Supplementary Information

An enzyme cascade for green synthesis of chrysanthemol and chrysanthemic acid: key precursors for pyrethrin and pyrethroids

Zhanyue Liao,[a] Chuncen Chen,[a] Junyi Lu,[a,c] Donghua Du,[b] Huailong Teng,[b] Yi Zhou,[a] Shuke Wu*[a]

a. National Key Laboratory of Agricultural Microbiology, College of Life Science and Technology, Huazhong Agricultural University, 1 Shizishan Street, Wuhan 430070, P.R. China. E-mail: shukewu@mail.hzau.edu.cn

b. College of Chemistry, Huazhong Agricultural University, 1 Shizishan Street, Wuhan 430070, P.R. China.

c. Current Address: School of Biochemistry, University of Bristol, University Walk, Bristol, BS8 1TD, UK

General information

Ethyl laurate, prenol, and (1*R*,3*R*)-chrysanthemic acid were purchased from Shanghai Haohong Scientific Co., Ltd. (Shanghai, China). Isoprenol, NAD^+, NADH, methyl tert-butyl ether (MTBE), and (±)-chrysanthemic acid were purchased from Shanghai Macklin Biochemical Technology Co., Ltd. (Shanghai, China). Ethyl chrysanthemate was purchased from Shaoyuan Technology Co., Ltd. (Shanghai, China). Sodium pyruvate was purchased from Bide Pharmatech Co., Ltd. (Shanghai, China). phosphoenolpyruvate (PEP) and ATP were purchased from Shanghai Dibo Biotech Co., Ltd. (Shanghai, China). pyruvate kinase (PK) was purchased from J&K Chemical Co., Ltd. (Shanghai, China). Other chemicals and reagents are purchased from commercial suppliers and used without further purification.

The genes of *Sp*LicA, *Sf*PhoN, *Sm*DAGK, *Mt*IPK, *Tc*ADH2, *Tc*ADH2low, *Tc*CDS were synthesized by SynBio Technologies (Suzhou, China) with codon optimization for *E. coli*. Other genes are cloned from the genome of microbes or the cDNA of plants (*Arabidopsis thaliana*, *Tanacetum cinerariifolium*).

Construction of Plasmids Used in This Study

Detailed cloning steps are described in the Molecular Biology Methods section within the Method section of the article. The details of primers

Linearized pETDuet, pACYCDuet, and pRSFDuet vectors were cloned using primers Duet-XhoI-F and Duet-1stATG-R. Linearized pET28a was cloned using primers pet28aChis-F and pet28a-R (for C-terminal His-tag), or pET28a-F and pET28aNhis-R (for N-terminal His-tag).

The *Ec*ThiM gene in pRSFDuet-*Ec*ThiM-CHis was cloned from the *Escherichia coli* genomic library using primers RS-C-Thim-F and RS-C-Thim-R.

The *Bs*ThiM gene in pRSFDuet-*Bs*ThiM-CHis was cloned from the *Bacillus subtilis* genomic library using primers RS-C-BsThim-F and RS-C-BsThim-R.

The *Sp*LicA gene in pRSFDuet-*Sp*LicA-CHis was cloned from synthesized *Sp*LicA gene using primers RS-C-LicA-F and RS-C-LicA-R.

The *Sc*CK gene in pRSFDuet-*Sc*CK-CHis was cloned from the *Saccharomyces cerevisiae* genomic library using primers RS-C-ScCK-F and RS-C-ScCK-R.

The *Sf*PhoN gene in pRSFDuet-*Sf*PhoN-CHis was cloned from synthesized *Sf*PhoN gene using primers RS-C-PhoN-F and RS-C-PhoN-R.

The *Bce*ThiM gene in pET28a-*Bce*ThiM-CHis was cloned from the *Bacillus cereus* genomic library using primers ET-C-Bce-F and ET-C-Bce-R.

The *Bli*ThiM gene in pET28a-*Bli*ThiM-CHis was cloned from the *Bacillus licheniformis* genomic library using primers ET-C-Bli-F and ET-C-Bli-R.

The *Cbu*ThiM gene in pET28a-*Cbu*ThiM-CHis was cloned from the *Clostridium butyricum* genomic library using primers ET-C-Cbu-F and ET-C-Cbu-R.

The *Ehu*ThiM gene in pET28a-*Ehu*ThiM-CHis was cloned from the *Enterobacter huaxiensis* genomic library using primers ET-C-Ehu-F and ET-C-Ehu-R.

The *Gst*ThiM gene in pET28a-*Gst*ThiM-CHis was cloned from the *Geobacillus stearothermophilus* genomic library using primers ET-C-Gst-F and ET-C-Gst-R.

The *Sma*ThiM gene in pET28a-*Sma*ThiM-CHis was cloned from the *Serratia marcescens* genomic library using primers ET-C-Sma-F and ET-C-Sma-R.

The *Sty*ThiM gene in pET28a-*Sty*ThiM-CHis was cloned from the *Salmonella typhimurium* genomic library using primers ET-C-Sty-F and ET-C-Sty-R.

The *Mt*IPK gene in pETDuet-IPK was cloned from the synthesized *Mt*IPK gene using primers PET-Ipk-F and KLD-IPK-R.

The *Sma*ThiM gene in pETDuet-*Sma*ThiM-IPK was cloned from pET28a-*Sma*ThiM-CHis using primers RS-Smathim-F and RBS-Smathim-R. The IPK gene was cloned from pETDuet-IPK using primers IDI-IPK-RBS-F and Duet-1stATG-R.

The *Bs*ThiM gene in pETDuet-*Bs*ThiM-IPK was cloned from pRSFDuet-*Bs*ThiM-CHis using primers RS-Bsthim-F and RBS-Bsthim-R. The IPK gene was cloned from pETDuet-IPK using primers IDI-IPK-RBS-F and Duet-1stATG-R.

The *Sty*ThiM gene in pETDuet-*Sty*ThiM-IPK was cloned from pET28a-*Sty*ThiM-CHis using primers RS-Stythim-F and RBS-Stythim-R. The IPK gene was cloned from pETDuet-IPK using primers IDI-IPK-RBS-F and Duet-1stATG-R.

The *Sm*DAGK gene in pETDuet-*Sm*DAGK-IPK was cloned from synthesized *Sm*DAGK gene using primers Duet-DAGKop-F and Duet-DAGKop-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Bce*ThiM gene in pETDuet-*Bce*ThiM-IPK was cloned from pET28a-*Bce*ThiM-CHis using primers Duet-Bce-F and Duet-Bce-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Bli*ThiM gene in pETDuet-*Bli*ThiM-IPK was cloned from pET28a-*Bli*ThiM-CHis using primers Duet-Bli-F and Duet-Bli-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Cbu*ThiM gene in pETDuet-*Cbu*ThiM-IPK was cloned from pET28a-*Cbu*ThiM-CHis using primers Duet-Cbu-F and Duet-Cbu-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Ehu*ThiM gene in pETDuet-*Ehu*ThiM-IPK was cloned from pET28a-*Ehu*ThiM-CHis using primers Duet-Ehu-F and Duet-Ehu-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Gst*ThiM gene in pETDuet-*Gst*ThiM-IPK was cloned from pET28a-*Gst*ThiM-CHis using primers Duet-Gst-F and Duet-Gst-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Ec*ThiM gene in pETDuet-*Ec*ThiM-IPK was cloned from pRSFDuet-*Ec*ThiM-CHis using primers Duet-Ec-F and Duet-Ec-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Sf*PhoN gene in pETDuet-*Sf*PhoN-IPK was cloned from pRSFDuet-*Sf*PhoN-CHis using primers Duet-PhoN-F and Duet-PhoN-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Sp*LicA gene in pETDuet-*Sp*LicA-IPK was cloned from pRSFDuet-*Sp*LicA-CHis using primers Duet-LicA-F and Duet-LicA-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Sc*CK gene in pETDuet-*Sc*CK-IPK was cloned from pRSFDuet-*Sc*CK-CHis using primers Duet-ScCK-F and Duet-ScCK-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

The *Tc*ALDH1 gene in pRSFDuet-*Tc*ALDH1 was cloned from the *Tanacetum cinerariifolium* cDNA library using primers RS-ALDH3-F and RS-ALDH3-R.

The *Tc*ADH2 gene in pRSFDuet-*Tc*ADH2-*Tc*ALDH1 was cloned from synthesized *Tc*ADH2 gene using primers RS-ADH-F and RS-ADH-ALDH-R; The *Tc*ALDH1 gene was cloned from pRSFDuet-*Tc*ALDH1 using primers RSALDH-F and Duet-1stATG-R.

The *Tc*ADH2 gene in pET28a-NHis-*Tc*ADH2 was cloned from synthesized *Tc*ADH2 gene using primers ET-His-ADH-F and ET-ADH-R.

The *Tc*ADH2low gene in pRSFDuet-*Tc*ADH2low-*Tc*ALDH1 was cloned from synthesized *Tc*ADH2low gene using primers RS-ADH3-F and RS-ADH3-R. The *Tc*ALDH1 gene was cloned from pRSFDuet-*Tc*ADH2-*Tc*ALDH1 using primers RSALDH4-F and Duet-1stATG-R.

The *Tc*Nudix gene in pRSFDuet-*Tc*Nudix was cloned from the *Tanacetum cinerariifolium* cDNA library using primers RS-Nudix-F and RS-Nudix-R.

The *Tc*CDS gene in pRSFDuet-*Tc*CDS was cloned from synthesized *Tc*CDS gene using primers RS-CDS-F2 and RS-CDS-R2.

The *Tc*Nudix gene in pRSFDuet-*Tc*Nudix-*Tc*CDS was cloned from pRSFDuet-*Tc*Nudix using primers RS-Nudix2-F and RS-Nudix2-R; The *Tc*CDS gene was cloned from pRSFDuet-*Tc*CDS using primers RSCDS-F and Duet-1stATG-R.

The *Tc*CDS gene in pRSFDuet-*Tc*CDS-*Tc*Nudix was cloned from pRSFDuet-*Tc*CDS using primers Duet-1stATG-F and CDS-Nudix-R; The *Tc*Nudix gene was cloned from pRSFDuet-*Tc*Nudix using primers CDS-Nudix-F and Duet-1stATG-R.

The *Tc*Nudix-*Tc*CDS gene in pACYCDuet-*Tc*Nudix-*Tc*CDS and pCDFDuet-*Tc*Nudix-*Tc*CDS was cloned from pRSFDuet-*Tc*Nudix-*Tc*CDS using primers Duet-1stATG-F and Duet-XhoI-R.

The *Tc*CDS-*Tc*Nudix gene in pACYCDuet-*Tc*CDS-*Tc*Nudix and pCDFDuet-*Tc*CDS-*Tc*Nudix was cloned from pRSFDuet-*Tc*CDS-*Tc*Nudix using primers Duet-1stATG-F and Duet-XhoI-R.

The *Tc*ADH2low-*Tc*ALDH1 gene in pRSFDuet-Sumo*Tc*ADH2low-*Tc*ALDH1 was cloned from pRSFDuet-*Tc*ADH2low-*Tc*ALDH1 using primers RSAAM-ADHm-F and RSAAM-R. The Sumo tag gene was cloned from commercially available plasmid using primers RSAAM-SUMO-F and RSAAM-SUMO-R.

The *Tc*ADH2low-*Tc*ALDH1 gene in pRSFDuet-Gst*Tc*ADH2low-*Tc*ALDH1 was cloned from pRSFDuet-*Tc*ADH2low-*Tc*ALDH1 using primers RSAAM-ADHm-F and RSAAM-R. The Gst tag gene was cloned from commercially available plasmid using primers RSAAM-Gst-F and RSAAM-Gst-R.

The *Tc*ADH2low-*Tc*ALDH1 gene in pRSFDuet-Trx*Tc*ADH2low-*Tc*ALDH1 was cloned from pRSFDuet-*Tc*ADH2low-*Tc*ALDH1 using primers RSAAM-ADHm-F and RSAAM-R. The Trx tag gene was cloned from the *Escherichia coli* genomic library using primers RSAAM-Trx-F and RSAAM-Trx-R.

The *Tc*ADH2-*Tc*ALDH1 gene in pRSFDuet-Sumo*Tc*ADH2-*Tc*ALDH1 was cloned from pRSFDuet-*Tc*ADH2-*Tc*ALDH1 using primers RSAAS-ADHs-F and RSAAM-R. The Sumo tag gene was cloned from commercially available plasmid using primers RSAAM-SUMO-F and RSAAS-SUMO-R.

The *Tc*ADH2-*Tc*ALDH1 gene in pRSFDuet-Gst*Tc*ADH2-*Tc*ALDH1 was cloned from pRSFDuet-*Tc*ADH2-*Tc*ALDH1 using primers RSAAS-ADHs-F and RSAAM-R. The Gst tag gene was cloned from commercially available plasmid using primers RSAAM-Gst-F and RSAAS-Gst-R.

The *Tc*ADH2-*Tc*ALDH1 gene in pRSFDuet-Trx*Tc*ADH2-*Tc*ALDH1 was cloned from pRSFDuet-*Tc*ADH2-*Tc*ALDH1 using primers RSAAS-ADHs-F and RSAAM-R. The Trx tag gene was cloned from the *Escherichia coli* genomic library using primers RSAAM-Trx-F and RSAAS-Trx-R.

The trigger factor gene in pCDFDuet-TriggerFactor was cloned from the *Escherichia coli* genomic library using primers Duet-TFactor-F and Duet-TFactor-R.

