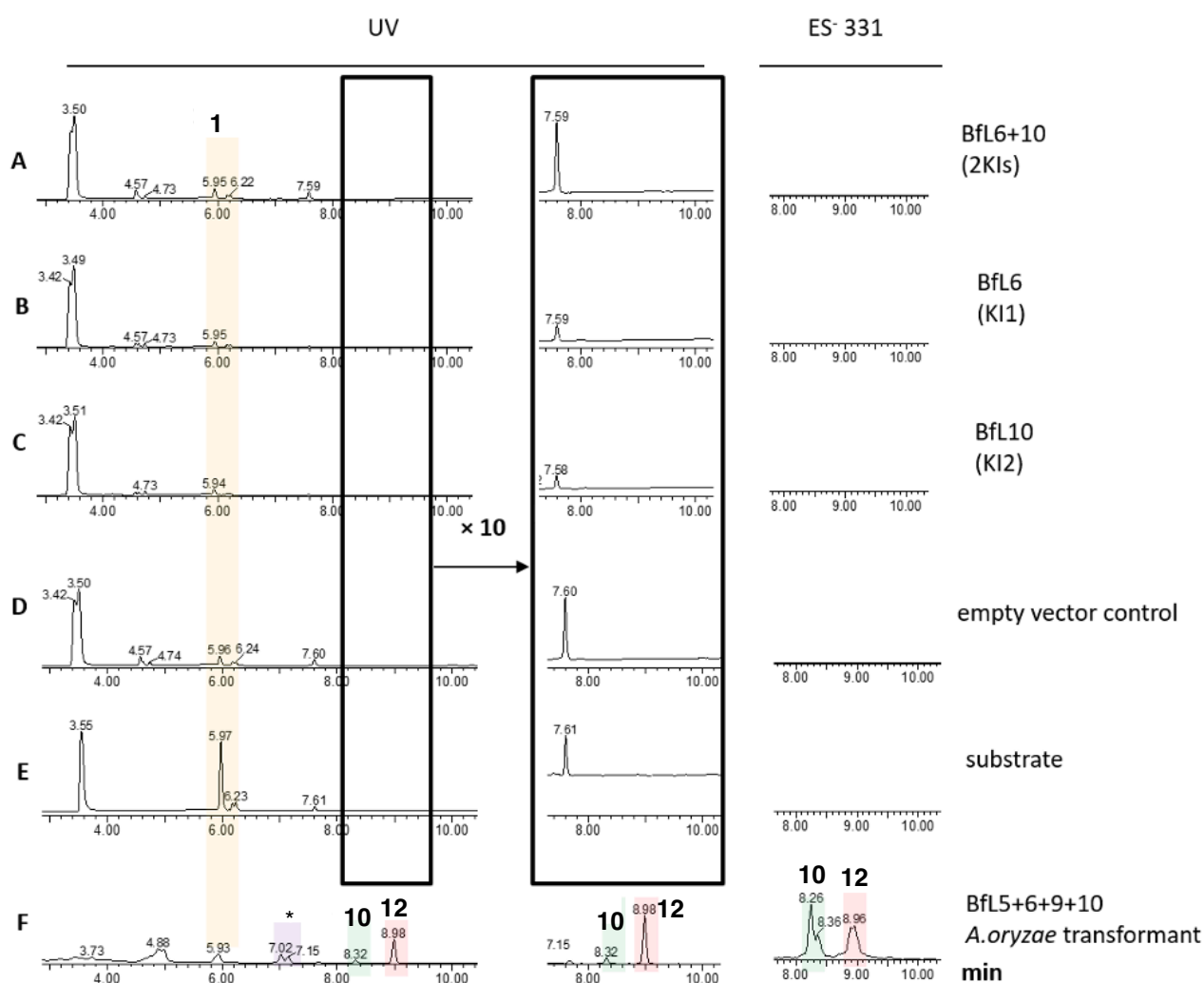


Fig. S2.4E LCMS analysis the *in vitro* assay using *E. coli* expression protein KSI and PEBP proteins.

* = diacid forms of byssochlamic acid **12**.

The crude *E. coli* cell extractions were used for *in vitro* assay. The experimental groups were set up as BfL6+10 (two KI **Fig. S2.4E A**), BfL6 (KI1 **Fig. S2.4E B**), BfL10 (**Fig. S2.4E C**) and empty vector control (**Fig. S2.4E D**). After the 8h reaction, reactions were stopped by addition 100 μ L of CH₃CN, then the samples were analysed by LCMS. From ES⁻ analysis, the result indicated that none of this experimental group could give any dimerised product compared with *A. oryzae* transformant extraction (**Fig. S2.4E F**). It might be because the loss of the N terminal signal peptide region made the proteins not stable or it can affect the proteins to get the right folding.

Decarboxylated compound 2 is not a substrate for KI

Even though the maleic anhydride monomer substrate decarboxylated, an *in vitro* test was still performed by using the compound **2** as a substrate. BfL6 (KI1) and BfL10 (KI2) from both yeast and *E. coli* expression (cell free extraction) were used in this assay. Empty vector transformants were used as controls. After three hours of incubation, the reaction was stopped by addition of 100 μ L of CH₃CN. Protein was precipitated by centrifugation and the reaction mixture was analysed by LCMS. No dimerization was observed (**Fig. S2.4F A and C**). The compound at retention time 7.6 min was converted to the compound at 4.5 min. And according to the UV spectrum of compound 4.5 min, it was the ring-opened form of **18**. At the same time, there was still a low concentration of ring-opened aconitate **17** at 3.2 min (**Fig. S2.4F A/B/C/D**).

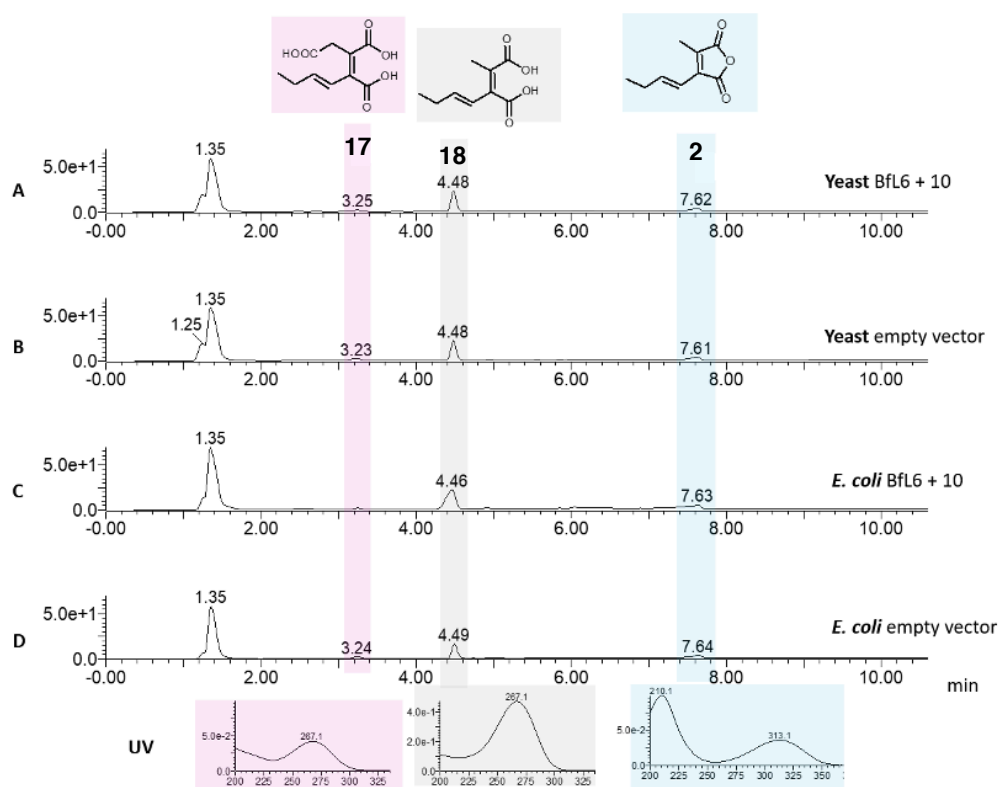


Fig. S2.4F *In vitro* study on decarboxylated substrate **2**.

3. Synthetic Chemistry

14b *E*-Hex-2-enoyl SNAC⁶

Bromoacetic acid (SigmaAldrich, 4.48g, 32.3 mmol), *N*-acetyl cysteamine **19**⁷ (3.50 g, 29.4 mmol) and dry CH₂Cl₂ (100 ml) were mixed under dry argon and cooled to 0 °C. EDCI.HCl (6.17 g, 32.3 mmol) and DMAP (0.36 g, 2.94 mmol) were added with stirring and the reaction was allowed to warm to RT. The solution was stirred for 48 h in the dark. Aqueous HCl (1 M, 100 ml) was added with mixing. The organic phase was extracted with saturated aqueous NaHCO₃, extracted with brine, then dried (anh. MgSO₄) and evaporated to afford bromoacetyl SNAC **19b** as a yellow oil (3.73 g, 15.5 mmol, 52.7%). This was used directly in the next step without further characterisation.

