Electronic Supplementary Information

Understanding the molecular properties of the E1 subunit (SucA) of α-ketoglutarate dehydrogenase complex from *Vibrio vulnificus* for the enantioselective ligation of acetaldehydes into *(R)*-acetoin

Pil-Won Seo[‡a], Hye-Jin Jo[‡b], In Yeub Hwang[‡c], Ha-Yeon Jeong[b], Jun-Hong Kim[a], Ji-Won Kim[a], Eun Yeol Lee*[c], Jin-Byung Park*[b], and Jeong-Sun Kim*[a]

[a] Department of Chemistry, Chonnam National University, Gwangju 61186, Republic of Korea

[b] Department of Food Science and Engineering, Ewha Womans University, Seoul 03760, Republic of Korea

[c] Department of Chemical Engineering, Kyung Hee University, Gyeonggi-do 17104, Republic of Korea

Table of contents

Chemicals and reagents

Acetaldehyde, glycolaldehyde, formaldehyde, calcium acetate, 2-(Bis(2-hydroxyethyl)amino)acetic acid, Polyethyleneglycol 300, thiamine diphosphate (ThDP), magnesium chloride, glycolaldehyde-1-^{13}C, and ampicillin were purchased from Sigma (St. Louis, MO, USA). *(R)*- and *(S)*-Acetoin was obtained from Tokyo Chemical Industry Co. (Tokyo, Japan). Ethyl acetate was purchased from Duksan Pure Chemical Co. (Ansan, Korea).

Microbial strains and culture conditions

The recombinant *Escherichia coli* BL21(DE3) strains, expressing the recombinant enzymes including the bacterial SucAs, were cultivated in the lysogeny broth (LB) medium supplemented with 50 μg/mL ampicillin (Table S1). The recombinant *E. coli* cultures were incubated at 37°C with shaking at 250 rpm (Jeiotech, Daejeon, Korea).

Gene cloning and construction of recombinant plasmids

The gene coding for $VvSucA_{\Delta 84}$ (Asp85-Asp941) and $MbSucA_{\Delta 361}$ (Asp362-Gly941) was amplified by polymerase chain reaction (PCR) from chromosomal DNA of *V. vulnificus* CMCP6 and *M. bovis* using the primers (Table S2), which were designed for Ligation-Independent Cloning (LIC) [1]. Each PCR product was separately treated with the T4 polymerase and ligated with the T4 polymerase treated vector pLIC-B3, a derivative pET-21a (Novagen, Madison, WI). The target gene was designed to translate the protein in a fused state with a 6-His and Tobacco etch virus (TEV) protease cleavage sequence at the N-terminus of the protein. EcSucA and ZmPDC genes were amplified using the chromosomal DNA of *E. coli* K12 and *Zymomonas mobilis* ZM4 using the designed primers (Table S2). FLS

was synthesized (Bioneer, Daejeon, Korea) and then amplified by PCR using the primers (Table S2). The genes of MaSucA and MtSucA were also amplified from the chromosomal DNAs of *M. alcaliphilum* 20Z and *Methylomonas* sp. DH-1 using the respective primers (Table S2). These five genes were respectively inserted into the *E. coli* expression vector pET21b(+) using *Nde*I and *Xho*I restriction enzymes. All the used primers to amplify the designated genes are tabulated (Table S2).

Site-directed mutagenesis of $VvSucA_{\Delta 84}$ was performed by overlap PCR method [2] on pLIC-B3-$VvSucA_{\Delta 84}$ (Table S2).

Expression and purification of the recombinant enzymes

All recombinant plasmids were expressed and purified following a similar protocol. The plasmid was transformed to *E. coli* BL21(DE3) Star and cultured in 1 liter of LB media [3] containing 50 μg/mL ampicillin. When OD_{600} reached 0.4-0.6, isopropyl β-D-1-thiogalactopyranoside (IPTG) was added (0.05 mM final concentration), and cultured for additional 16 h at 18 ℃. Cells were harvested by centrifugation at 8,660 g for 10 min at 4 ^{o}C.