The GroES-GroEL gene in pCDFDuet-GroES-GroEL was cloned from the *Escherichia coli* genomic library using primers Duet-GroESL-F and Duet-GroESL-R.

The DnaK-DnaJ gene in pCDFDuet-DnaK-DnaJ was cloned from the *Escherichia coli* genomic library using primers Duet-DnaKJ-F and Duet-DnaKJ-R.

The ClpB gene in pCDFDuet-ClpB was cloned from the *Escherichia coli* genomic library using primers Duet-clpB-F and Duet-clpB-R.

pACYCDuet-*Tc*Nudix-$(GSG)_3$-*Tc*CDS was generated from pACYCDuet-*Tc*Nudix-*Tc*CDS by inverse PCR using primers (GSG)3-CDS-F and Nudix-R.

pACYCDuet-*Tc*Nudix-$(PT)_4$P-*Tc*CDS was generated from pACYCDuet-*Tc*Nudix-*Tc*CDS by inverse PCR using primers (PT)4P-CDS-F and Nudix-R.

pACYCDuet-*Tc*CDS-$(GSG)_3$-*Tc*Nudix was generated from pACYCDuet-*Tc*CDS-*Tc*Nudix by inverse PCR using primers (GSG)3-CDS-F and CDS-R.

pACYCDuet-*Tc*CDS-$(PT)_4$P-*Tc*Nudix was generated from pACYCDuet-*Tc*CDS-*Tc*Nudix by inverse PCR using primers (PT)4P-Nudix-F and CDS-R.

The *Tc*CDS gene in pET28a-NHis-*Tc*CDS was cloned from pACYCDuet-*Tc*Nudix-*Tc*CDS using primers 28a-NHis-CDS-F and 28a-CDS-R.

The *Tc*Nudix gene in pET28a-*Tc*Nudix-CHis was cloned from pACYCDuet-*Tc*Nudix-*Tc*CDS using primers 28a-Nudix-F and 28a-Nudix-CHis-R.

The *Mt*IPK gene in pET28a-NHis-IPK was cloned from pETDuet-IPK using primers 28a-NHis-IPK-F and 28a-IPK-R.

The *Tc*CDS gene in pET28a-NHis-*Tc*CDS was cloned from pACYCDuet-*Tc*Nudix-*Tc*CDS using primers 28a-NHis-CDS-F and 28a-CDS-R.

The *Tc*CDS-$(PT)_4$P-*Tc*Nudix gene in pRSFDuet-Sumo*Tc*CDS-$(PT)_4$P-*Tc*Nudix was cloned from pACYCDuet-*Tc*CDS-$(PT)_4$P-*Tc*Nudix using primers TAG-CDS-F and Duet-Nudix-R. The Sumo tag gene was cloned from pRSFDuet-Sumo*Tc*ADH2-*Tc*ALDH1 using primers SUMO-CDS-R and Duet-XhoI-F.

The *Tc*CDS-$(PT)_4$P-*Tc*Nudix gene in pRSFDuet-Gst*Tc*CDS-$(PT)_4$P-*Tc*Nudix was cloned from pACYCDuet-*Tc*CDS-$(PT)_4$P-*Tc*Nudix using primers TAG-CDS-F and Duet-Nudix-R. The Gst tag gene was cloned from pRSFDuet-Gst*Tc*ADH2-*Tc*ALDH1 using primers GST-CDS-R and Duet-XhoI-F.

The Sumo*Tc*CDS-$(PT)_4$P-*Tc*Nudix gene in pACYCDuet-Sumo*Tc*CDS-$(PT)_4$P-*Tc*Nudix was cloned from pRSFDuet-Sumo*Tc*CDS-$(PT)_4$P-*Tc*Nudix using primers Duet-1stATG-F and Duet-XhoI-R.

The Gst*Tc*CDS-$(PT)_4$P-*Tc*Nudix gene in pACYCDuet-Gst*Tc*CDS-$(PT)_4$P-*Tc*Nudix was cloned from pRSFDuet-Gst*Tc*CDS-$(PT)_4$P-*Tc*Nudix using primers Duet-1stATG-F and Duet-XhoI-R.

The *Tc*Nudix gene in pACYCDuet-*Tc*Nudix-Sumo*Tc*CDS was cloned from pACYCDuet-*Tc*Nudix-*Tc*CDS using primers Duet-XhoI-F2 and Nudix-RBS-R. The Sumo*Tc*CDS gene was cloned from pRSFDuet-Sumo*Tc*CDS-$(PT)_4$P-*Tc*Nudix using primers RBS-SUMO-F and Duet-CDS-R.

The *At*FKI gene in pETDuet-*At*FKI-IPK was cloned from the *Arabidopsis thaliana* cDNA library using primers Duet-ATFKI-F and ATFKI-IPK-R. The IPK gene was cloned from pETDuet-*Sma*ThiM-IPK using primers IK-IPK-F and Duet-1stATG-R.

Gene sequence

> Sequences encoding *Sm*DAGK[1]

ATGCCGATGGATCTGCGCGACAACAAACAGTCCCAGAAAAAATGGAAAAACCGC
ACCCTGACCAGTAGTCTGGAATTTGCACTGACCGGTATCTTCACCGCGTTCAAAGA
AGAGCGCAACATGAAAAAACACGCGGTTAGCGCTCTGCTGGCGGTGATTGCGGGT
CTGGTTTTTAAAGTCTCCGTCATTGAGTGGCTGTTTCTGCTGCTGAGCATCTTCCTG
GTCATCACCTTCGAGATCGTCAACAGCGCGATTGAAAACGTCGTCGATCTGGCAA
GCGATTATCACTTCAGCATGCTGGCGAAAAACGCCAAAGATATGGCAGCAGGTGC
GGTTCTGGTTATTAGCGGTTTTGCAGCACTGACCGGTCTGATCATCTTCCTGCTGAA
AATCTGGTTCCTGCTGTTCCACTAA

> Sequences encoding *Bs*ThiM

ATGGATGCACAATCAGCAGCAAAATGTCTTACGGCTGTCCGCCGGCATAGCCCACT
GGTGCATAGCATAACCAACAATGTCGTAACGAATTTCACAGCAAACGGCCTGCTCG
CGCTCGGCGCATCGCCCGTTATGGCGTACGCAAAAGAAGAGGTCGCCGATATGGC
GAAAATTGCGGGTGCACTCGTTTTAAATATCGGAACACTGAGCAAGGAGTCAGTC
GAAGCGATGATCATCGCGGGAAAATCAGCTAATGAACATGGCGTTCCCGTCATTCT
TGATCCTGTCGGTGCCGGAGCAACACCGTTCCGCACTGAATCGGCACGTGACATC
ATTCGTGAGGTGCGCCTTGCTGCAATCAGAGGAAATGCGGCGGAAATTGCCCATAC
CGTCGGCGTGACCGATTGGCTGATCAAAGGTGTTGATGCGGGTGAAGGTGGAGGC
GACATCATCCGGCTGGCTCAGCAGGCGGCACAAAAGCTAAACACGGTCATTGCGA
TAACTGGTGAAGTTGATGTCATAGCCGACACGTCACATGTATACACCCTTCATAAC
GGCCACAAGCTGCTGACAAAAGTGACAGGCGCCGGTTGCCTGCTGACTTCCGTCG
TCGGTGCGTTTTGCGCTGTGGAAGAAAATCCATTGTTTGCTGCTATTGCGGCCATTT
CTTCGTATGGGGTCGCCGCTCAGCTTGCCGCACAGCAGACGGCTGACAAAGGCCC
TGGAAGCTTTCAGATTGAATTGCTGAACAAGCTTTCAACTGTTACTGAACAAGAC
GTCCAAGAATGGGCGACTATAGAAAGGGTGACTGTCTCATGA.

> Sequences encoding *Ec*ThiM[2]

ATGCAAGTCGACCTGCTGGGTTCAGCGCAATCTGCGCACGCGTTACACCTTTTTCA
CCAACATTCCCCTCTTGTGCACTGCATGACCAATGATGTGGTGCAAACCTTTACCG
CCAATACCTTGCTGGCGCTCGGTGCATCGCCAGCGATGGTTATCGAAACCGAAGAG
GCCAGTCAGTTTGCGGCTATCGCCAGTGCCTTGTTGATTAACGTTGGCACACTGAC
GCAGCCACGCGCTCAGGCGATGCGTGCTGCCGTTGAGCAAGCAAAAAGCTCTCA

AACACCCTGGACGCTTGATCCAGTAGCGGTGGGTGCGCTCGATTATCGCCGCCATT
TTTGTCATGAACTTTTATCTTTTAAACCGGCAGCGATACGTGGTAATGCTTCGGAAA
TCATGGCATTAGCTGGCATTGCTAATGGCGGACGGGGAGTGGATACCACTGACGCC
GCAGCTAACGCGATACCCGCTGCACAAACACTGGCACGGGAAACTGGCGCAATCG
TCGTGGTCACTGGCGAGATGGATTATGTTACCGATGGACATCGTATCATTGGTATTC
ACGGTGGTGATCCGTTAATGACCAAAGTGGTAGGAACTGGCTGTGCATTATCGGCG
GTTGTCGCTGCCTGCTGTGCGTTACCAGGCGATACGCTGGAAAATGTCGCATCTGC
CTGTCACTGGATGAAACAAGCCGGAGAACGCGCAGTCGCCAGAAGCGAGGGGCC
AGGCAGTTTTGTTCCACATTTCCTTGATGCGCTCTGGCAATTGACGCAGGAGGTGC
AGGCATAA

> Sequences encoding *Sf*PhoN[3]

ATGTCTATTCCACCAGGTAACGACGTAACTACAAAGCCAGATTTATATTATTTGACT
AACGACAACGCTATTGACTCATTAGCCCTTCTTCCGCCGCCACCTCAGATCGGTTC
CATTGCCTTCCTTAACGACCAGGCCATGTATGAAAAGGGTCGCTTGTTGCGGAACA
CTGAGCGGGGTAAGCTCGCTGCTGAGGACGCTAACCTGTCAAGCGGCGGGGTGG
CGAACGTCTTTAGCGCCGCTTTCGGCAGCCCAATCACTGCCAAGGATTCCCCAGA
GCTTCACAAGTTACTGACTAATATGATTGAAGATGCAGGCGATTTGGCGACCCGCT
CGGCAAAAGAGTACTACATGCGTATCCGCCCATTCGCTTTTTACGGTGTGAGTACC
TGTAACACTAAGGAGCAAGACACATTGAGCCGTAATGGCTCTTATCCATCAGGCCA
TACTAGTATTGGTTGGGCGACAGCCTTGGTGCTGTCCGAAATCAATCCGGCACGTC
AAGACACAATTCTCAAGCGTGGCTACGAATTGGGGGACTCTCGCGTAATCTGTGG
GTATCACTGGCAGAGCGATGTAGATGCCGCCCGTATTGTTGGTTCAGCTATCGTGG
CGACATTACATTCTAATCCAGTTTTCCAGGCACAATTACAAAAAGCAAAAGACGAA
TTTGCTAATAACCAGAAGAAGTAA

> Sequences encoding *Sc*CK[4]

ATGGTACAAGAATCACGTCCAGGGAGTGTAAGAAGTTACTCGGTCGGTTACCAAG
CAAGGTCCAGATCGAGTTCTCAAAGAAGACATTCGTTAACACGCCAACGTTCCTC
GCAAAGACTGATTAGAACCATCAGTATCGAGTCTGATGTGTCTAATATTACTGACG
ATGACGATTTGAGAGCTGTCAATGAGGGAGTAGCGGGTGTGCAACTGGACGTCTC
TGAAACCGCAAATAAGGGACCAAGAAGAGCATCAGCAACTGATGTCACAGATAGT
TTGGGTTCGACTTCGTCGGAATATATTGAGATTCCCTTTGTTAAGGAAACATTGGAT
GCAAGTTTACCTTCGGATTATCTGAAGCAGGACATATTAAATCTCATTCAGAGTTTG
AAGATATCCAAATGGTATAACAACAAGAAAATCCAACCGGTAGCACAAGATATGAA
CTTAGTCAAGATCTCTGGTGCGATGACAAACGCAATTTTCAAAGTTGAATACCCTA