Bromoacetyl SNAC **19b** (3.7 g, 15.4 mmol) and PPh₃ (4.04 g, 15.4 mmol) were mixed in dry toluene (20 ml) which was heated to 80 °C for 18h. The toluene was removed *in vacuo* and the solid residue suspended in EtOAc (30 ml) which was washed with water (3 x 30 ml). The combined aqueous layers were basified with NaOH to pH 13, then extracted with CH₂Cl₂ (3 x 30 ml). The combined organic extracts were dried (MgSO₄) and evaporated to give the expected ylide as pale yellow crystals (4.57 g, 10.8 mmol, 70 %) which were used without further purification or characterisation.

Ylide **19c** (100 mg, 0.24 mmol), CH₂Cl₂ (500 μL) and freshly distilled butyraldehyde (100 μL, 1.11 mmol) were mixed with stirring for 72 h at RT. CH₂Cl₂ (25 mL) was added and the solution washed with brine (2 x 20 mL). The organic layer was dried (anh. MgSO₄) and evaporated to give a solid residue which was purified by flash chromatography (50:50 Et₂O:Petroleum ether) to give the product as a colourless oil (37 mg, 17.2 mmol, 72 %). ¹H NMR (CDCl₃): 0.94 (t, 3H, *J* = 8.4), 1.51 (m, 2H), 1.97 (s, 3H), 2.19 (m, 2H), 3.09 (t, 2H, *J* = 6.4), 3.46 (dt, 2H, *J* = 6.4, *J* = 6.4), 6.13 (m, 2H), 6.92 (dt, 1H, *J* = 6.9, *J* = 15.6). ¹³C NMR (CDCl₃): 13.7, 21.2, 23.2, 28.2, 34.2, 39.8, 128.4, 146.5, 170.5, 190.4. MS (ESI⁺): 216.4 (MH⁺), 120.1 (CH₃CONHCH₂CH₂SH₂⁺, 65%)

14c *E*-Hex-2-enoyl Pantetheine⁶

Bromoacetic acid (96 mg, 0.69 mmol) and pantetheine dimethylacetal **20a** (200 mg, 0.63 mmol) were dissolved in CH₂Cl₂ (2 ml). EDCI (132 mg, 0.69 mmol) and DMAP (7.5 mg, 0.06 mmol) were added with stirring. The reaction mixture was stirred at RT for 24 h then quenched by addition of aqueous HCl (1 M, 2 ml). The aqueous phase was washed with CH₂Cl₂ (1 x 5 ml). The combined organic phases were washed with saturated aqueous NaHCO₃ (5 ml) and brine (5 ml), then dried (anh. MgSO₄) and evaporated to afford a

colourless oil **20b** 160 mg, 0.36 mmol, 52 %) which was used without further characterisation in the next step.

20b (1.38 g, 3.15 mmol), triphenyl phosphine (826 mg, 3.15 mmol) and toluene (4 ml) were mixed at RT. The reaction mixture was heated to 80 °C for 16h, then the toluene was removed *in vacuo* and the solid residue suspended in EtOAc (30 ml) which was washed with water (3 x 30 ml). The combined aqueous layers were basified with NaOH to pH 13, then extracted with CH₂Cl₂ (3 x 30 ml). The combined organic extracts were dried (anh. MgSO₄) and evaporated to give the expected ylide **20c** as a pale yellow solid (400 mg, 0.65 mmol, 20.5 %).

Ylide **20c** (400 mg, 0.64 mmol) and freshly distilled butyraldehyde (115 µl, 1.28 mmol) were dissolved in CH₂Cl₂ (6 ml) at RT. The mixture was heated to 40 °C for 3 d. Solvent was removed *in vacuo* and the residue purified by flash chromatography (50:50 Et₂O:Petroleum ether) to yield the expected protected product as colourless crystals which were dissolved in anhydrous THF (1.8 ml) to which trifluoroacetic acid (200 µl) was added at RT with stirring. After 60 min H₂O (1.0 ml) was added and stirring continued overnight. Solvent was removed *in vacuo* and the residue purified by flash chromatography (EtOAc) to give the expected product (123 mg, 0.33 mmol, 51.3%). ¹H NMR: 0.93 (s, 3H, CH₃-16), 0.96 (t, 3H, *J* = 7.4, CH₃-6), 1.01 (s, 3H, CH₃-17), 1.52 (m, 2H, CH₂-5), 2.20 (dt, 2H, *J* = 6.6, *J* = 6.8, CH₂-4), 2.44 (t, 2H, *J* = 5.9, CH₂-10), 3.09 (m, 2H, CH₂-7), 3.45 (m, 4H, CH₂-15 and CH₂-11), 3.57 (m, 2H, CH₂-8), 4.01 (s, 1H, CH-13), 6.14 (d, 1H, *J* = 5.3, CH-2), 6.60 (brt, 1H, NH), 6.94 (dt, 1H, *J* = 6.9, *J* = 15.6, CH-3), 7.46 (brt, 1H, NH). MS: ESI⁺, 771.7 ([M₂]Na⁺, 20 %), 397 ([M]Na⁺, 30 %), 375 ([M]H⁺, 100 %), 245 (30 %); ESI⁻, 419.2 ([M + formic acid - H]⁻, 80 %), 373.2 ([M-H]⁻, 100 %). u_vmax (CH₃CN/H₂O): 226, 262 nm.

14d E-Hex-2-enoyl CoA⁸

E-Hex-2-enoic acid (6.0 mg, 0.052 mmol) was dissolved in anhydrous CH₂Cl₂ (1.5 ml). Et₃N (7.5 µL, 0.052 mmol) was added by syringe. The solution was cooled to 0 °C and ethylchloroformate (5.5 µl, 0.052 mmol) was added by syringe. After 2 h the CH₂Cl₂ was removed under a gentle flow of dry N₂. DMF (1.5 ml) was added. Coenzyme A trilithium salt (21.4 mg, 0.026 mmol) was dissolved in aqueous NaHCO₃ (0.4 M, 1.5 ml) and this solution added to the mixed anhydride solution at RT. After 10 minutes the reaction mixture was acidified by addition of formic acid (100 µl) and added to water (50 ml). The aqueous solution was frozen and lyophilised to afford the product as a colourless solid (53 mg). MS: ESI⁻, 884.7 ([M + Na - 2H]⁻, 5 %), 862.3 ([M-H]⁻, 85 %), 430.7 (100 %); ESI⁺, 886.5 ([M]Na⁺, 5 %), 864.5 ([M]H⁺, 30 %), 432.9 (100 %). u_vmax: (CH₃CN/H₂O) 258 nm.

15b Hexanoyl SNAC⁹

Synthesised according to Moon Il Kim, Seok Joon Kwon, and Jonathan S. Dordick, *Org. Lett.*, **2009**, 11(17), 3806-3809.

15c Hexanoyl Pantetheine¹⁰

Pantetheine dimethylacetal (0.5 g, 1.57 mmol), hexanoic acid (0.39 ml, 3.14 mmol), DMAP (0.38 g, 3.14 mmol) and EDCI (1.38 g, 7.22 mmol) were dissolved in CH₂Cl₂ (8 ml) at RT with stirring. After 2h the reaction was quenched by the addition of aqueous HCl (0.1 M, 10 ml). The mixture was extracted with CH₂Cl₂ (4 x 10 ml) and the combined organic phases dried (anh. MgSO₄) and evaporated *in vacuo*. The organic residue was purified by flash chromatography (1:1 EtOAc: Petroleum ether to 100 % EtOAc) to afford the desired product as a colourless oil (410 mg).

The protected pantetheine (200 mg, 0.48 mmol) was dissolved in a mixture of THF (2 ml) and water (10 ml) and then 5 drops of TFA were added at RT. After 4 h solvent was removed *in vacuo* and the residue was purified by flash chromatography (100% EtOAc to 95% EtOAc plus 5% MeOH) to give the product as a colourless oil (140 mg, 0.37 mmol, 24 %). MS: ESI⁻, 421.2 ([M + formic acid - H]⁻, 100 %), 375.2 ([M-H]⁻, 85 %); ESI⁺, 399.4 ([M]Na⁺, 5 %), 377.4 ([M]H⁺, 100 %), 359.4 (M + H - H₂O)⁺, 40%), 247.3 (80 %). uV_{max} : (CH₃CN/H₂O) 232 nm. ¹H NMR (CDCl₃, 400 MHz) 0.92 (t, 3H, J = 7.1 Hz, CH₃-6), 0.96 (s, 3H, CH₃-16), 1.04 (s, 3H, CH₃-17), 1.33 (m, 4H, CH₂-4 and CH₂-5), 1.67 (m, 2H, CH₂-3), 2.47 (m, 2H, CH₂-2), 2.46 (t, 2H, J = 6.9, CH₂-10), 3.05 (m, 2H, CH₂-7), 3.40 (m, 2H, CH₂-11), 3.54 (s, 2H, CH₂-15), 3.60 (m, 2H, CH₂-8), 3.98 (s, 1H, CH-13), 6.27 (brt, 1H, OH-15), 6.44 (brt, 1H, NH), 7.50 (brt, 1H, NH), 8.12 (brs, 1H, OH-13).