The cell pellet was resuspended in lysis buffer [20 mM Tris-HCl (pH 7.5), 500 mM NaCl, and 1mM β -mercaptoethanol (β-ME)], containing 1 mM phenylmethylsulfonyl fluoride (PMSF), and disrupted by sonication. The supernatant was collected by centrifugation at 31,660 g for 30 min at 4 ^{o}C and loaded onto 5ml His-bind agarose resin column (Elpis Biotech) that was pre-equilibrated with the lysis buffer. The bound protein was eluted by a step gradient of imidazole in the lysis buffer.

For crystallization of $VvSucA_{\Delta 84}$, the recombinant protein was digested with a recombinant TEV protease to remove the 6xHis tag, and was dialyzed by a buffer containing 20 mM Tris-HCl (pH 7.5), 50 mM NaCl, and 1mM β-ME. $VvSucA_{\Delta 84}$ was further isolated by using a HiTrap Q column (GE healthcare) operated with a linear NaCl gradient. Gel filtration chromatography using HiLoad 26/600 Superdex 200pg (GE Healthcare) was applied at the final purification step in a buffer of 20 mM Tris-HCl (pH 7.5) and 200 mM NaCl.

General procedures for the enzyme assay and biotransformations

The carboligating activities of the enzymes including the bacterial SucAs were determined by quantification of the products accumulating in the reaction medium. The reactions were initiated by adding the reaction substrate (i.e., 25 mM acetaldehyde) into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgCl_2$, and 0.1 mg/mL of the enzymes. The reaction mixture were incubated at 30^{o}C and 300 rpm for 10 to 24 h.

Structure determination of $VvSucA_{\Delta 84}$

For crystallization, the purified $VvSucA_{\Delta 84}$ protein was concentrated to 5 mg/ml in 20 mM Tris-HCl (pH7.5), 200 mM NaCl, 0.1 mM ThDP, and 0.1 mM $MgCl_2$. The protein concentration was determined by considering the extinction coefficient of 1.142 $mg^{-1}{\cdot}ml{\cdot}cm^{-1}$, which was calculated from its amino-acid sequence. To obtain the protein crystals, the screening using commercial solution kits was attempted by sitting-drop vapor-diffusion method at 294 K. The initial crystals were grown from a precipitant solution consisting of 200 mM calcium acetate, 100 mM sodium cacodylate (pH 6.5), and 40% (v/v) polyethylene glycol 300 (PEG300). The crystals suitable for diffraction experiments were obtained in 220 mM

calcium acetate, 100 mM 2-(Bis(2-hydroxyethyl)amino)acetic acid (BICINE) (pH 9.0), 50% (v/v) PEG 300, 0.1 mM ThDP, and 0.1 mM $MgCl_2$ using hanging-drop vapor-diffusion method. For diffraction experiments, the crystals were immersed into the precipitant solution containing an additional 12% (v/v) glycerol as a cryoprotectant for 5 s and placed in the 100 K nitrogen-gas stream. Diffraction data of $VvSucA_{\Delta 84}$ were collected on the beamline BL-11C of the Pohang Accelerator Laboratory (PAL) in Korea. Indexing, integration, and scaling of the reflections were done using the HKL2000 suite [4]. The electron-density map of $VvSucA_{\Delta 84}$ was calculated from molecular replacement with the PHENIX program [5] using the structure of EcSucA (PDB ID 2JGD) as a search model. Further model building was done manually using WinCoot [6], and subsequent refinement was performed with PHENIX [5]. To get the glycolaldehyde-1-^{13}C-bound structure, the crystals were soaked for 30 mins in 100 mM HEPES pH 7.5, 220 mM calcium acetate, 50% (v/v) PEG300, 0.1 mM ThDP, 0.1 mM $MgCl_2$, 12% (v/v) glycerol, and 4 mM glycolaldehyde-1-^{13}C. The diffraction data were also obtained at the same beamline and the structure was determined also by molecular replacement using the determined VvSucA structure. The data collection and refinement statistics are summarized in Table S3.