AGTTACCATCGTTGCTATTGAGAATATACGGACCGAATATTGATAATATCATTGACAG
GGAATATGAATTGCAGATTTTGGCTAGGCTTTCATTGAAAAATATAGGTCCTTCCCT
TTACGGCTGTTTTGTAAACGGTAGATTTGAGCAGTTTCTGGAGAATTCTAAGACTT
TAACAAAAGACGACATTAGAAACTGGAAGAACTCTCAAAGGATTGCAAGGAGAA
TGAAGGAGTTACATGTAGGTGTTCCTCTCTTGAGTTCAGAAAGGAAGAACGGGTC
GGCTTGTTGGCAAAAGATTAACCAGTGGTTGCGCACGATTGAAAAAGTCGACCAA
TGGGTGGGGGATCCTAAAAACATTGAAAACTCTTTATTATGTGAGAATTGGTCCAA
GTTTATGGATATTGTCGATAGATATCACAAGTGGCTTATTTCTCAAGAACAGGGTAT
AGAGCAAGTCAACAAAAATCTTATATTCTGCCATAATGATGCCCAATACGGCAATTT
ACTTTTCACTGCTCCTGTGATGAACACACCGAGCCTATACACTGCACCTTCGTCTA
CATCATTGACTTCCCAATCAAGTTCCTTATTTCCTTCGAGCTCCAATGTCATTGTAG
ATGATATAATCAACCCGCCAAAGCAGGAGCAAAGCCAAGATTCCAAATTGGTCGTC
ATTGATTTTGAATATGCAGGTGCCAATCCCGCCGCATATGATTTAGCGAATCATCTTT
CCGAGTGGATGTATGATTACAACAATGCTAAGGCCCCACATCAGTGCCACGCTGAT
AGATATCCCGATAAAGAACAGGTTTTGAATTTCTTATACTCTTATGTTTCGCATCTAA
GGGGTGGTGCTAAGGAACCCATAGATGAAGAGGTTCAAAGACTCTATAAGTCAAT
CATTCAATGGAGACCCACTGTACAACTATTTTGGTCGCTCTGGGCCATCCTACAAA
GTGGTAAATTAGAGAAAAAAGAAGCCTCCACTGCCATCACTAGAGAAGAAATTGG
ACCCAATGGAAAAAAATATATCATCAAGACTGAACCCGAATCCCCTGAAGAAGAC
TTTGTTGAAAATGACGACGAGCCTGAAGCTGGCGTCAGCATTGACACGTTCGATTA
TATGGCTTATGGTCGTGACAAGATTGCGGTCTTTTGGGGCGACCTTATTGGCTTAGG
CATAATCACCGAAGAAGAATGCAAAAATTTCAGCTCTTTCAAGTTCCTCGATACTA
GTTATTTGTAA

> Sequences encoding *Sp*LicA

ATGGAAAAGATCATCAAAGAGAAGATTAGCAGCCTGCTGAGCCAAGAAGAAGAA
GTTCTGAGCGTTGAACAGTTAGGTGGTATGACCAATCAGAATTATCTGGCCAAAAC
CACCAACAAACAGTACATCGTTAAATTCTTTGGCAAAGGCACCGAGAAACTGATT
AATCGTCAGGATGAAAAGTACAACCTGGAACTGCTGAAAGATTTAGGTCTGGATG
TGAAAAACTACCTGTTCGATATTGAAGCCGGTATCAAAGTGAACGAATATATTGAA
AGCGCGATTACCCTGGATAGCACCAGCATTAAAACCAAATTCGATAAAATCGCACC
GATCCTGCAAACCATTCATACCAGCGCAAAAGAACTGCGTGGTGAATTTGCACCGT
TTGAAGAGATCAAAAAATACGAAAGCCTGATCGAAGAACAAATCCCGTATGCAAA
TTATGAAAGCGTTCGTAATGCAGTGTTCAGCCTGGAAAAACGTCTGGCCGATCTGG
GTGTTGATCGTAAAAGCTGTCATATTGATCTGGTGCCGGAAAACTTTATTGAAAGT
CCGCAGGGTCGTCTGTATCTGATTGATTGGGAATATAGCAGCATGAATGATCCGATG
TGGGATTTAGCAGCACTGTTTCTGGAAAGCGAATTTACCAGTCAAGAAGAGGAAA

CCTTTCTGAGCCATTATGAGAGCGATCAGACACCGGTTAGCCATGAAAAAATTGCC
ATCTACAAAATCCTGCAAGACACCATTTGGAGCCTGTGGACCGTTTATAAAGAAGA
ACAGGGCGAAGATTTTGGTGATTATGGTGTTAATCGTTATCAGCGTGCAGTTAAAG
GTCTGGCAAGCTATGGTGGTAGTGATGAAAAATAA

> Sequences encoding *Bce*ThiM

ATGAATATGAAAGAAATTAGTAAAGTAGTGGATTTGGTGAGAGAATCTAATCCGCT
TGTTCATAATATTACAAATGTTGTTGTAACAAATTTTACTGCTAACGGTTTGTTGGC
ACTTGGGGCATCGCCTGTAATGGCGTATGCAAAAGAAGAAGTGGCAGAAATGGCT
AGCATTGCTGGAGCGTTAGTATTAAATATGGGAACACTTCGTCCTGAAGAAGTAGA
AGCGATGTTACTTGCCGGTAAATCAGCAAATGTGAACAATGTACCAGTACTGTTTG
ATCCAGTTGGTGCAGGAGCAACATCGTATCGGACAGAAGTTGCGAGACATATTCCG
GCTGAAATTGAATTAGCGAGTATTCGCGGAAATGCAGCTGAAATAGCGAACGTTAT
TAATGAGAGATGGGAAATTAAAGGGGTAGACGCTGGTACTGGAAATGGCAATGTT
GTAAGTATTGCAAGGCAGGCAGCAGATGAACTAAATACAGTTGCGGTCATTACTGG
AAAAGAAGATGTTGTTACAGATGGAGAGAGAACTATTGTTATTCGAAATGGTCACC
CTATTTTAACGAAAGTTACGGGAACAGGTTGTTTACTAACTTCTGTAATAGGAGCAT
TTGTAGCAGTGGAAAAAGATTATGTAAAAGCAGCAGTAGCAGCATTAACGTTTTAT
GGTGTAGCTTCGGAACTGGCAGCTGCTAAGACCGTGGAAAAGGGACCTGGTAGTT
TCCAAATTGAATTTTTAAATCAGTTAGCGAATACTACTTCTAGTGATATTGAGAAGT
ATGGAAAGATTGAGGAGTTAGAGTAA

> Sequences encoding *Bli*ThiM

ATGAATATGGAAGATGCCGCAAGGGGAATTGAGTTAGTCAGAAAGCAAAAACCGC
TCGTACACAATATGACAAACAATGTCGTCACCAACTTTACGGCAAACGGTTTGCTG
GCTCTCGGCGCATCGCCGGTGATGGCGTATGCGAGGGAAGAGGCCGCCGACATGG
CAAAGATCGCCGGAGCCCTTGTGCTGAACATCGGCACGCTCAGCCGTGAATCTGT
TGAAGCGATGATTATTGCAGGAAAGTCAGCTAATGAACATGGGGTGCCGGTTATTT
TTGATCCGGTCGGTGCGGGGGCGACGCCGTTTCGAACGGAATCGGCTCAAGCGAT
CATGCGCAAAGTGAAGGTTTCTGCCGTCAGAGGCAATGCTGCGGAAATTGCCAAC
ACGCTTGGGGAAGACTGGCTGATCAAAGGCGTGGACGCGGGTGAAGAGAGCGGC
GACAGAACGGAGCTCGCAAAAAAAGCCGCGAAATTATGGGATACAGCAGTCATCA
TAACCGGTGCTGAAGACGTGATAACGGATGGTACAAAAACGTATACCGTGGGCAA
CGGACATCAGCTGCTTACGAAAGTGACGGGGACAGGCTGCTTGTTTACCTCTGTC
ATCGGTGCGTTTTGCGCCGTTGAAAAAGACGTTTGCCGTGCCGCTGTATCTGCAGC
CGCTTTTTATGGAACAGCGGCCGAGTTGGCCGCGGCGAAAACGGAAAGCCGGGG

GCCCGGAAGTTTCCAAATCGAGTTTTTGAATCAGCTTGCTGCCGTTGAAGCTATGG
ATATAAAAGAACGCGGCAGGATCAGAGAGGTGGAATAA

> Sequences encoding *Cbu*ThiM

ATGAATATAGATATAGCACAAAAGGTAGTAGAGCTATTAAACAGATTAAAAAATAA
AAAACCATTGATTCATAACATAACTAATTATGTTACAGTAAATGATTGTGCTAATATA
CTTCTTGCAATAGGTGCATCACCAATAATGGCGGATGACTTAAAAGAAGCAGAAG
ATATTACATCAATTGCATCAGCACTTGTAATAAATATTGGAACTTTAAATGAAAGAA
CTATAGAATCAATGATTGCATCAGGGAAAAAGGCTAATGAATTGAATATTCCCGTTG
TACTAGATCCGGTTGGAGCGGGTGCATCATTATTTAGAAATGAAACAACTAAAAGA
ATTCTTGAAGAAATAAAAATCAGTGTTTTACGTGGAAATATGTCGGAAATAAAATTT
ATTGCTGGATTAGAGTCTGAAACAAAGGGTGTGGATGCATCAGAGTCAGATTTAAA
AAGTGATAGTGATGAAGGCATAAGAGTAGCTAAATATTTAGCAAAAAGATTTAATT
GTACAGTAGCTATAACTGGAGTATGTGATATAGTTTCTGATGGAGAAAAATCTGTTA
CCATAGAAAATGGAACAAAAATGCTTTCTAATGTTACAGGGACTGGATGTATGACC
ACAGCTTTAGTTGGTGGGTATTTAGGAGCTTACGAAACAAAAGAAGATTTATTTAT
AGCTGCTGTTTCAGGAATAGTATCTATGGGAATCTGTGGAGAAATTGCAGAAGAAA
GAGCTGGAAGTATAGGTCTTGGAAGTTTTCATATGGCAATAATAGATGCAGTAAGT
AATTTAGATGAAGAAGATTTGCTAAAAAGAAGTAAAATTAAATAA

> Sequences encoding *Ehu*ThiM

ATGCAGCCTGACCTGCTCGATCTACACGTTTTACATCAGTTCCGAACCCGTTCCCC
GCTTACCCACTGTATGACTAACAATGTCGTACAAACCTTTACCGCCAATGTCCTGCT
CGCACTTGGCGCCTCGCCGGCGATGGTGATTGAAGCCGAAGAAGCCGAACAGTTT
GCGACGCTTGCCGATGCGCTGCTGATCAACGTCGGTACGCTGACCGCACCGCGCG
CCCAGTCGATGCGTCGGGCAATAGAAAGTGCGGTGGCGGCAGGCAAACCCTGGA
CGCTCGATCCGGTTGCCGTTGGCGCGCTCGCCTTTCGCACCCGCTTTTGTCATCAA
ATTCTCTCCCTTAAGCCTGCCGCAATTCGTGGTAACGCCTCGGAAATTCTCGCCCTT
GCAGGGATGAGCGCGGGCGGTCGCGGCGTAGATACAACCGATACGGCGGCCAGC
GCGCTGCCTGCCGCCCAGGCACTGGCGCGCCAGATCAACGCCATTGTGGTGGTAA
CCGGCGAGGTGGATTACATCACCGACGGGCAGAGAACCCGCACCGTATCGGGGGG
CGATCCGATAATGACCCGCGTTGTCGGCACAGGCTGCGCGCTCTCCGCCGTCGTTG
CGGCCTGCTGTTCGCTGCCCGGCGACCGGCTGGATAATATCGCCGCCGCCTGCGGA
TGGATAAAGCGTGCCGGTACGGTCGCCGTGGCGCACTCCCGCGGTCCCGGCAGCT
TCGCCAGCGCGTTTCTGGACGCGCTTTATACGCTGGAGGAGCAGGCATGA

> Sequences encoding *Sma*ThiM

ATGATCGCTCGACCTGATGCTCTACCCGGCGCCCGCGCCGCGGCCTGTCTGACCCA
ATTCCAACAGCGGCCTCCGCTGATCCACTGCCTGACCAACGAAGTGGTGCAGTCG
CTGACTGCCAACGTGCTGCTGGCGCTCGGCGCCTCCCCGGCGATGGTGGTGGAGC
CGCAAGAGGCGGCGCAGTTCAGCGCGCTTGCCGACGGCCTGCTGATCAACGTCGG
TACCCTCAACGCGCTGCGCGCCGAATCGATGCGGGCGGCGGTGGCCGCCGCCAAT
CAGGCCGGCAAGCCCTGGACGCTGGATCCGGTGGCGGTCGGCGCGCTCTCCTACC
GTACCGCCTTCGCCAAACAGCTGCTGGACGAAAAACCGGCGGCGATACGCGGCA
ACGCCTCGGAAATCATGGCGCTCAACGGGTTGCTGGCCAGCGGCCGCGGCGTGGA
CAGCGGCGACGATTCCCTGGCGGCATTGCCCGCCGCGCGTGAGCTGGCGCGGCGC
AGCGGCGCGGTGGTGGCGGTCACCGGCGCGGTGGATTACGTCACCGACGGCCAA
CGCGACTGGGCGATAGAAGGCGGCTCCCCGCTGATGACGCGGGTGGTGGGCACC
GGCTGCGCGCTGTCGGCGGTGGCGGCGGCGTTTTGCGCCCTGCCGGGCGATCGCC
TGGACAACGTGGCGACCGCCTGTCGGGTGATGTCACACTGCGGCGGCCTGGCGGC
CAGGCAGGCCGCCGGGCCGGGCAGCTTTACGCCGGCGTTCCTCGACGCCCTGTAT
CAACTGCGCGGAGAGGATCTGTGA