16a E, 3S*, 4R* - 3-hydroxy-3,4-dicarboxy-oct-5-ene

16b E, 3S*, 4S* - 3-hydroxy-3,4-dicarboxy-oct-5-ene

Triester **28** (50 mg) was suspended in aqueous HCl (3 M, 25 ml) and refluxed for 3 h. The reaction mixture was extracted with Et₂O (3 x 25 ml). The combined ethereal extracts were extracted into saturated aqueous NaHCO₃ (2 x 20 ml). The combined aqueous extracts were acidified with aq HCl and re-extracted with Et₂O (3 x 25 ml). The ethereal extracts were dried (anh. Na₂SO₄) and the solvent was evaporated. The mixture was dissolved in CH₃CN:H₂O (1:1, 1 ml) and sufficient material for enzyme assays and further analysis was purified by automated mass-directed reverse-phase chromatography. The *anti* diastereomer **16a** was eluted first, followed by the *syn* diastereomer **16b**.



D₂O, 600 / 150 MHz

position	δ_{H} / ppm	J / Hz	δ_{C} / ppm	δ_{H} / ppm	J / Hz	δ_{C} / ppm
6	5.82 (dt)	15.3, 6.3	140.0	5.88 (dt)	15.8, 6.2	141.0
5	5.60 (ddt)	15.4, 9.9, 1.4	120.6	5.53 (ddt)	15.6, 9.8, 1.5	120.0
4	3.38 (d)	9.8	57.0	3.51 (d)	9.7	56.7
2a	3.14 (d)	16.5	41.9	2.91 (d)	15.8	41.7
2b	2.87 (d)	16.5	41.9	2.85 (d)	16.0	41.7
7	2.06 (m; <i>pseudo</i> -ddq)	7.6, 1.4	25.1	2.12 (m; <i>pseudo</i> -ddq)	7.5, 1.4	25.1
8	0.97 (t)	7.5	12.4	1.01 (t)	7.5	12.4
3	-	-	76.8	-	-	76.2
1	-	-	174.0	-	-	174.0
10	-	-	175.5	-	-	175.7
9	-	-	176.2	-	-	176.2

24 Dimethyloxaloacetate

Dimethylacetylenedicarboxylate (16.3 mmol, 2.32 g) was mixed with water (60 ml) with stirring at RT. Morpholine (19.6 mmol, 1.7 ml) was added. The solution warmed, clarified and became yellow. Stirring was continued for 90 min and then the solution was extracted with EtOAc (2 x 50 ml). The EtOAc extract was dried (anh. Na₂SO₄) and evaporated. The residue was dissolved in Et₂O (60 ml). Oxalic acid (19.6 mmol, 1.76g) was dissolved in Et₂O (30 ml) and EtOH (6 ml), and this solution added to the first solution to give a turbid mixture. Water (300 μ l) was added, and the mixture stirred overnight at RT. The resulting colourless precipitate was removed by filtration. The filtrate was washed with water (3 x 15 ml), dried (anh. Na₂SO₄), and evaporated to give yellow crystals which were recrystallised from Et₂O and dried under high vacuum to give the product as pale yellow crystals (1.18g, 7.3 mmol, 44.8%). This was used in following steps without further analysis or purification.

26 Methyl *E*-hex-2-enoate

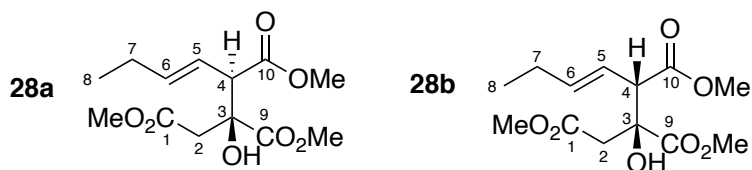
E-hex-2-enoic acid (1.0 eq., 5.0 g, 43.8 mmol) was dissolved in THF (100 mL) and cooled with an ice bath. Oxalylchloride (1.0 equiv., 3.8 mL, 43.8 mmol) and DMF (0.3 equiv., 1.0 mL, 13.0 mmol) were added dropwise. After stirring for 30 minutes, methanol (5.6 equiv., 10.0 mL, 246.6 mmol) was added dropwise and the ice bath was removed. The mixture was stirred for 2 h at room temperature. After evaporating the solvent, the residue was washed with 6 % NaHCO₃-solution (100 mL) and extracted with 2 x 100 mL diethyl ether. The organic phase was washed with saturated aqueous NaCl and the ether was evaporated. The residue was purified by distillation (180 mbar, 120 °C) giving the product as a colourless liquid (3.293

g, 57 %). ¹H NMR (400 MHz, Chloroform-d) δ 6.97 (4-CH, dt, *J* = 15.6, 7.0 Hz, 1H), 5.82 (5-CH, dt, *J* = 15.6, 1.6 Hz, 1H), 3.72 (7-CH₃, s, 3H), 2.18 (3-CH₂, ddt, *J* = 7.1, 1.6 Hz, 2H), 1.49 (2-CH₂, tq, *J* = 7.4 Hz, 2H), 0.93 (1-CH₃, t, *J* = 7.4 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 167.3 (6-C), 149.7 (4-C), 121.1 (5-C), 51.5 (7-C), 34.4 (3-C), 21.4 (2-C), 13.8 (1-C). LCMS 6.4 min, *m/z* = 129.17 [M+H⁺].

27 Methyl *E*-4-bromohex-2-enoate.

Methyl ester **26** (1.0 equiv., 2.082 g, 16.2 mmol) and NBS (1.0 equiv., 2,891 g, 16.2 mmol) were dissolved in CCl₄ (23 mL) with exclusion of moisture. Dibenzoylperoxide (0.006 equiv., 0.2522 g, 0.1 mmol) was added and the solution was refluxed for 6 h then stirred over night at room temperature. The mixture was cooled and the formed succinimide was filtered off. After washing with water (3 x 4.6 mL), drying and evaporating the solvent the residue was purified by distillation (4 mbar, 95 °C) giving the racemic product as a colourless liquid (2.586 g, 77%). ¹H NMR (400 MHz, Chloroform-d) δ 6.97 (4-CH, dd, *J* = 15.5, 9.1 Hz, 1H), 5.94 (5-CH, dd, *J* = 15.5, 0.8 Hz, 1H), 4.52 – 4.42 (3-CH, m, 1H), 3.76 (7-CH₃, s, 3H), 1.97 (2-CH₂, ddq, *J* = 7.2, 5.2 Hz, 2H), 1.03 (1-CH₃, t, *J* = 7.3 Hz, 3H). ¹³C NMR (101 MHz, CDCl₃) δ 166.3 (6-C), 146.7 (4-C), 121.7 (5-C), 53.1 (3-C), 51.9 (7-C), 31.3 (2-C), 12.2 (1-C). LCMS 6.9 min, *m/z* = 207.09, 209.07 [M+H⁺]

28 (±) methyl *E*-3-hydroxy, 3,4-dicarboxymethyl oct-5-enoate



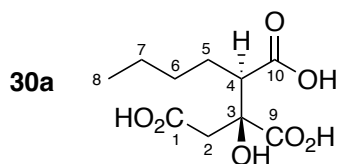
Dimethyl oxaloacetate **24** (1.0 equiv., 0.103 g, 0.6 mmol) and ester **27** (1.5 equiv., 0.200 g, 1.0 mmol) were dissolved in dry DMSO (3.2 mL). Zinc powder (2.0 eq., 84 mg, 1.3 mmol) was added and the reaction mixture was stirred for 48 h with exclusion of moisture. The solution was diluted with diethyl ether (22 mL) and washed with water (7 x 8 mL). The organic layer was dried (anh. Na₂SO₄) and evaporation of the solvent gave a crude product which was purified by preparative LCMS giving 40 mg of the desired product (21 % yield) as a mixture of diastereoisomers (*ca* 2:1 ratio).