Product analysis using gas chromatography/mass spectrometry

The 125 µL of the reaction mixture was extracted by using 700 µL of ethyl acetate. The organic phase was collected by centrifugation and was dried over $MgSO_4$. The organic phase was then subjected to gas chromatography/mass spectrometry (GC/MS) analysis. The GC/MS (Agilent Technologies, USA) equipped with a chiral column (Supelco β-DEX™ 120, 30 m length, 0.25-mm inner diameter) was used to determine the concentrations and

enantioselectivities of acetoin. Isoamylol (10 mM) was used as an internal standard. Column temperature was increased from 50 to 85°C at a rate of 17°C/min and to 90°C at a rate of 3.5°C/min, and then maintained at 90°C for 3 min. Helium was used as the carrier gas at a flow rate of 5 mL/min.

Product analysis using high performance liquid chromatography

The 200 µL of the reaction mixture was captured and quenched by adding 10 % H_2SO_4, centrifuged (13,000 rpm, 5 min) and analyzed by high performance liquid chromatography (HPLC). The HPLC system was equipped with a Bio-Rad Aminex HPX-87H column. The mobile phase consisted of 10 mM H_2SO_4 at 0.5 mL/min (55°C). A refractive index detector and UV detector were used for quantification of glycolaldehyde and erythrulose.

Supplementary tables

Table S1. Plasmids used in this study

Plasmids	Features	Description	Reference
pLIC-B3	T7, Amp	pET-21a derivative	[1]
pLIC-B3 $VvSucA_{\Delta 84}$	T7, Amp	pET-21a derivative	This study
LIC-B3-$MbSucA_{\Delta 361}$	T7, Amp	pET-21a derivative	This study
pET-21a-EcSucA	T7, Amp	pET-21a	This study, [7]
pET-21b-MaSucA	T7, Amp	pET-21b	This study
pET-21b-MtSucA	T7, Amp	pET-21b	This study
pET-21b-ZmPDC	T7, Amp	pET-21b	This study
pET-21b-FLS	T7, Amp	pET-21b	This study, [8, 9]
pET-21b-MtputSucA	T7, Amp	pET-21b	This study

Table S2. Oligonucleotide primers used for PCR or site-directed mutagenesis

Name	Primer sequence (5’ - 3’)
Oligonucleotides used for PCR	
$VvSucA_{\Delta 84}$	(F)-GGCGGTGGTGGCGGCGATGTCGATGCTAAGCAAGTT
	(R)-GTTCTTCTCCTTTGCGCCCTTAATCCGAAGTCTTTTCGTTC
$MbSucA_{\Delta 361}$	(F)-GGCGGTGGTGGCGGCGACTCGATCGTCGACAAGAACGC
	(R)-GTTCTTCTCCTTTGCGCCCTCAGCCGAACGCCTCGTCGAG
EcSucA	(F)- GCAGGTCGACAAGCTTATGCAGAACAGCGCTTTG
	(R)- ATGCGGCCGCAAGCTTTTATTCGACGTTCAGCGCG
ZmPDC	(F)-GCCAATCATATGAGTTATACTGTCGGTAC
	(R)-CGGTTACTCGAGGAGGAGCTTGTTAACAGGCT
FLS	(F)-GCCGCCCATATGGCTATGATTACTGGT
	(R)-CGGCGGCTCGAGCGCGCCGGATTGGAAATA
MaSucA	(F)-GCCACCCATATGAGCGAACTCCTCAAGTT
	(R)-GGCGGTCTCGAGTTTTCCTGTGTGTTTTCCAA
MtSucA	(F)-GGAGATATACATATGATGAGTAGCCTGCTTAAAGA
	(R)-TGGTGGTGGTGCTCGAGAAGCTCATGTTTGTTTCC
MtputSucA	(F)-GGAGATATACATATGATGGAAATCAGCCAAAAATC
	(R)-TGGTGGTGGTGCTCGAGGAAAAGGTCGAGAATTTG
Oligonucleotides used for site-directed mutagenesis of $VvSucA_{\Delta 84}$	
K228L	(F)-GCAAAATTCCCAGGAGCG**CTG**CGTTTCTCTCTCGAAGGT
	(R)-ACCTTCGAGAGAGAAACG**CAG**CGCTCCTGGGAATTTTGC