> Sequences encoding *Gst*ThiM

ATGGCCGATTGGCGTGAATGGGCGGAACTGCGCGAGCGGGTGAGAGAGGCCAAT
CCGCTCGTCCACAATATGACGAATGTCGTCGTGACGAATTTTACGGCGAACGGATT
GCTGGCGCTTGGCGCTTCGCCGGTGATGGCGTATGCGAAAGAGGAAGTGGCAGAG
ATGGCGAAGCTTGCCGGCGCGCTCGTATTGAACATCGGGACGCTGAACGCTGAGG
AAGTAGAGGCGATGTTGATCGCCGGCCGGGCGGCGAACGAAGCCGGTGTGCCGG
TGGTGTTCGACCCGGTCGGAGCCGGGGCCACGCGCTATCGAACGGAGACAGCGC
GCCGCATCGCCGAACAGGTGAAGCTTGCAGCGGTCCGCGGCAATGCGGCGGAAAT
CGCCAATATGATCGGCGAGACGTGGGCGATTAAAGGCGTCGACGCCGGGGAAGGA
AGCGGCGATGCCGCGGCTTTGGCGAAGCGGGCGGCGGCCAAGCTTGGCACCGTT
GTGATCATCACCGGAAAGGACGATGTCGTCACGGACGGAAACACGACGTATTTGG
TTCACAACGGCCACCCGCTGTTGACGAAAGTGACCGGGACGGGCTGTTTGTTGAC
GTCGGTGATTGGGGCGTTTGCCGCTGTTGAGCGTGATATGGTGAAGGCAGCGGTG
GCGGCGCTCGCCTATTATGGCGTCGCGGCGGAGCGGGCGGCTGCAGAAGCTGGGG
AGCGCAGTCCGGGCCGTTTTCAGTCGTTGTTTCTAGACGCGTTGGCGCACACAAG
CGCCGATGATGTGAAGCGTGACGGACAGGTCGAACAACGGTAA

> Sequences encoding *Sty*ThiM

ATGCAGCCTGACCTGCACTGCCGCACGCTTGCGGCGCATACGTTAAAACACTTTCG
CGCGCTCTCCCCGCTTACGCACTGTATGACGAATGACGTCGTACAAACGTTTACCG
CCAATACGTTGCTGGCGCTGGGCGCTTCACCCGCCATGGTGATTGATCCTGTCGAG
GCCAGACCGTTTGCCGCCATCGCCAACGCCTTGCTGATTAATGTCGGAACATTAAC
TGCCTCACGCGCTGACGCGATGCGTGCGGCGGTAGAAAGCGCTTATGATGCCAAA
ACGCCGTGGACGCTTGATCCTGTCGCGGTGGGCGCGCTGGAATTTCGTCGGCGATT
TTGCCTGGATCTCTTGTCCCTGCGCCCGGCGGCAATACGCGGCAACGCCTCGGAA
ATCCTGGCGTTATCCGGCATGGCGCTGGGCGGACGTGGCGTAGATACCACCGAGG
CGGCGTTGGCCGCACTGCCTGCGGCGCAGGCGCTGGCGCGTCAGATAGACTGCAT
CGTTGTGGTTACCGGAGAGATAGATTACGTCACTAATGGTCAGCGCACGCTGAGCA
TTCCCGGCGGCGATCCGTTAATGACTCGCATTGTAGGCACCGGCTGCGCGTTGTCG
GCGGTCGTCGCCGCCAGTTGCGCGTTACCGGGCGCCGCGCTGGACAATGTCGCGT
CGGCCTGCTGCTGGATGAAACTGGCTGGACAGGCCGCGGCAGAGCGTAGCGAAG
GACCGGGTAGCTTCATCCCGGCTTTTCTTGATGCGCTCTATCATCTGGATGTGGAGG
CGGCCAATGCAACGAATTAA

> Sequences encoding *Mt*IPK

ATGATTATTCTGAAACTTGGTGGTAGCGTTATTACCCGTAAAGATAGCGAAGAACC
GGCAATTGATCGTGATAATCTGGAACGTATTGCAAGCGAAATTGGTAATGCAAGCC
CGAGCAGCCTGATGATTGTTCATGGTGCAGGTAGCTTTGGTCATCCGTTTGCCGGT
GAATATCGTATTGGTTCCGAAATTGAAAACGAAGAGGATCTGCGTCGTCGTCGTTT
TGGTTTTGCACTGACCCAGAATTGGGTGAAAAAACTGAATAGCCATGTTTGTGATG
CACTGCTGGCAGAAGGTATTCCGGCAGTTAGCATGCAGCCGAGCGCATTTATTCGT
GCACATGCAGGTCGTATTAGCCATGCAGATATTAGCCTGATTCGTAGCTATCTGGAA
GAAGGCATGGTTCCGGTTGTTTATGGTGATGTTGTTCTGGATAGCGATCGTCGTCTG
AAATTTAGCGTGATTAGCGGTGATCAACTGATCAATCATTTTAGCCTGCGTCTGATG
CCGGAACGTGTTATTCTGGGCACCGATGTTGATGGTGTTTATACCCGTAATCCGAAA
AAACATCCTGATGCACGTCTGCTGGATGTTATTGGTAGCCTGGATGATCTGGAAAG
TCTGGATGGCACCCTGAATACCGATGTTACCGGTGGTATGGTTGGCAAAATTCGTG
AACTGCTGCTGCTGGCCGAAAAAGGTGTTGAAAGCGAAATCATTAATGCAGCCGT
TCCGGGTAATATTGAACGTGCCCTGCTGGGTGAAGAAGTTCGCGGTACACGTATTA
CAGGTAAACACTAA

> Sequences encoding *Tc*Nudix

ATGGCGATGACAGTTGGTTTAGGAACTCGTGGGATCTCATTAGTTGTGGATCGCGG
TCTAACCTCGGGGAAGAATGCTAAATCATCGCAAACTGTTTACAAATGCATTAATTC
GTCTCAACAGCGTTGTGTTATGCGACCCGTTAAGTCCTTGATGAAGATTGTTGAAC
AAAACAAGGAACGAGCATTTTCTCCAAATCGTTCGATCTTTCAAACTTCGGATACT
AGACTAATAGATGTGGAAGCATCATCTGTGGAGAATGCCGCAAAGTCTCGAGTGC
CAGAGGTTGGGGTTACGGTGTTCATAGTAAAAGACGACAAGATACTACTAGGACG
TCGCGGCTCAACACTCGTTGTTGGCAACAGTTTTCACCTCCCTGGCGGACACCTTG
AGTACGGGGAAAGTTTTGAGGAATGTGCAACAAGAGAAGTGAAGGAGGAAACAG
GACTAGACATAAATAATTTAAAGGTGGTAACAATAGTCAACCATGTCCTTGCTGATA
TACATGCTGTTGTTGTTTACATGCGCGCAAATTTATCAGACCCAAATCAGACGCCG
CAAACTATTGAACCTCAAAGGTGTGAAGGATGGGATTGGTATGACTTGGAAAACC
TACCCGAACCAATGTTTCAACCATTGAAAGAATTGCTACAAGCTGGTAGTTTCAAT
ATTTTCACTACTCATTCTTGA

> Sequences encoding *Tc*CDS

ATGACTACGACATTGAGCAGCAATCTGGACTCACAATTCATGCAAGTTTATGAGAC
TTTGAAATCTGAGCTGATTCATGACCCGTCATTTGAGTTTGATGACGATTCTCGTCA
GTGGGTGGAGCGTATGATTGACTACAATGTACCTGGTGGTAAGATGGTCCGTGGCT
ATTCTGTTGTTGACAGCTACCAATTGCTTAAAGGTGAAGAATTGACGGAAGATGAA
GCTTTTCTCGCGTGTGCTCTGGGTTGGTGCACTGAATGGCTGCAAGCCTTTATTCT
GGTCCTGGATGACATTATGGATGGCTCGCACACACGTCGTGGTCAACCCTGTTGGT
TTCGTCTGCCCGAGGTTGGTGTAGTTGCTATTAATGATGGTGTTCTGCTGCGCAACC
ATGTGCATCGCATCCTGAAGAAATATTTCCAAGGTAAGCCTTATTACGTGCATCTTC
TGGACCTCTTCAATGAGACCGAATTTCAAACAATCTCTGGCCAAATGATTGATACA
ATCTGTCGTCTGGCTGGCCAAAAAGATCTCTCAAAGTATACTATGACTCTTAACCGT
CGCATTGTTCAGTACAAAGGTTCTTACTACTCATGTTACCTGCCAATTGCGTGTGCA
CTGCTGATGTTTGGAGAGAATCTGGAAGATCATGTTCAAGTGAAAGACATCCTTGT
AGAATTGGGTATGTATTATCAAATTCAGAATGATTATCTCGACACTTTTGGTGATCCT
GATGTTTTTGGCAAGACGGGTACAGATATTGAAGAATGCAAGTGTTCATGGTTGAT
TGCAAAAGCACTGGAACTTGCCAACGAGGAACAAAAGAAAATTTTAAGCGAAAA
CTATGGGATCAACGATCCATCAAAGGTAGCAAAAGTGAAGGAATTATACCATGCTC
TGGATCTGAAGGGTGCGTATGAAGATTATGAGACAAATCTGTATGAGACGTCAATG
ACATCAATTAAAGCTCATCCAAACATTGCAGTGCAAGCGGTGTTGAAATCTTGTCT
GGAAAAGATGTATAAGGGACATAAGTAA

> Highly codon-optimized sequences encoding *Tc*ADH2

ATGAGCCTGAACACCCCGGATGTGATTATTTGCAAAGCGGCGGTGGTGCGCGAAC
TGGGCCGCAGCGTGATGGTTGAAGAAATTAAAGTGGATCCGCCGAAAGCGACCGA
AGTGCGCATCAAAATGCTGTTTGCGAGCATCTGCCACACCGATATGCTGTGCTTTG
ATGGCTTTCCGACCCCGCTGTTTCCGCGCATTCCGGGCCATGAAGGTGTTGGCATG
GTGGAAAGCGTGGGCGAAGATATTAAAACCAAGCTGAAACCAGGCGATATCGTGA
TGCCGCTGTTTATGGGCGAATGCGGCCAGTGCCTGAACTGCAAAAGCAAACGCAC
CAACCTGTGCCATGCGTATCCGCTGACCCTGAGCGGCCTGCTGTTAGATGGCACCA
GCCGCATGAGCATTGCGAAAACCGAAGAAACCATTTATCATCACCTGAGCTGCAG
CACCTGGAGCGAATACATGGTGATTGATATCAACTACGTGCTGAAAATTGACCCGA
AAATGCACCTGCCGTATGCGAGCTTTCTGAGCTGCGGCTTTACCACCGGCTTTGGC
GCGCCGTGGAAAGAGACCCAGATTACCAAAGGCAGCATTGTGGCGGTGTTTGGCC
TGGGCGCGGTGGGCTTAGGTGCAATTAAAGGTGCGCAGATGCAGGGCGCGAGCAT
TATTATTGGCGTGGATATTAACGAAAACAAAGCGGCGAAAGGCAAAGCGTTTGGC
ATGACCCACTTCATTAACCCGAAAGATCATCCGAACCAGCTGGTGAGCGATATGGT
GCGCGATATTACCGATGGCCTGGGCGTGGATTATTGCTTTGAATGCACCGGCATTGC
GAGCCTGCTGAAAGAGATCATCGAAGCGAGCAAAATTGGCTTTGGCACCACCATT
CTGATTGGCGCGGCGCCGGATAACGTGCCGATTAGCAGCCTGAGCCTGATTAACGG
CCGCACCCTGAAAGGCACCACCTTTGGCGGTGTTCGTACTCGCTCTGATCTGCCGA
TCATTCTGCAGAAATGCATGAACGAAGAAATTGAGCTGGATGAGCTGATGAGCCAT
GAAATCCGCCTGGAAAACATCCATGAAATTTTCGAGATTCTGAAAAAGCCGGACT
GCGTGAAGATCCTGATTAACTTTGAT

> Moderately codon-optimized sequences encoding *Tc*ADH2 (*Tc*ADH2low)

ATGTCCTTAAACACTCCGGATGTTATCATTTGTAAAGCTGCAGTAGTACGTGAGCTT
GGCCGCTCGGTGATGGTAGAAGAGATTAAGGTAGATCCACCAAAAGCAACCGAAG
TGCGTATCAAGATGTTGTTCGCTAGCATTTGTCATACCGACATGTTATGCTTCGATG
GCTTTCCTACTCCTTTGTTTCCTCGTATTCCTGGTCATGAAGGTGTCGGGATGGTTG
AGAGTGTCGGGGAAGACATTAAAACAAAGTTGAAACCAGGAGACATCGTGATGC
CCCTCTTTATGGGTGAATGTGGTCAATGCTTAAATTGTAAATCTAAACGTACCAATT
TGTGTCATGCTTACCCTCTGACACTCAGTGGCCTTTTGCTTGACGGCACTTCTCGC
ATGTCCATTGCTAAAACCGAAGAAACTATTTATCACCATTTAAGTTGTTCCACATGG
TCCGAGTACATGGTTATTGACATCAACTATGTACTGAAAATTGACCCCAAAATGCAT
CTCCCTTATGCTAGTTTCCTCTCATGTGGTTTCACGACTGGCTTTGGTGCTCCATGG
AAAGAAACTCAAATCACTAAAGGTTCGATTGTAGCTGTTTTTGGCCTGGGTGCCGT
TGGACTGGGGGCAATCAAAGGGGCACAAATGCAAGGGGCTTCCATCATCATTGGA
GTAGACATCAATGAAAACAAGGCGGCTAAAGGTAAGGCGTTTGGGATGACACATT

TCATCAACCCTAAGGACCATCCTAATCAATTAGTCTCGGATATGGTGCGTGATATCA
CTGATGGTTTAGGCGTTGACTACTGCTTCGAGTGCACAGGAATCGCCTCATTGTTG
AAAGAAATCATTGAAGCCTCTAAAATTGGTTTTGGTACCACAATCTTAATTGGAGC
TGCTCCAGATAACGTTCCAATTAGTAGCCTATCGCTGATTAATGGTCGCACTTTAAA
GGGTACAACTTTTGGTGGTGTACGTACACGCTCTGATCTGCCAATCATTCTTCAAA
AATGTATGAATGAGGAGATTGAACTGGATGAACTTATGTCACATGAAATTCGCTTG
GAGAATATACATGAAATCTTTGAGATCTTGAAGAAGCCTGACTGCGTCAAGATTCT
TATCAACTTTGATTGA