major diastereomer ¹H NMR (600 MHz, Chloroform-d) δ 5.74 – 5.51 (4-CH, 5-CH₂, m, 2H), 3.81 (13-CH₃, s, 3H), 3.71 (CO₂CH₃, s, 3H), 3.67 (CO₂CH₃, s, 3H), 3.32 (3-CH, d, *J* = 9.5 Hz, 1H), 3.02 (11-CH₂, d, *J* = 16.3 Hz, 1H), 2.79 (11-CH₂, d, *J* = 16.3 Hz, 1H), 2.06 (6-CH₂, dt, *J* = 13.6, 6.9 Hz, 2H), 0.98 (7-CH₃, t, *J* = 7.5 Hz, 3H). ¹³C NMR (151 MHz, CDCl₃)

δ 173.5, 171.6, 170.9, 139.2, 121.6, 76.7, 57.6, 53.2, 52.4, 52.1, 41.9, 25.7, 13.5. LCMS 5.6 min, $m/z = 289.1$ [M + H⁺]

minor diastereomer ¹H NMR (600 MHz, Chloroform-d) δ 5.99 – 5.37 (4-CH, 5-CH₂, m, 2H), 3.80 (13-CH₃, s, 3H), 3.71 (CO₂CH₃, s, 3H), 3.67 (CO₂CH₃, s, 3H), 3.41 (3-CH, d, $J = 9.9$ Hz, 1H), 2.88 (11-CH₂, d, $J = 16.1$ Hz, 1H), 2.77 (11-CH₂, d, $J = 16.1$ Hz, 1H), 2.21 – 1.99 (6-CH₂, m, 2H), 1.01 (7-CH₃, t, $J = 7.5$ Hz, 3H). ¹³C NMR (151 MHz, CDCl₃) δ 174.2, 172.2, 170.8, 140.2, 120.8, 76.4, 56.1, 53.2, 52.4, 52.1, 41.1, 25.8, 13.5. LCMS 5.6 min, $m/z = 289.1$ [M + H⁺]

30a 3S*,4R* 3-hydroxy, 3,4-dicarboxy octanoic acid

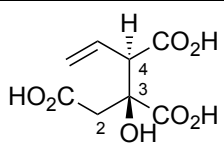
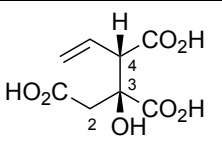


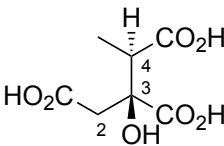
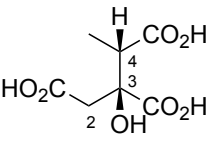
HCl-activated zinc dust (0.42g, 6.46 mmol) was suspended in dry DMSO (17 ml). Dimethyloxaloacetate **24** (1.034g, 6.46 mmol) and methyl 2-bromohexanoate **25** (1.05 ml, 6.46 mmol) were dissolved in dry DMSO (8 ml) and this solution added to the suspension of zinc dust with stirring at RT. The suspension was stirred at RT for 60 min, then warmed to 40 °C and stirred overnight. Saturated aqueous NH₄Cl (50 ml) was added slowly, then Et₂O (100 ml). The mixture was extracted with Et₂O (4 x 50 ml). The combined organic extracts were dried (anh. Na₂SO₄) and evaporated to afford a yellow oil (1.3 g). MS: ES⁺, 603 ([M₂]Na⁺, 15 %), 312.9 ([M]Na⁺, 45 %), 291 ([M]H⁺, 100 %).

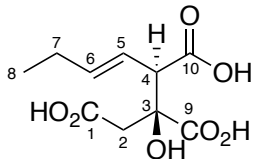
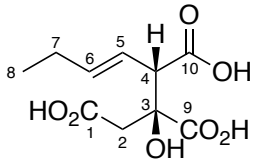
Trimethylbutylcitrate **29** (100 mg, 0.34 mmol) was suspended in aqueous HCl (3 M, 1.5 ml) and refluxed for 5 h. The cooled solution was extracted with Et₂O (2 x 3ml). The organic layers contained starting material which was recycled. The aqueous layer contained mainly the *anti* diastereomer. The desired triacid was purified by preparative LCMS (2 mg). MS: ESI⁻, 495.1 ([M₂ - H]⁻, 10 %), 247.0 ([M - H]⁻, 100 %), 229.1 ([M - H₂O - H]⁻, 5 %); ESI⁺, 497.1 ([M₂]H⁺, 100 %), 249.0 ([M]H⁺, 60%), 231 ([M - H₂O]H⁺, 65 %). ¹H NMR (500 MHz, D₂O): 0.78 (3H, t, $J = 6.9$ Hz, CH₃-8), 1.16 (2H, m, CH₂-7), 1.18 (1H, m, CH-6A), 1.25 (1H, m, CH-6B), 1.37 (1H, m, CH-5A), 1.66 (1H, m, CH-5B), 2.66, (1H, dd, $J = 3.2$ Hz, $J = 12.0$ Hz, CH-4), 2.67 (1H, d, $J = 16.5$ Hz, CH-2A), 3.06 (1H, d, $J = 16.5$ Hz, CH-2B). ¹³C NMR (125 MHz, D₂O): 13.0 (C-8), 21.6 (C-6), 26.2 (C-5), 29.0 (C-7), 41.7 (C-2), 53.7 (C-4), 76.1 (C-3), 174.1 (C-9), 177.0 (C-10), 177.1 (C-1).

4. Assignment of Stereochemistry

Both vinyl¹¹ and methyl¹² citrates have been synthesised and studied by ¹H NMR and x-ray crystallography. Chemical shifts of CH₂-2 and CH-4 are diagnostic. In the *anti* series H-2 resonates consistently at higher field and diastereotopic H-4 protons are consistently more separated. Selected NMR data is shown.

		 a anti, 3S, 4R		 b syn, 3S, 4S	
Position	δ_H / ppm	J / Hz	δ_H / ppm	J / Hz	
4	3.25 (d)	5.0	3.45 (d)	4.5	
2a	3.05 (d)	10.8	2.78 (s)	-	
2b	2.75 (d)	10.8			

		 a anti, 3S, 4R		 b syn, 3S, 4S	
Position	δ_H / ppm	J / Hz	δ_H / ppm	J / Hz	
4	2.86 (q)	7.0	2.99 (q)	7.2	
2a	2.89 (d)	16.2	2.89 (d)	16.0	
2b	3.08 (d)	16.2	3.01 (d)	16.0	

		 16a		 16b	
position	δ_H / ppm	J / Hz	δ_H / ppm	J / Hz	
4	3.38 (d)	9.8	3.51 (d)	9.7	
2a	3.14 (d)	16.5	2.91 (d)	15.8	
2b	2.87 (d)	16.5	2.85 (d)	16.0	

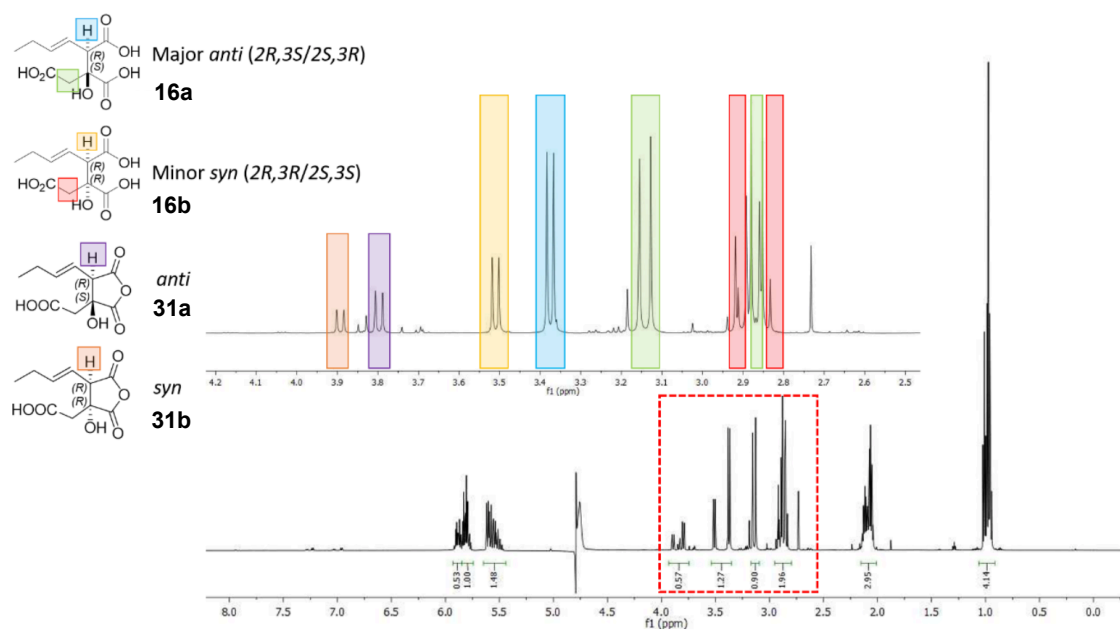


Figure S4.1. ^1H NMR (H_2O suppression) of a mixture of citrate diastereomers in D_2O