> Sequences encoding *Tc*ALDH

ATGAGTTCAGGAGCTAATGGAAATTCTAAATCTTTAGCTTATGACATCAAGTTCACA
AAGCTCTTCATCAATGGCGAATTTGTAGACTCAATATCAGGAAGCACATTTGAGAC
GATAGATCCAGCAACAGAGGAAGTATTAGCAACTGTGGCTGAAGGGAGAGAGGA
AGACGTTGATCTGGCCGTCAAGGCTGCACGTGAAGCTTTCGACAATGGACCCTGG
CCTCGTTTATCTGGCGAGGCACGTCGGAAGATTCTCTTAAAATTTGCTGACTTAATC
GAGGAAAATGCTGATGAAATAGCAACCCTAGAAGTAATCGACACAGGAAAGCCGT
TTCAAATAGCAAGGTACGTTGAAAACTCATGGACGTCAGAAACATTCCGCTACTTT
GCTGGTGCAGCCGATAAAATCCGTGGAGCCACTCTAAAAATGTCTAGTGATTTTCA
AGCTTACACGTTACGCGAACCAATCGGTGTAGTTGGACACATCATTCCTTGGAATG
CCCCTGCCTATTTATTTGCCATGAAAGTCGCGCCAGCTTTAGCAGCTGGGTGCACC
GTGGTAATCAAGCCAGCTGAAAACACGCCTCTTGTGGGTCTCTTTATGGCTTACTT
GTCTAAACTGGCCGGTGTTCCTGATGGGGTAATTAACGTGGTCAATGGGTTCGGAA
GCACGGCTGGTGCTGCCGTTAGCTCTCATATGGATATTGACGCGGTTACTTTTACGG
GTTCAACCAAAGTTGGACGTACCATAATGCAAGCCGCGGCAGCAAGCAATCTGAA
ACCTGTATCACTCGAATTGGGAGGAAAATCACCTTTTATTGTATTTGATGATGCAGA
TATTGAAAAAGCCGCAGAGATTGCTGTTTTGGGAGTTCTATCGAATAAAGGTGAAT
TGTGCGTGGCAGGATCTCGTGTGTTTGTACACGAAGGAATTTATGATGCTTTTGTC
AAGAAACTGGAAGCAACGGTGAAAAATTGGGCAACAGGAGACCGTTTTGATGCC
GCCACTCGTCATGGTCCTCAGAATAACAAACAACAATACGAAAAAGTTCTTTCGTA
CATTGAGCTTGGCAAGAAGGAAGGTGCAACTTTGGTAACTGGTGGCAAGCCGTTT
GGAAATAAAGGATACTACATCGAGCCGACTTTATTCACGAACGTTACGGATGAAAT
GACCATAGCAAAGGAAGAAATATTTGGCCCTGTTATAATGGTCCTCAAATTCAAGA
CCATTGAAGAAGTGATAAGAAGAGCAAATGCCACAACATATGGACTTGCGGCAGG
TATAATGACGAAAAATATCGACATTGCAAACACAGTTACAAGATCTATTCGAGCTG
GCTCGGTATGGGTCAACTGTTATCTTGCACTTGACCGGGATACACCTTTTGGAGGA
TACAAAATGAGTGGGTTCGGAAGAGAACAAGGATTGGAGGCACTTGAACATTATC
TTCAGGTTAAGACCGTGACTACCCCCATATACAATTCTCCATGGCTCTAG

Supporting Tables

Table S1 Kinetic parameters of *Tc*ADH2 catalytic oxidation of **1** and **3**

Substrate	k_{cat} (s^{-1})	K_m (μM)	K_i (μM)	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
1	6.21×10^{-1}	7.56×10^{2}	5.86×10^{3}	8.21×10^{-4}
3	2.73×10^{-1}	1.07×10^{2}	1.55×10^{3}	2.56×10^{-3}

Table S2 Kinetic parameters of *Ec*ThiM and *Sma*ThiM catalytic phosphorylation of **1**

Enzyme	k_{cat} (s^{-1})	K_m (mM)	k_{cat}/K_m ($\mu M^{-1}\cdot s^{-1}$)
*Ec*ThiM	5.70×10^{-2}	10.61	5.37×10^{-6}
*Sma*ThiM	1.21×10^{-1}	13.07	9.26×10^{-6}

Supporting Figures

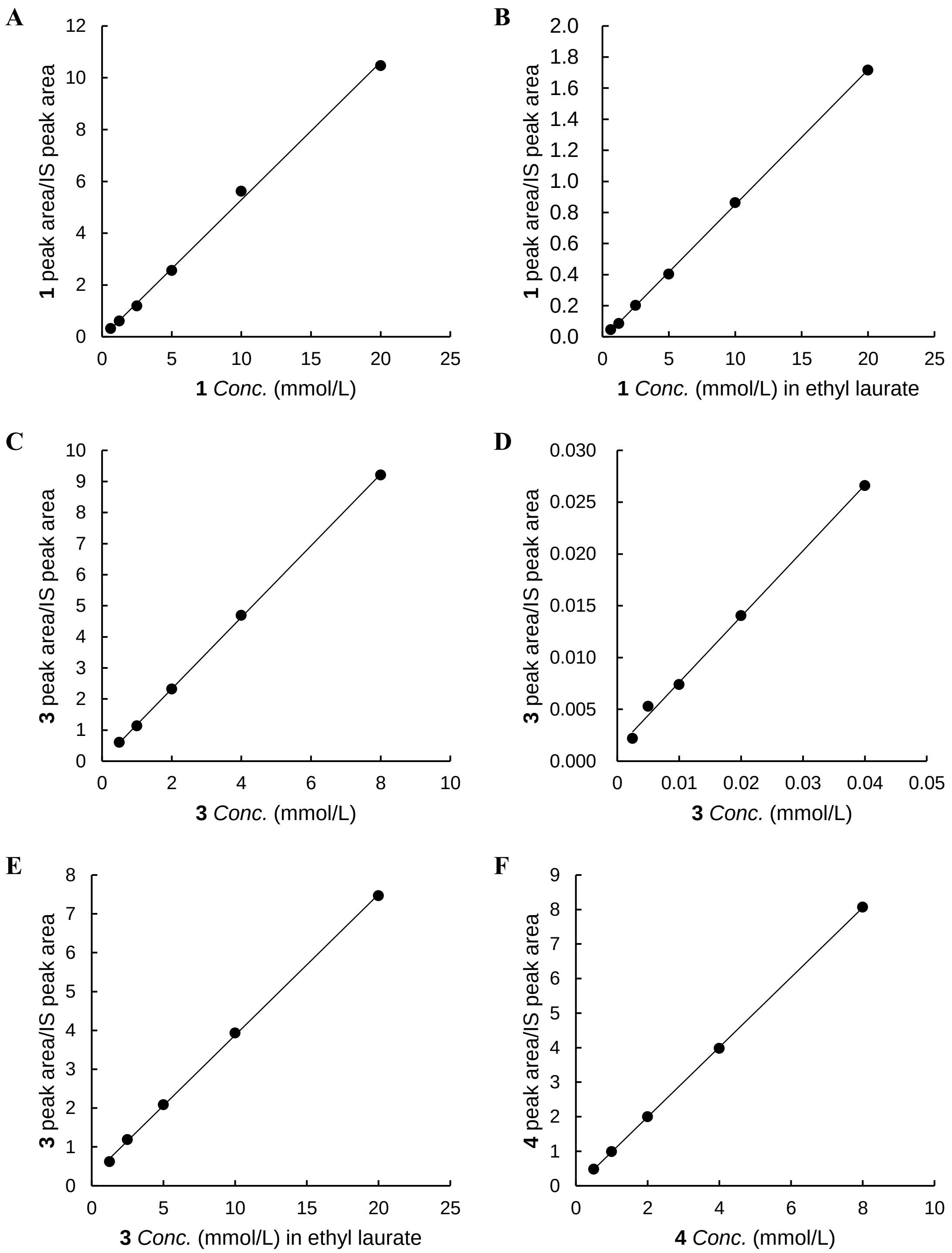
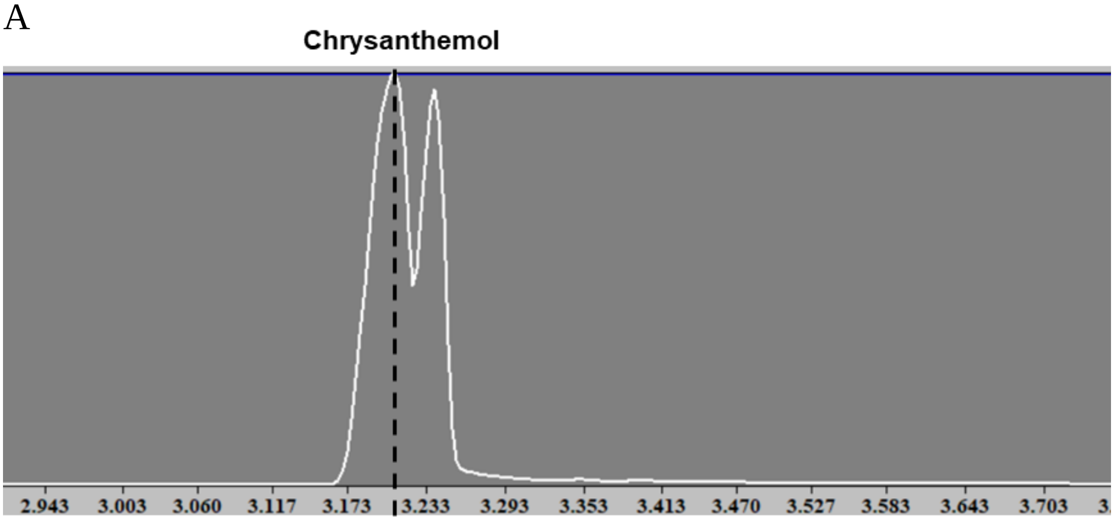


Fig. S1 A_p/A_{IS}-p *Conc.* standard curves of **1**, **3**, **4**. (A) standard curves of **1**; (B) standard curves of **1** in ethyl laurate; (C) standard curves of **3**; (D) standard curves of **3** in the low-concentration range; (E) standard curves of **3** in ethyl laurate; (F) standard curves of **4**.

A



B

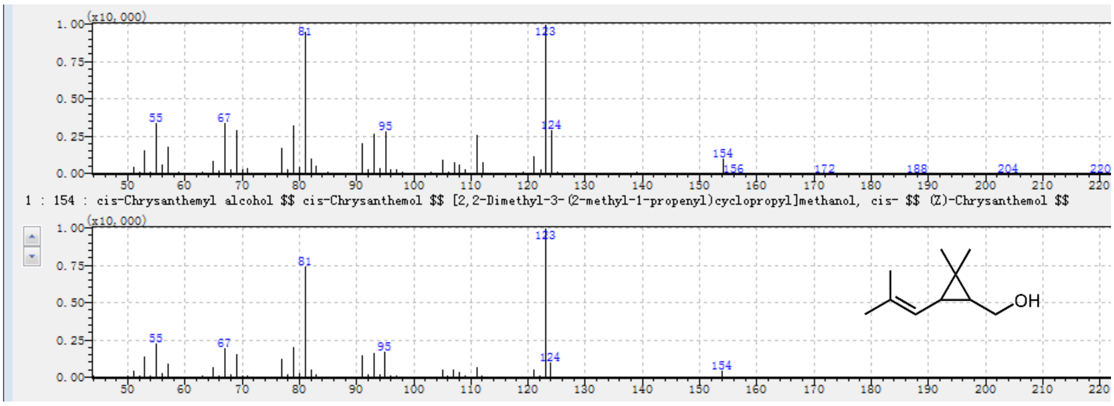
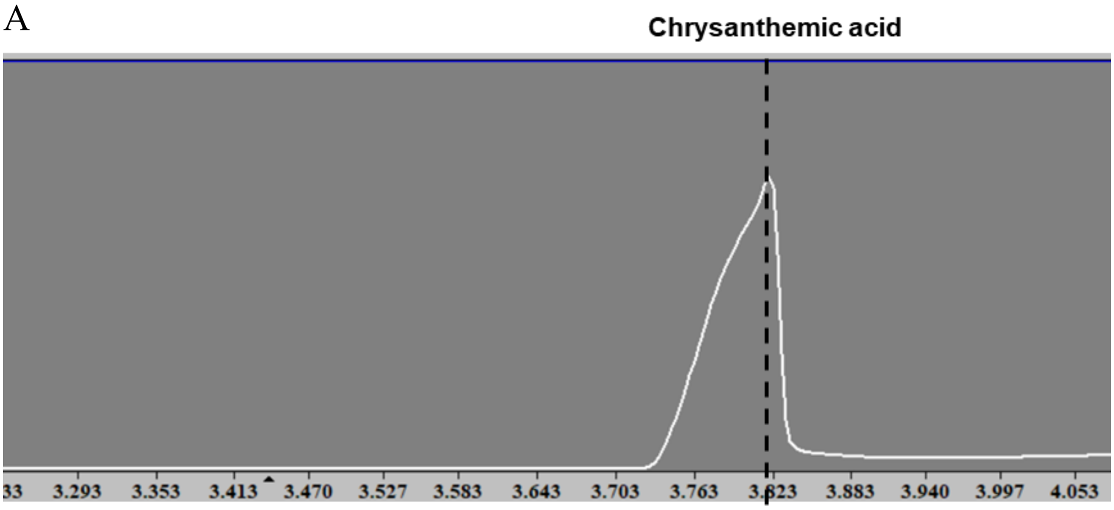


Fig. S2 GC-MS analysis of chrysanthemol **3** standard. (A) GC analysis of chrysanthemol standard; (B) Mass spectra are shown for the experimental sample (upper panel) and the standard chrysanthemol from the NIST database (lower panel).

A



B

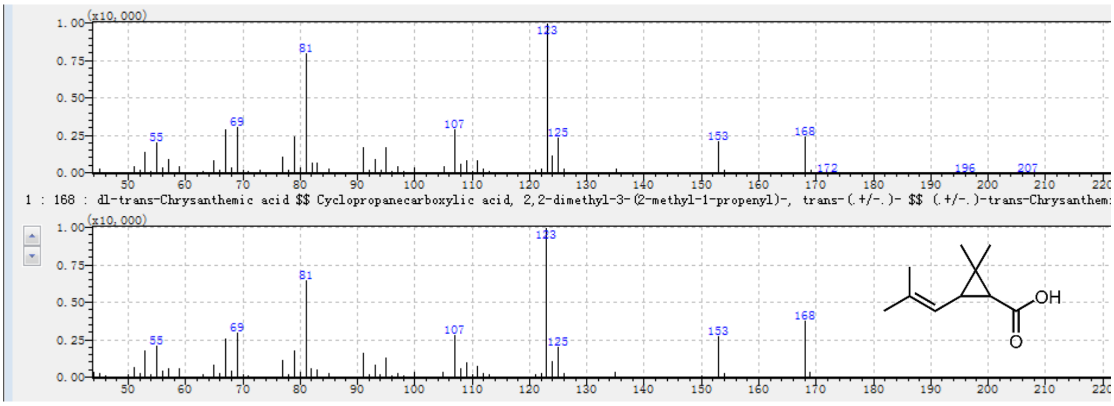


Fig. S3 GC-MS analysis of chrysanthemic acid **4** standard. (A) GC analysis of chrysanthemic acid standard; (B) Mass spectra are shown for the experimental sample (upper panel) and the standard chrysanthemic acid from the NIST database (lower panel).

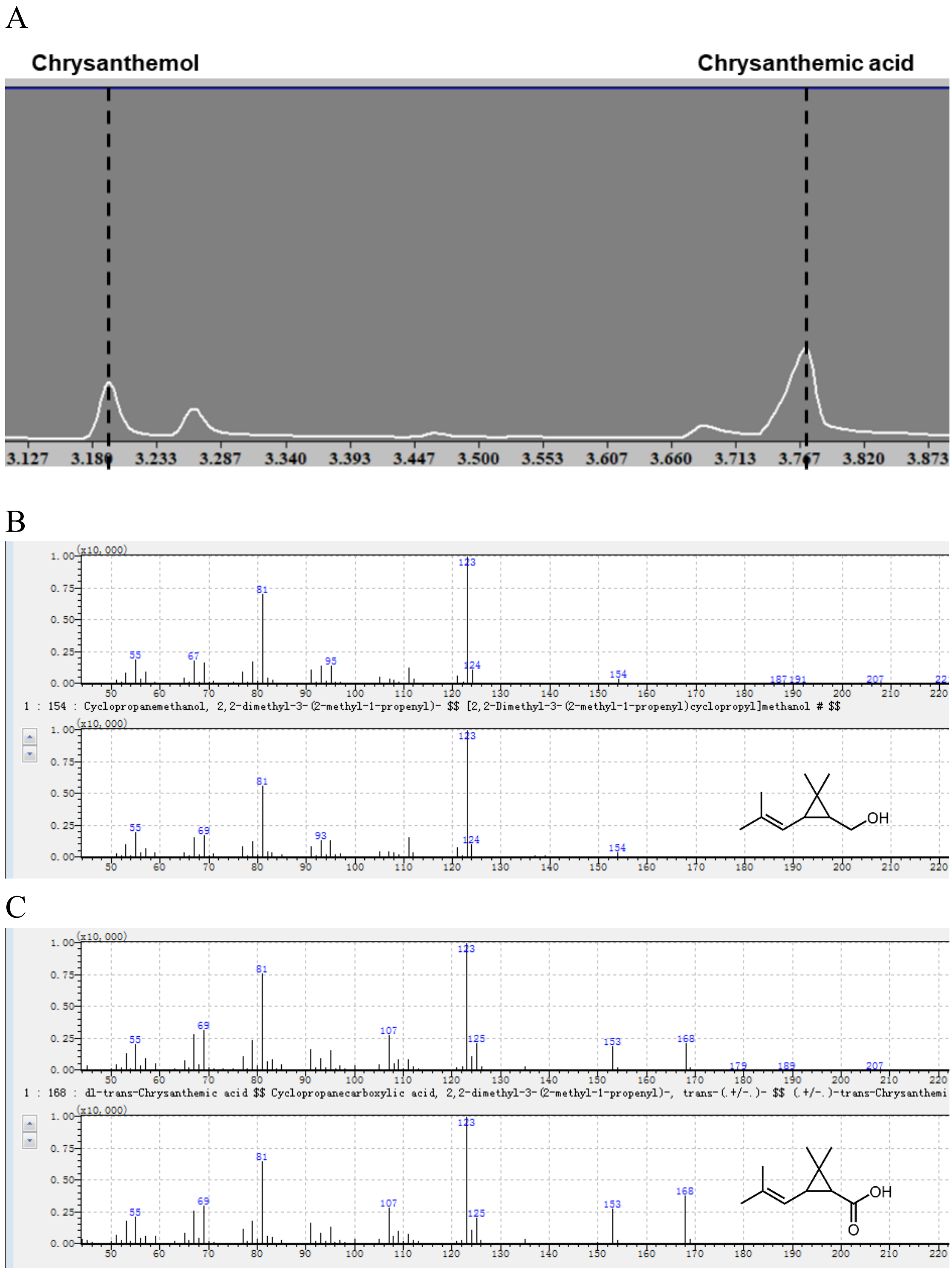
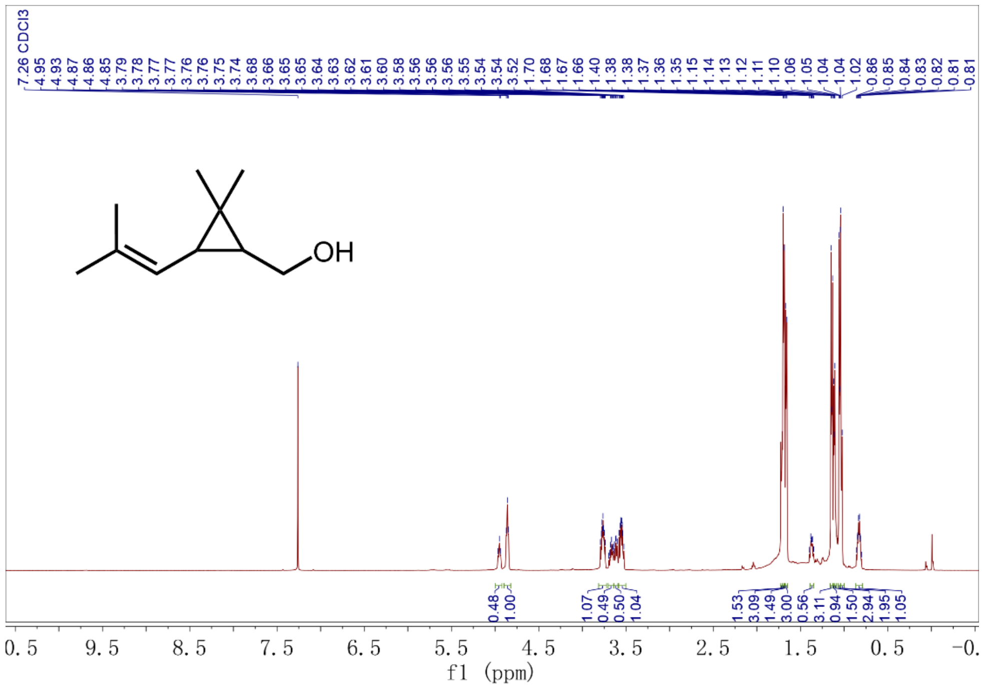


Fig. S4 GC-MS analysis of product **3** and **4**. (A) GC analysis of product sample; (B) Mass spectra are shown for the chrysanthemol peak (upper panel) and the standard chrysanthemol from the NIST database (lower panel); (C) Mass spectra are shown for the chrysanthemic acid peak (upper panel) and the standard chrysanthemic acid from the NIST database (lower panel).

A



B

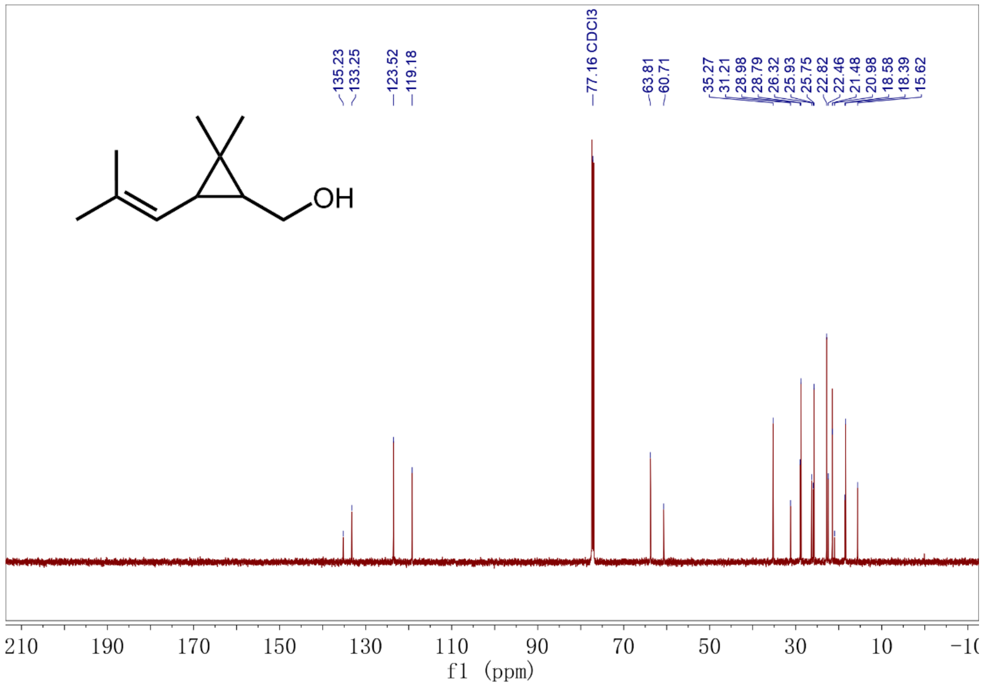


Fig. S5 Nuclear magnetic resonance analysis of chrysanthemol standard. **A:** ^{1}H NMR; **B:** ^{13}C NMR. Characterization Data: ^{1}H NMR (600 MHz, Chloroform-*d*) (dr: 2:1) major isomer δ 4.86 (t, *J* = 6.0 Hz, 1H), 3.77 (ddd, *J* = 18.0, 12.0, 6.0 Hz, 1H), 3.55 (dd, *J* = 12.0, 6.0 Hz, 1H), 1.68 (s, 3H), 1.66 (s, 3H), 1.13 (s, 3H), 1.12 (d, *J* = 6.0 Hz, 1H), 1.06 (s, 3H), 0.89 – 0.75 (m, 1H). ^{13}C NMR (151 MHz, Chloroform-*d*) δ 133.2, 123.5, 63.8, 35.2, 28.8, 25.7, 22.8, 21.4, 18.4, 15.6.