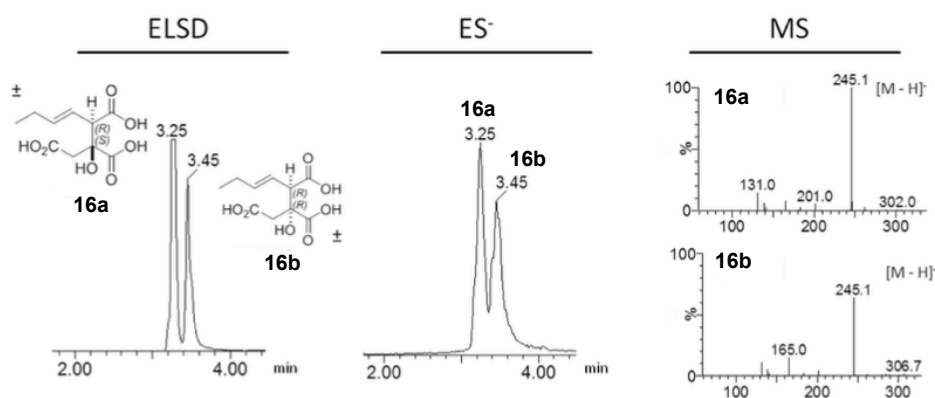


Figure S4.2. LCMS analysis of a synthetic mixture of citrate diastereomers.

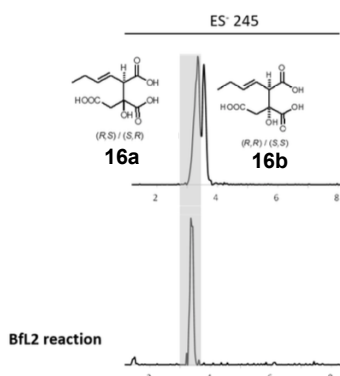


Figure S4.3. LCMS (EIC- 245) comparison of synthetic mixture of diastereomers with the product from BfL2 reaction.

5. LCMS Analysis and Purification

5.1 Analytical LCMS

LCMS data were obtained with using a Waters LCMS system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex column (2.6 μ , C₁₈, 100 Å, 4.6 \times 100 mm) equipped with a Phenomenex Security Guard precolumn (Luna C₅ 300 Å) eluted at 1 ml/min. Detection was by Waters 2998 Diode Array detector between 200 and 600 nm; Waters 2424 ELSD and Waters SQD-2 mass detector operating simultaneously in ES + and ES- modes between 100 m/z and 650 m/z . Solvents were: **A**, HPLC grade H₂O containing 0.05% formic acid; **B**, HPLC grade MeOH containing 0.045% formic acid; and **C**, HPLC grade CH₃CN containing 0.045% formic acid. Gradients were as follows. *Method 1*. Kinetex/CH₃CN: 0 min, 10% C; 10 min, 90% C; 12 min, 90% C; 13 min, 10% C; 15 min, 10% C.

5.2 Semi-Preparative LCMS and Compound Purification.

Purification of polar compounds was generally achieved using a Waters mass-directed autopurification system comprising of a Waters 2767 autosampler, Waters 2545 pump system, a Phenomenex Kinetex Axia column (5 μ , C₁₈, 100 Å, 21.2 \times 250 mm) equipped with a Phenomenex Security Guard precolumn (Luna C₅ 300 Å) eluted at 20 ml/min at ambient temperature. Solvent **A**, HPLC grade H₂O + 0.05% formic acid; Solvent **B**, HPLC grade CH₃CN + 0.045% formic acid. The post-column flow was split (100:1) and the minority flow was made up with HPLC grade MeOH + 0.045% formic acid to 1 ml·min⁻¹ for simultaneous analysis by diode array (Waters 2998), evaporative light scattering (Waters 2424) and ESI mass spectrometry in positive and negative modes (Waters SQD-2). Detected peaks were collected into glass test tubes. Combined tubes were evaporated (vacuum centrifuge), weighed, and residues dissolved directly in solvent for use or analysis.

6. Preparation and Purification of 1

Maleic anhydride monomer **1** was obtained at large scale by fermentation of *A. oryzae* transformants containing core genes for byssochlamic acid biosynthesis (Dr. K. Williams).¹³ In particular this included the genes PKS, hyd (*bfL1*), CS (*bfL2*), 2MCDH (*bfL3*) plus KSIs (*bfL6* and *10*) and PEBPs (*bfL5* and *9*). After 7 days of fermentation mature nonadrides and heptadrides were observed, but **1** was not present (Figure S6.1A).

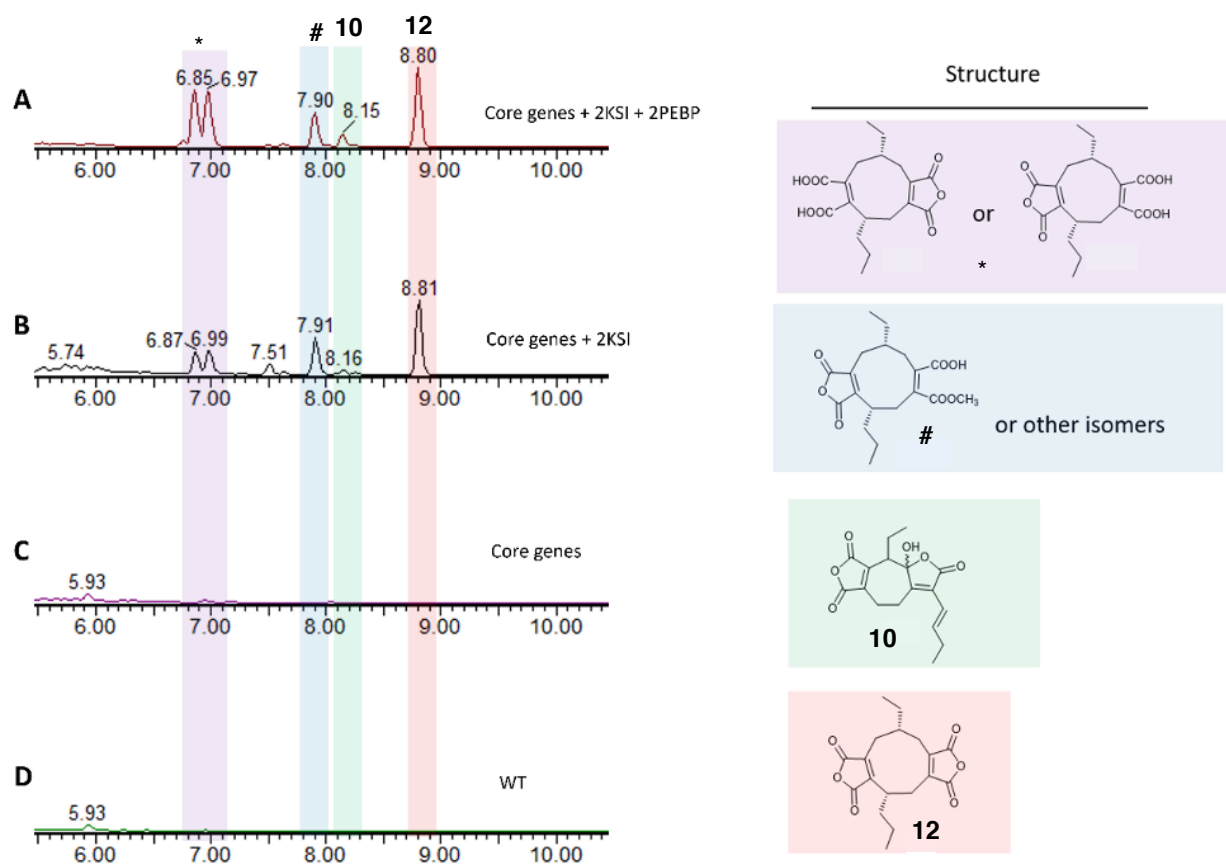


Fig S6.1A Chromatography of *A. oryzae* transformants with genes from byssochlamic acid pathway grown for 7 days: **A**, core genes with 2 KSI and 2 PEBP; **B**, core genes with 2 KSI; **C**, core genes; **D**, wild type. All these chromatograms show at the same scale.