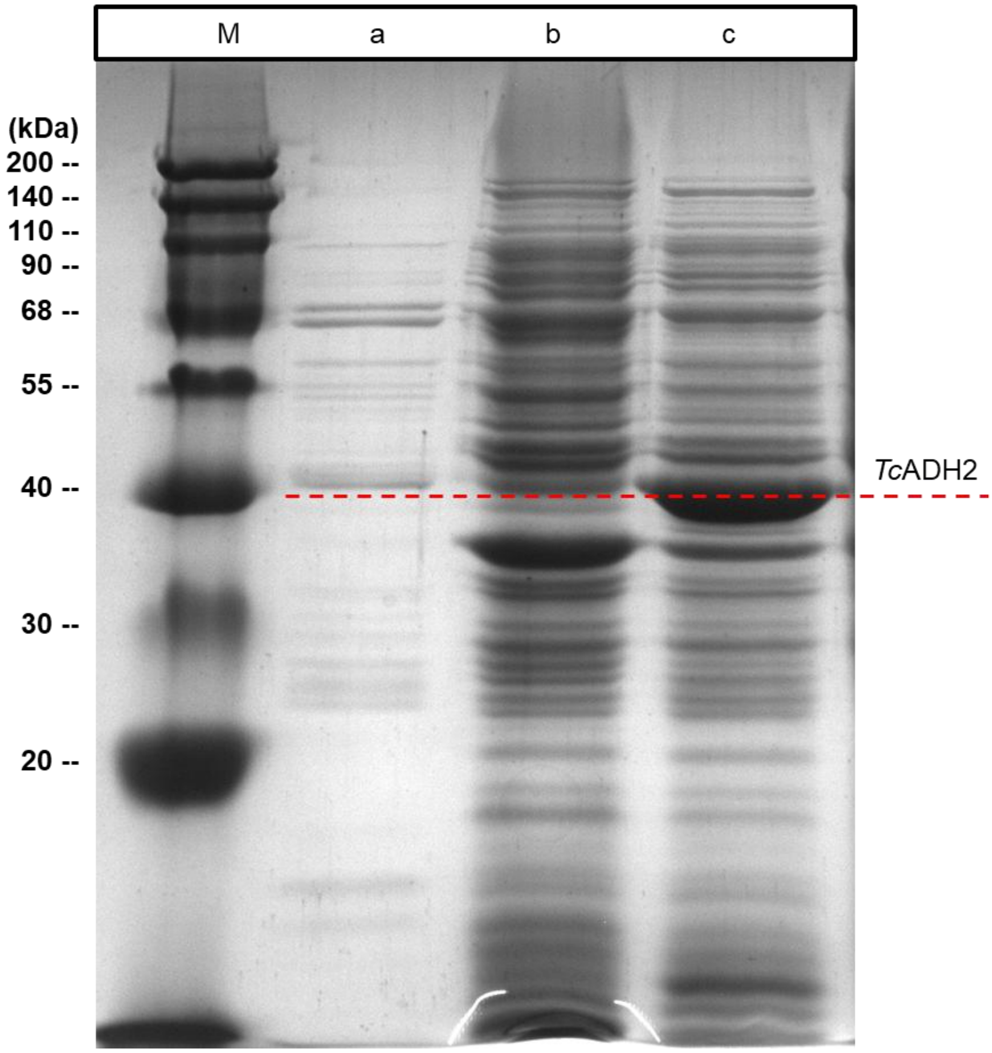


Fig. S6 SDS-PAGE analysis of *Tc*ADH2 (41.4 kDa) purification process. M: Marker; a: Purified *Tc*ADH2 solution; b: cell-free extract (supernatant) of *E. coli* expressing *Tc*ADH2; c: whole-cell protein of *E. coli* expressing *Tc*ADH2.

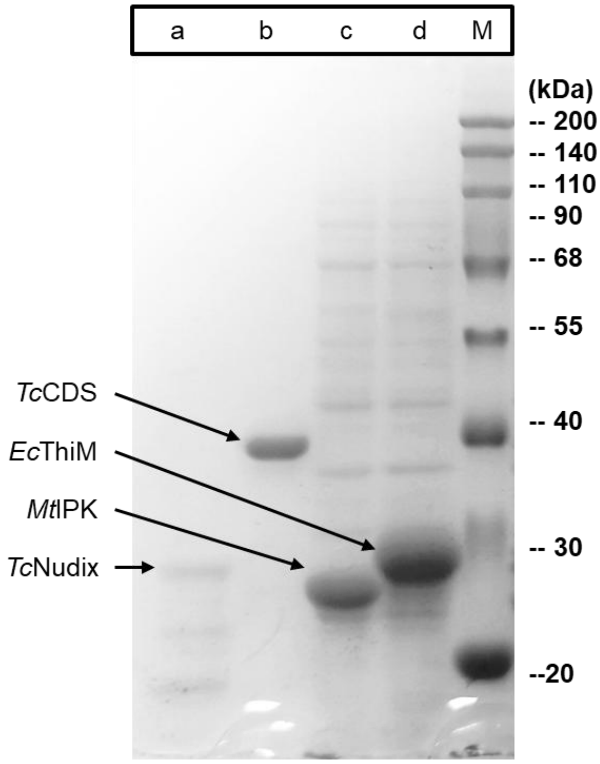


Fig. S7 SDS-PAGE analysis of the purified enzymes involved in the cascade for synthesis of chrysanthemol. a: *Tc*Nudix; b: *Tc*CDS; c: *Mt*IPK; d: *Ec*ThiM; M: Marker.

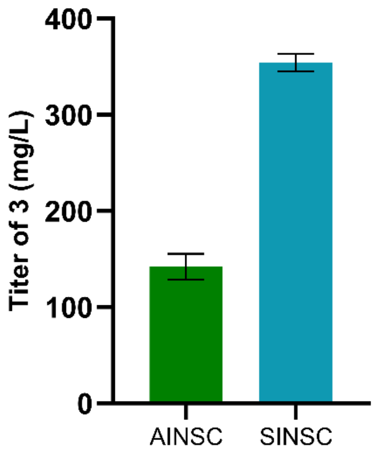


Fig. S8 Titer of **3** using different IKs. The AINSC strain, expressing *At*FKI as the IK, was transformed with pETDuet-*At*FKI-IPK and pACYCDuet-*Tc*Nudix-Sumo*Tc*CDS. In contrast, the SINSC strain, expressing *Sma*ThiM as the IK, was transformed with the same set of expression vector.

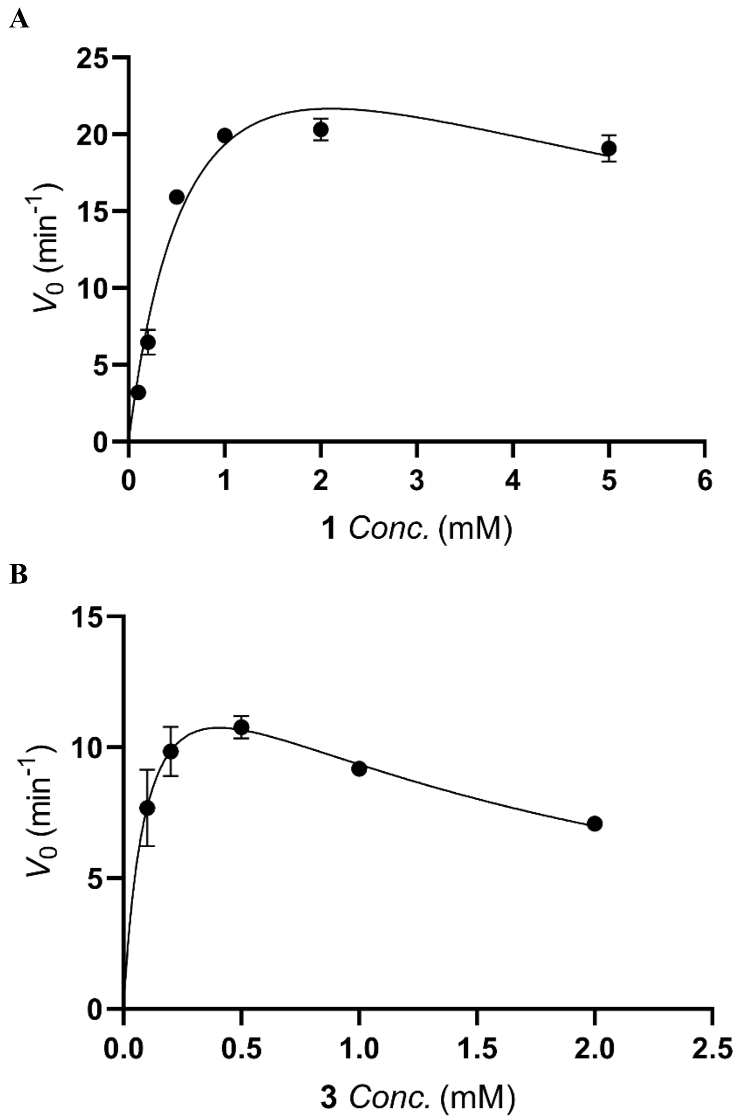


Fig. S9 Haldane curves of *Tc*ADH2-catalyzed oxidation of **1** and **3**. (A) Haldane curves of *Tc*ADH2-catalyzed oxidation of **1**; (B) Haldane curves of *Tc*ADH2-catalyzed oxidation of **3**.

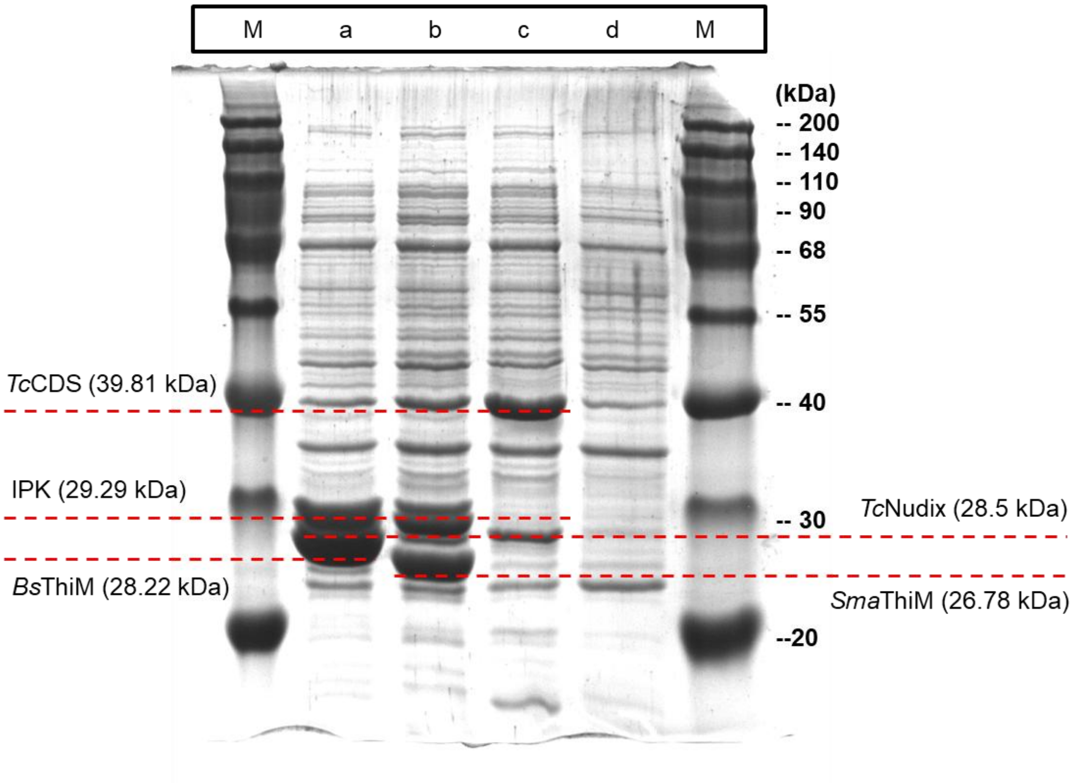


Fig. S10 SDS-PAGE analysis of **3**-producing strains, which co-expresses pETDuet-*Sma*ThiM/*Bs*ThiM-IPK and pACYCDuet-*Tc*Nudix-*Tc*CDS. a: **3**-producing strains (expressing *Bs*Thim); b: **3**-producing strains (expressing *Sma*Thim); c: the control group that was transferred by pACYCDuet-*Tc*Nudix-*Tc*CDS only; d: the blank control group that was transferred by pETDuet-1 and pACYCDuet-1 without any target gene; M: Marker. According to the electrophoresis results, the 4 pathway enzymes IK (including *Sma*Thim (26.78 kDa) and *Bs*Thim (28.22 kDa)), IPK (29.29 kDa), *Tc*Nudix (25.46 kDa) and *Tc*CDS (39.81 kDa) in the **3**-producing strains were successfully expressed.

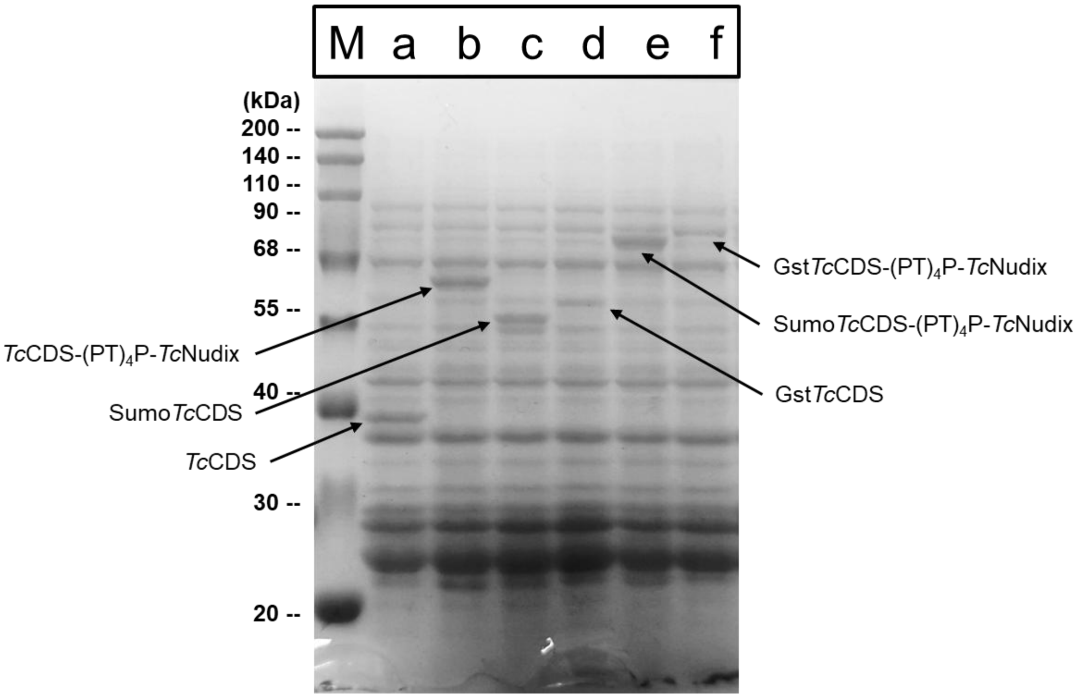
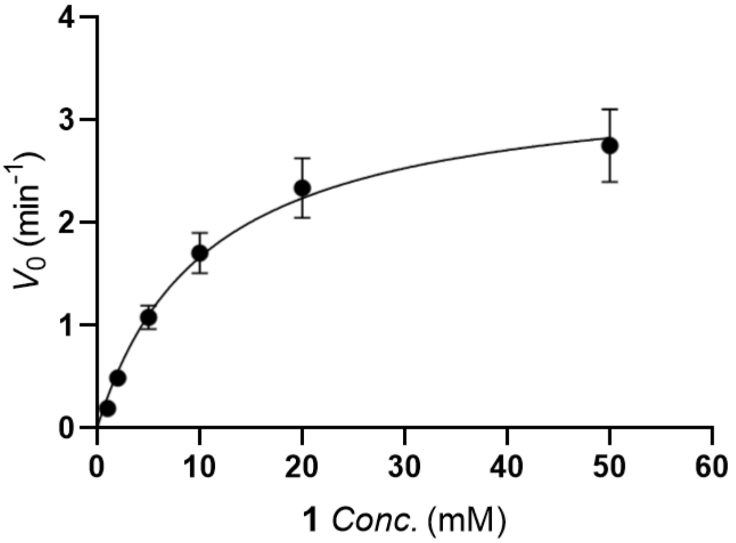