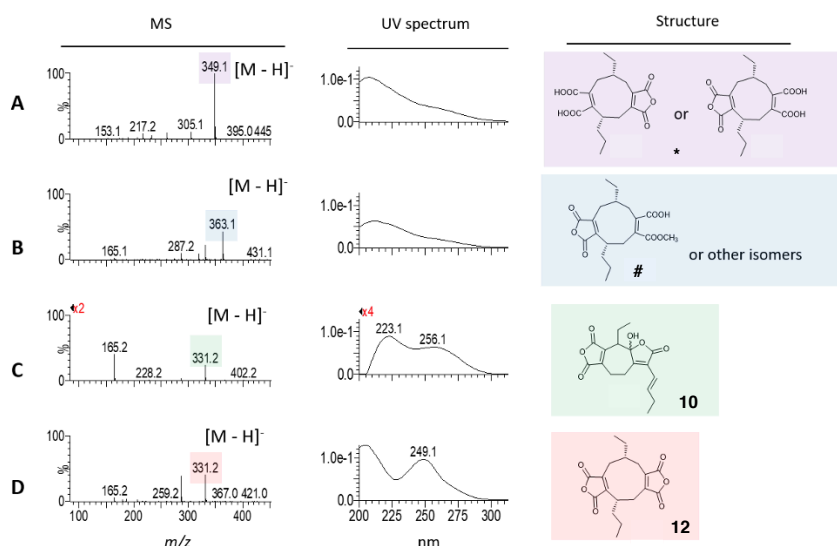


Fig S6.1B MS spectrum, UV spectrum and structures of the products: **A**, ring opened byssochlamic acid; **B**, methylated byssochlamic acid; **C**, agnestadride A **10**; **D**, byssochlamic acid **12**.

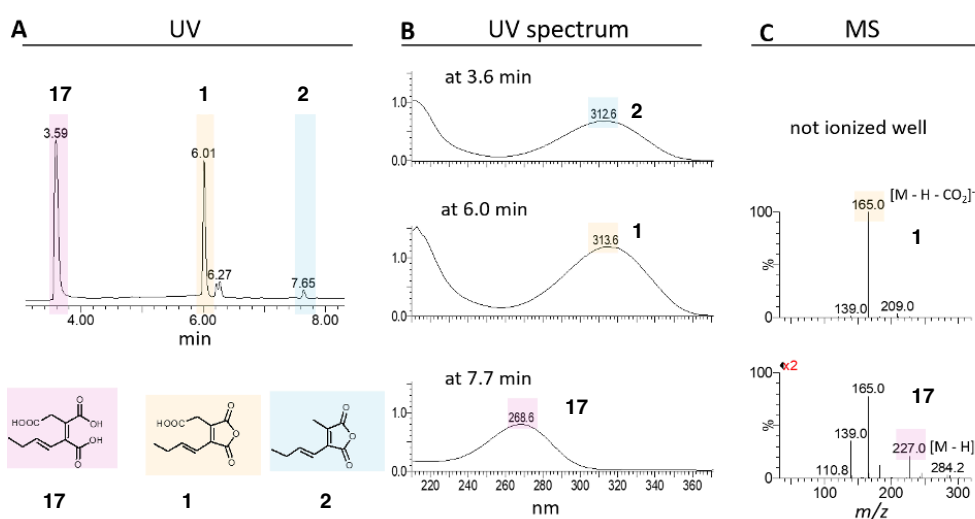


Fig. S6.1C LCMS analysis of metabolites extracted after 4 days: **A**, chromatogram; **B**, UV spectrum; **C**, MS.

However, fermentation for 4 days lead to good production of **1**. This was purified by LCMS. Under these conditions **1** was obtained pure in a mixture of H₂O and CH₃CN. These solvents were then directly removed by rotary evaporation at 5 mbar and 40 °C. After evaporation, 200 μL ddH₂O was added to give a maleic anhydride monomer **1** suspension. The suspension was analysed by LCMS. The compound **1** elutes in the sample at 6.0 min. The ring-opened citrate **17** elutes at 3.6 min (**Fig. S6.1A A**). The UV and MS spectra of these three compounds were identical to the previous study (**Fig. 67.1C B and C**). Therefore, this method could be used for making the substrate for the *in vitro* study of KSI and PEBP enzymes.

7. Protein Modelling

Multiple alignment and modelling of citrate synthase

As described in section 1.5.3, BfL2 is a citrate synthase like enzyme encoded by the byssochlamic acid BGC. BfL2 catalyses the formation of an alkyl citrate using oxaloacetate and polyketide CoA as substrate. *In vitro* assay of BfL2 showed that BfL2 could use the tetraketide CoA (**14d**, **15d**) and oxaloacetic acid to produce (*2R*, *3S* or *2S*, *3R*) distereoisomers (**16a/b** or **30a/b**). To additionally investigate the absolute configuration of the BfL2 product, protein modelling and multiple alignment of BfL2 was used to analyse the catalysis mechanism. The known CS crystal structures and sequence were used as models and templates.

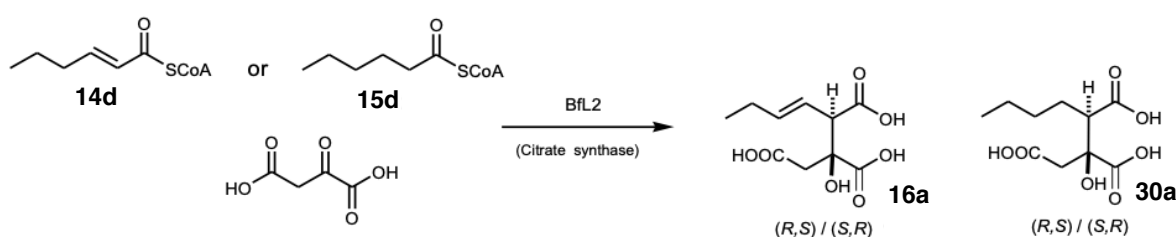


Figure S7.1A Proposed BfL2 reactions.

Karpusas *et al* obtained a crystal structure of *Thermus thermophilus* citrate synthase complexed with oxaloacetate and the unreactive alkyl CoA derivative carboxymethyl CoA.¹⁴ Based on this structure, the mechanism for the condensation reaction of citrate synthase was deduced.¹⁵ The overall CS reaction is thought to proceed through three partial reactions and involves both closed and open conformational forms of the enzyme. These steps begin with generation of the carbanion (or equivalent) from acetyl-CoA by base abstraction of a proton. The side-chain carboxylate of Asp-375 removes an acidic α proton from acetyl-CoA, while the side-chain NH of His-274 protonates the carbonyl oxygen, giving an enol. In the second step the nucleophilic attack of this carbanion on the *Si* face of oxaloacetate generates citryl-CoA. His-274 deprotonates the acetyl-CoA enol, which adds to the ketone carbonyl group in an aldol-like reaction. His-320 simultaneously protonates the oxaloacetyl carbonyl oxygen, producing *S*-citryl CoA. Finally, the hydrolysis of citryl-CoA to produce citrate and CoA is catalysed. *S*-citryl CoA is hydrolyzed to citrate by a typical nucleophilic acyl substitution reaction, catalyzed by the same citrate synthase enzyme (Figure S7.1B).¹⁶

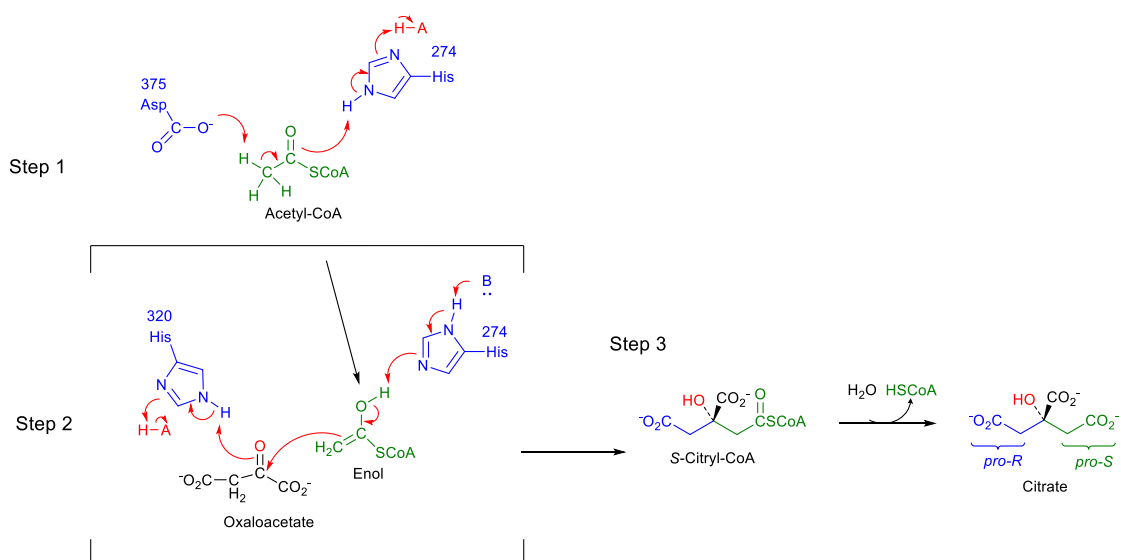


Figure S7.1B Catalytic mechanism of citrate synthase in primary metabolism.