Fig. S11 SDS-PAGE analysis of **3**-producing strains that fusing different solubility tags to the *Tc*CDS N-terminus (in the AC-N-C vector) or to the fusion protein N-terminus (in the AC-C-(PT)$_4$P-N vector): M: Marker; a: pACYCDuet-*Tc*Nudix-*Tc*CDS; b: pACYCDuet-*Tc*CDS-(PT)$_4$P-*Tc*Nudix; c: pACYCDuet-*Tc*Nudix-Sumo*Tc*CDS; d: pACYCDuet-*Tc*Nudix-Gst*Tc*CDS; e: pACYCDuet-Sumo*Tc*CDS-(PT)$_4$P-*Tc*Nudix; f: pACYCDuet-Gst*Tc*CDS-(PT)$_4$P-*Tc*Nudix.

A



B

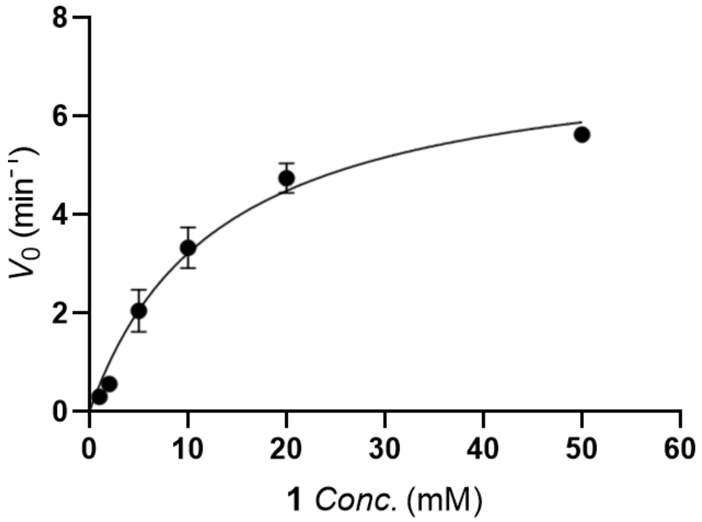
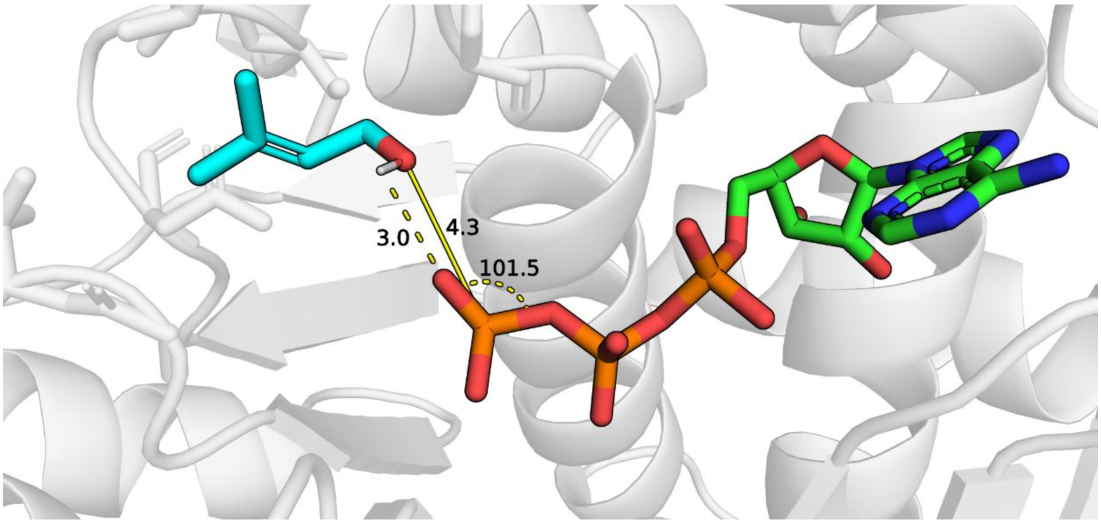


Fig. S12 Michaelis–Menten curves of phosphorylation of **1** by *Ec*ThiM (A) and *Sma*ThiM (B).

A



B

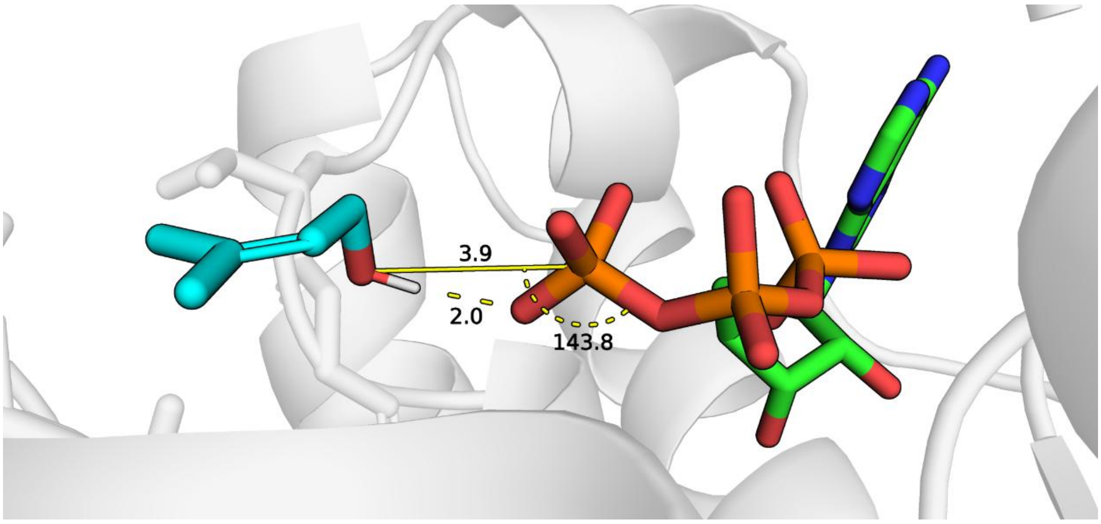
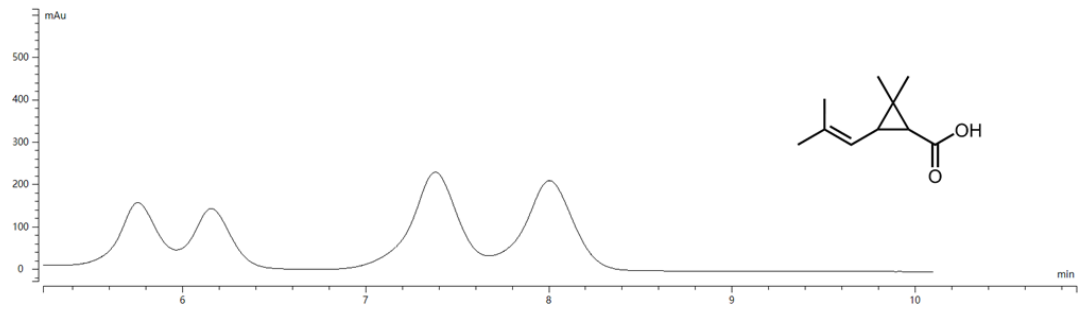
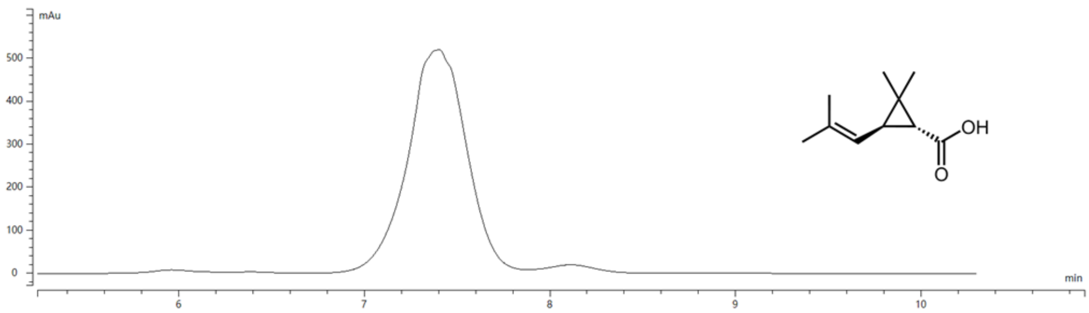


Fig. S13 Molecular docking results of *Ec*ThiM (A) and *Sma*ThiM (B). The angle formed by the prenol hydroxyl oxygen, the ATP γ-phosphorus atom, and the leaving group oxygen is 143.8° in *Sma*ThiM, compared to 101.5° in *Ec*ThiM. The former more closely approaches the optimal geometry for an in-line attack, thereby conferring a catalytic advantage.[5, 6] Furthermore, hydroxyl hydrogen atom of **1** is positioned closer to the γ-phosphate oxygen atom (which may serve as a proton acceptor) in *Sma*ThiM (2.0 Å) than in *Ec*ThiM (3.0 Å). Additionally, the hydroxyl oxygen atom is nearer to the γ-phosphorus atom of ATP in *Sma*ThiM (3.9 Å) than in *Ec*ThiM (4.3 Å). Under the substrate-assisted catalysis mechanism,[7] these proximity effects allow prenol in *Sma*ThiM to be more readily deprotonated to form an oxyanion, which can then more easily attack the γ-phosphate group of ATP to drive the reaction forward.

A



B



C

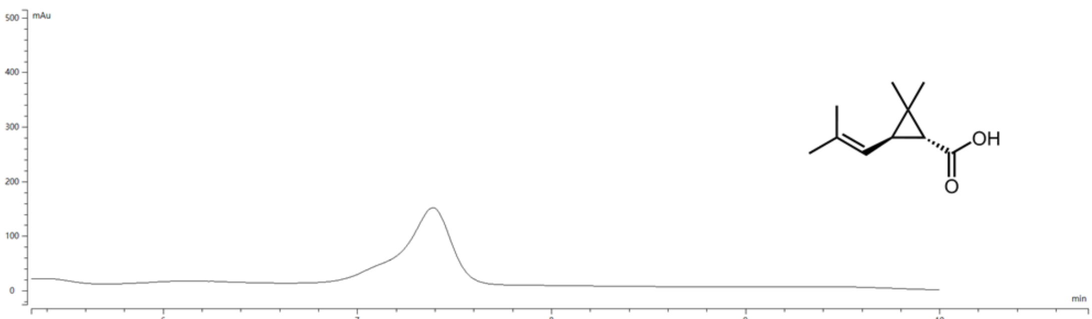


Fig. S14 Chiral HPLC analysis of (±)-chrysanthemic acid standard, (1*R*, 3*R*)-chrysanthemic acid standard, and product **4**. (A) (±)-chrysanthemic acid standard; (B) (1*R*, 3*R*)-chrysanthemic acid standard; (C) Product **4** produced by the SINSCSAA strain.

Reference

1. S. Lund, R. Hall and G. J. Williams, ACS Synth. Biol., 2019, **8**, 232-238.
2. L. A. Johnson, A. Dunbabin, J. C. Benton, R. J. Mart and R. K. Allemann, Angew. Chem. Int. Ed., 2020, **59**, 8486-8490.
3. W. Du, Z. Cheng, X. Pan, C. Liu, M. Yue, T. Li, Z. Xiao, L.-L. Li, X. Zeng, X. Lin, F. Li and L.-B. Dong, Angew. Chem. Int. Ed., 2025, **64**, e202419463.
4. A. O. Chatzivasileiou, V. Ward, S. M. Edgar and G. Stephanopoulos, Proc. Natl. Acad. Sci. USA, 2019, **116**, 506-511.
5. Z. Wang and P. A. Cole, in *Methods Enzymol.*, ed. K. M. Shokat, Academic Press, 2014, vol. 548, pp. 1-21.
6. G. K. Smith, Z. Ke, H. Guo and A. C. Hengge, J. Phys. Chem. B, 2011, **115**, 13713-13722.
7. I. Bordes, R. Castillo and V. Moliner, J. Phys. Chem. B, 2017, **121**, 8878-8892.