In order to understand the reaction of the BfL2 citrate synthase better, we compared it to the well-studied primary metabolism citrate synthases from *T. thermophilus*, *Pyrococcus furiosus* and *E. coli* that have been previously studied by crystallography and site direct mutation.^{14,17,18}

According to the basic information of the CS from primary metabolism of *T. thermophilus*, His-274, His-320 and Asp-375 are the conserved residues in the active site. So the amino acid sequences of primary metabolic CS from *P. furiosus* and *E. coli* (crystal structure already known), as well as secondary metabolic CS from *B. fulva* (BfL2 of byssoclamic acid **12** pathway) and *P. variotii* (PvL7 of cornexstin **13** pathway) were chosen to make a multiple alignment with *ClustalX*. The result indicated that the three residues (**Fig. S7.1C** highlighted with “*”) in the active site are conserved in both primary and secondary metabolic CS. Furthermore, the residues around the active site are also highly conserved.

According to the result of multiple alignment with primary and secondary metabolic CS, the residues around the active site are conserved between these two kinds of CS. It means that the protein structures of secondary metabolic CS are likely to be similar to the primary ones. In order to look into the protein structure and the active site information of BfL2, protein modelling was carried out in the next step.

ANF07286.1	1	MSSGTLFVK--DSRTSLNYPEIPIHRNATAA--TAFKKIKAPVSGSDPADKVDGGLRVHDP-----GLQNTTVVET	66
AAA31017.1	1	MALLTAARLFGAKNASCLVLAARHASASTNLKDILADLIPKEQARIKTFRQQHGNTVVGQITVDMMYGMRGMKGLVY	80
WP_153827436.1	1	MADTKAKLTLNGDTAVEL-----DVLKGTGLGQDVIDIRTLGSK-----GVFTFDP--GFTSTASCES	55
BBL93869.1	1	-----ME-----VARGLLEGVLFTEES	15
AAB35835.2	1	-----MNTEKYLAKGLEVDVYIDQT	19
ANF07286.1	67	DISFSNSDSGLLLFRGYSLDQLWDS-----DFEELFHLVWGKYPTRVQKDDLSNTLANVMKNVPEVNVKAIRSF	136
AAA31017.1	81	ETSVLDPDEG--IRFRGYSIPECQKMLPKAKGGEEPLPEGLFWLLVVGQIPTEEQVSWLSKEWAK--RAALPSHVVTMLDNF	158
WP_153827436.1	56	KITFDIDGDEGILLHRGFPIDQLATD-----SNYLEVCYIILNGEKPTQEQYDEFKTTVTR--HTMIHEQITRFLHAF	125
BBL93869.1	16	RMICYIDGQQKLYYYGPIQELAEK-----SFEEITFLHLHGRLLRRQEELEEFSAALAR--RRALPAHLLSFKRY	85
AAB35835.2	20	NICYIDGKEGKLYYRGYSVEELAEK-----STPEEVVYLLWVGKLPSSLSELENFKKELAK--SRGLPKVEIIMEAL	89
ANF07286.1	137	PRSTSPMPMVIAGLAAYIASDPESIPAANG-----GNIYQGNTAQTDLGLKTVSAYAVVLGLVASHRKDIPF--VPA	207
AAA31017.1	159	PTNLHPMSQLSAAITALNSESINFARAYAEGIHRTKYWELIYE--DCMDLIAKLPVCAAK-----IYRNLYREGSSIGAI	230
WP_153827436.1	126	RRDSDHPAMVCMGITGALAAFYHDSLDVNNP-----RHREI--AAFRLLSKMPMTMAAM-----CCKHSIGQPF--VYP	188
BBL93869.1	86	PVSAHPMSFLRTAVSEFGMLDPT--EGDISR-----EALYE--KGLDLIAKFATIVAA-----NKRLKEGKEP--IPP	147
AAB35835.2	90	PKNTHPMGALRTIISYLGNIIDSDGPIVTP-----EEVYR--IGISVTAKIPTIVAN-----WYRIKNGLEY--VPP	152
ANF07286.1	208	SSENSYENLFIIMGMVDRVSGRPDTLQLSCFRRFAALSCDGM--ALSVFATLVCASSLADPISCLISALAAAYGPLIFG	286
AAA31017.1	231	DSKLDWSHNFTNMLGYTDAQ-----FTELMRLYLTIHSDTEGGNVSAHTSHLVGSALSADPFLSFAAAMNGLAGPLGL	303
WP_153827436.1	189	RNDLSYAGNFLNMFSTPCEPYEVNPIILERAMDRILILHADTEQ--NASTSTVTRTAGSSGANPFACIAAGIASLWGPAIGG	267
BBL93869.1	148	REDLSHAANFLYMANGVEPSPE-----QARLMDAALILHAEVGF--NASTFTAIAAFSTETDLYSAITAAVASLKGPRIGG	221
AAB35835.2	153	KEKLSHAANFLYMLHGEEPPKE-----WEKAMDVALILYAEVEI--NASTLAVMTVGSSTLSDYSAIAGLIGALKGPIIGG	226
ANF07286.1	287	ATEAAHRAHQEIGSV--ERVPDFLEQVKRGERK----LFTYGHVYKGTDPVPIPKK--LLEDNSATSNPL--IEIAKS	356
AAA31017.1	304	ANQEVLVWLTLQKQKEVGKDVSDKLRDYIWNLTNSGRVVPVYGHVLRKTDPRYTCQREFALKHLPHPMPFK--LVAQLYK	382
WP_153827436.1	268	ANEALAMLEEISSV--KHIFEVRRRAKDKNDS--FRLMFFSHVYKNDPRATVMRETCHVSLKELGTKDLLEVAME	342
BBL93869.1	222	ANEAVMRMIQEIETP--ERAREVWREKLAKKER----IMDMGHVYKAFDPRAGVLEKLARLVAEKHGHSK--EYQILKI	293
AAB35835.2	227	AVEEAIKQFMEIGSP--EKVEEWFKALQQRK----IMDACHVYKTYDPRARIFPKYA----SKLGDKK--LPEIAER	294
ANF07286.1	357	IEIHASTDDYFKSGLSAIAAFYGNVVFSAIGFDP--DFIPVAMLAQIIGIMAHWREYMLKRGKLFPSHITYTGNTEPL	434
AAA31017.1	383	IVPNVLEQGGK--ANPWPVVAHSGVLLQYYGMTEMNYYTVLFGVSAALGVLAQLIWSRALGFPLEPKSMSTDGL--IKL	460
WP_153827436.1	343	LENIALNDPYFIEKLYPNVIFYSGIILKAMGIPS--SMFTVIFAMATVGVIAHWSMHSDGMKIASPRQLYTGYE--KRD	420
BBL93869.1	294	VEEAGKVLNP--GIYPNVIFYSGVVYSDLGFSL--EFFTPIFAVATISGVVGHILEYQELDNRLLEPGAKYIGELDVY	370
AAB35835.2	295	LERLVEEYLSK--GISIVVYVWGLVFGMKIPI--ELYTTIFAMGATAGWTAHLAEVYVSHN--RIIPRLQVYGEIGKKY	370
ANF07286.1	435	CNFSPKL 441 ANF07286.1 <i>B. fulva</i> * Catalytic Residue	
AAA31017.1	461	VDSK-- 464 AAA31017.1 <i>S. scrofa</i> Oxaloacetate / Acyl Coa Binding	
WP_153827436.1	421	FKSDIKR 427 WP_153827436.1 <i>E. coli</i>	
BBL93869.1	371	VPLEARE 377 BBL93869.1 <i>T. thermophilus</i>	
AAB35835.2	371	LPIELRR 377 AAB35835.2 <i>P. furiosus</i>	

Fig. S7.1C Multiple alignment of primary metabolic and secondary metabolic CS sequences.

By using online software SWISS-MODEL¹⁹ and BLASTp, the closest structural model for BfL2 was found. A structure of citrate synthase (PDB ID: 2H12)²⁰ complexed with oxaloacetate (OAA) and carboxymethyldehia coenzyme A (CMX) from *Acetobacter aceti* was chosen as the structural model. Taking crystal structure of 2H12 as the template, a BfL2 protein structure model was built. The 2H12 and the BfL2 modelling structure were displayed and analysed on the software *PYMOL*.²¹ The BfL2 model has a very similar structure as the 2H12 which is a dimer structure. The tertiary structure of the BfL2 (in green) and 2H12 (in blue) were highly similar in general (**Fig. S7.1D A**). In particular, the active site residues around the bound substrates OAA and CMX were very similarly located between the two structures (**Fig. S7.1D B**). In addition, the residues around the OAA and CMX within 4 Å were displayed. The result showed that these active site residues were very closely overlapped between the two structures (**Fig. S7.1D C and D**). The key residues (His-284, His-323 and Asp-377) in BfL2 modelling are also very similarly located with 2H12 structure (**Fig. S7.1D E**). According to the polar contacts (yellow line) between substrate OAA / CMX and the active site residues, substrates are very likely to be bound in identical orientations. In particular, activation of oxaloacetate is likely to be identical in both cases. So the stereochemistry of BfL2 is extremely likely to be the same as the primary CS 2H12, resulting in an *S*-centre at the newly formed alcohol.

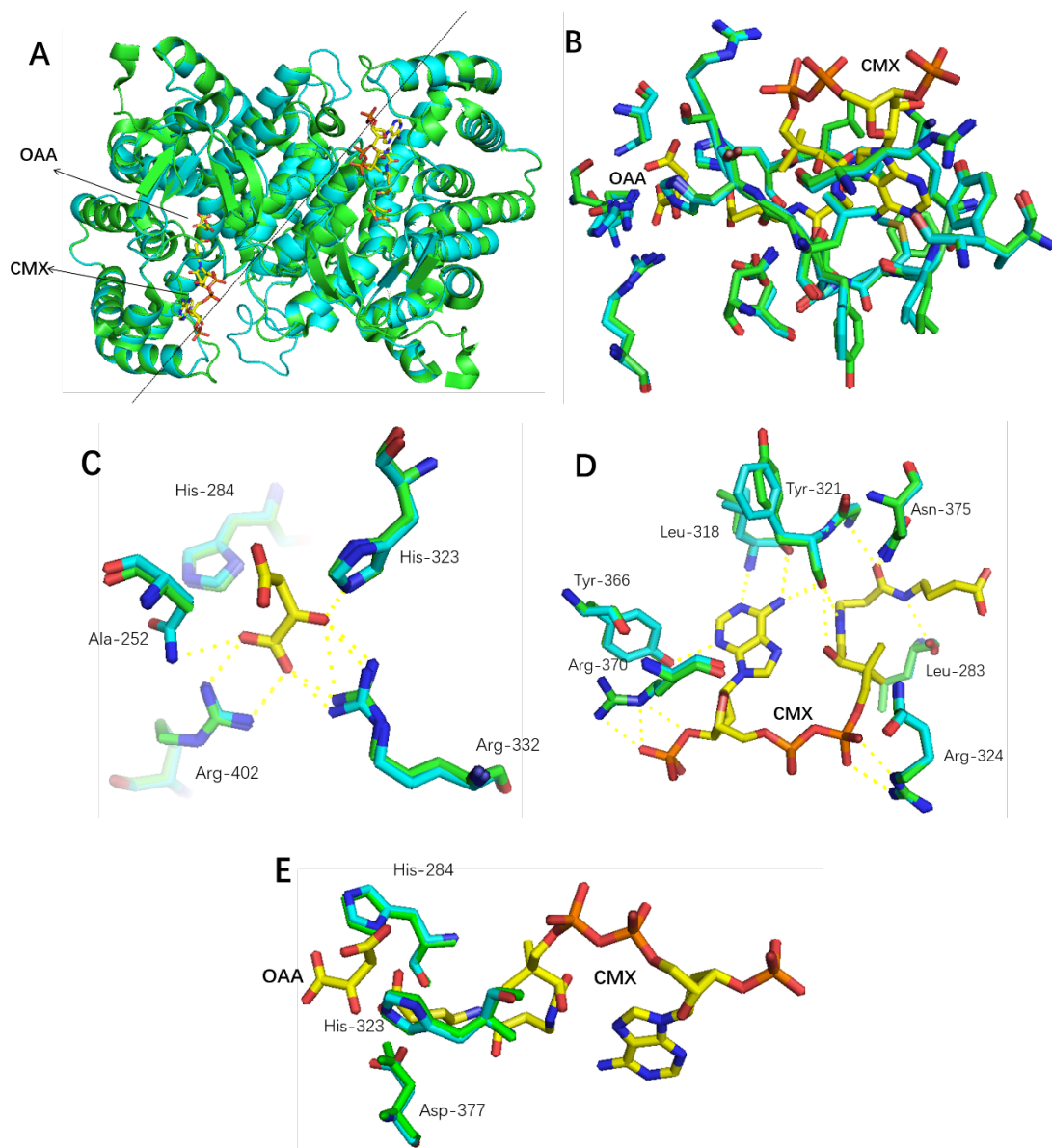


Fig. S7.1D Modelling of Bfl2 (green) in byssochlamic acid pathway and structure of 2H12: **A**, alignment of Bfl2 and 2H12 structures; **B**, the residues around the OAA and CMX within 4 Å; **C**, residues around the OAA; **D**, residues around the CMX; **E**, key residues.

8. References

1. Dolle, Roland E., et al. "Preparation of (±)-(erythro)-and (±)-(threo)-2-vinyl citric acids as potential mechanism-based inhibitors of ATP-citrate lyase." *Tetrahedron letters* 32.35 (1991): 4587-4590.
2. Reddick, Jason J., et al. "First biochemical characterization of a methylcitric acid cycle from *Bacillus subtilis* strain 168." *Biochemistry* 56.42 (2017): 5698-5711.
3. Krawczyk, Hanna, and Tomasz Martyniuk. "Characterisation of the ¹H and ¹³C NMR spectra of methylcitric acid." *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 67.2 (2007): 298-305.
4. Szwalbe, Agnieszka J., et al. "Novel nonadride, heptadride and maleic acid metabolites from the byssochlamic acid producer *Byssochlamys fulva* IMI 40021—an insight into the biosynthesis of maleidrides." *Chemical Communications* 51.96 (2015): 17088-17091.
5. Williams, Katherine, et al. "Heterologous production of fungal maleidrides reveals the cryptic cyclization involved in their biosynthesis." *Angewandte Chemie International Edition* 55.23 (2016): 6784-6788.
6. D. M. Roberts, C. Bartel, A. Scott, D. Ivison, T. J. Simpson and R. J. Cox, *Chemical Science*, 2016, **8**, 1116–1126.
7. J. Wunderlich, T. Roß, M. Schröder and F. Hahn, *Org. Lett.*, 2020, **22**, 4955–4959.
8. B. B. Bond-Watts, A. M. Weeks and M. C. Y. Chang, *Biochemistry*, 2012, **51**, 6827–6837.
9. M. I. Kim, S. J. Kwon and J. S. Dordick, *Org Lett*, 2009, **11**, 3806–3809.
10. L. Ziesche, J. Rinkel, J. S. Dickschat and S. Schulz, *Beilstein J Org Chem*, 2018, **14**, 1309–1316.
11. Dolle, R. E.; Novelli, R.; Saxty, B. A.; Wells, T. N. C.; Kruse, L. I.; Camilleri, P.; Eggleston, D. *Tetrahedron Lett.* **1991**, 32 (35), 4587–4590.
12. Krawczyk, H.; Martyniuk, T. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2007**, 67 (2), 298–305.
13. K. Williams, A. J. Szwalbe, N. P. Mulholland, J. L. Vincent, A. M. Bailey, C. L. Willis, T. J. Simpson and R. J. Cox, *Angewandte Chemie Int Ed*, 2016, **55**, 6784–6788.
14. Karpusas, Mihail, Bruce Branchaud, and S. James Remington. "Proposed mechanism for the condensation reaction of citrate synthase: 1.9-Å structure of the ternary complex with oxaloacetate and carboxymethyl coenzyme A." *Biochemistry* 29.9 (1990): 2213-2219.
15. Wiegand, Georg, and Stephen J. Remington. "Citrate synthase: structure, control, and mechanism." *Annual review of biophysics and biophysical chemistry* 15.1 (1986): 97-117.
16. Weitzman, P. D. J., and Michael J. Danson. "Citrate synthase." *Current topics in cellular regulation*. Vol. 10. Academic Press, 1976. 161-204.
17. Russell, Rupert JM, et al. "The crystal structure of citrate synthase from the hyperthermophilic archaeon *Pyrococcus furiosus* at 1.9 Å resolution." *Biochemistry* 36.33 (1997): 9983-9994.
18. Man, Wai-Jin, et al. "The binding of propionyl-CoA and carboxymethyl-CoA to *Escherichia coli* citrate synthase." *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology* 1250.1 (1995): 69-75.
19. Guex, Nicolas, and Manuel C. Peitsch. "SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling." *electrophoresis* 18.15 (1997): 2714-2723.
20. Ferraris, Davide M., et al. "Structures of citrate synthase and malate dehydrogenase of *Mycobacterium tuberculosis*." *Proteins: Structure, Function, and Bioinformatics* 83.2 (2015): 389-394.
21. DeLano, Warren L. "Pymol: An open-source molecular graphics tool." *CCP4 Newsletter on protein crystallography* 40.1 (2002): 82-92.