K228S	(F)-GCAAAATTCCCAGGAGCG**TCG**CGTTTCTCTCTCGAAGGT
	(R)-ACCTTCGAGAGAGAAACG**CGA**CGCTCCTGGGAATTTTGC
V298-H301-to-GGG	(F)-ACGGGCGAT**GGAGGTGGC**CAAGGTTTCTCAGCGGAC
	(R)-GAAACCTTG**GCCACCTCC**ATCGCCCGTACCCCAACT
Y300H	(F)-GGTACGGGCGATGTAAAA**CAT**CACCAAGGTTTCTCAGCGGAC
	(R)-GTCCGCTGAGAAACCTTGGTG**ATG**TTTTACATCGCCCGTACC
S324N	(F)-GCCCTTGCGTTTAACCCA**AAT**CATTTAGAGATCGTTAAC
	(R)-GTTAACGATCTCTAAATG**ATT**TGGGTTAAACGCAAGGGC
S324T	(F)-GCCCTTGCG TTT AAC CCA **ACT**CATTTAGAGATCGTTAAC
	(R)-GTTAACGATCTCTAAATG**AGT**TGGGTTAAACGCAAGGGC

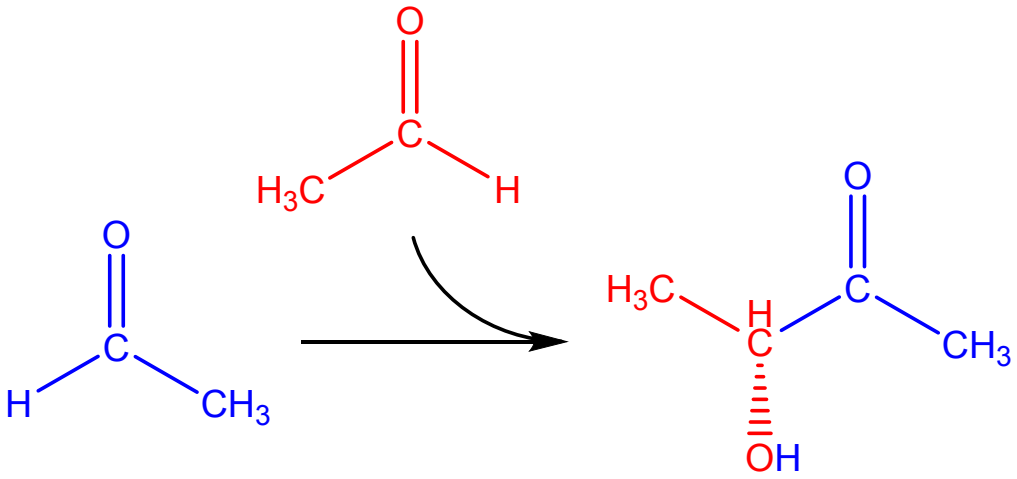
* Bold bases represent the coding sequence of the mutated amino acid.

Table S3. Data Collection and Structure Refinement Statistics.

Data Collection	$VvSucA_{\Delta 84}$	Glycolaldehyde-1-^{13}C-$VvSucA_{\Delta 84}$
Space group	$P2_1$	P1
Unit cell dimensions		
a, b, c (Å)	84.27, 76.71, 144.58	76.82, 84.56, 145.22
α, β, γ (°)	90, 100.56, 90	79.51, 88.47, 89.92
Wavelength (Å)	0.9792	0.9794
Resolution (Å)	50-2.75 (2.80-2.75)[a]	50-2.3 (2.34-2.30)
R_{sym}	27.2 (82.5)	17.7 (71.7)
R_{pim}	12.0 (42.4)	10.8 (53.0)
$I/\sigma(I)$	6.4 (1.2)	6.2 (0.8)
Completeness (%)	98.7 (99.1)	91.4 (89.7)
Redundancy	5.8 (4.3)	3.5 (2.4)
Refinement		
No. of reflections	47911	148359
R_{work} / R_{free}	19.2 (28.1) / 24.9 (31.5)	20.0 (31.3) / 23.0 (34.4)
No. atoms		
protein / water / ligand	13345 / 139 / 93	26456 / 1243 / 218
bond lengths (Å) / angles (º)	0.006 / 0.86	0.002 / 0.58
Average B-values (Å^2)		
protein / water / ligand	43.84 / 36.28 / 43.89	37.17 / 36.86 / 41.05
Ramachandran plot (%)		
favored / allowed / outliers	96.8 / 3.2 / 0.0	96.6 / 3.4 / 0.0

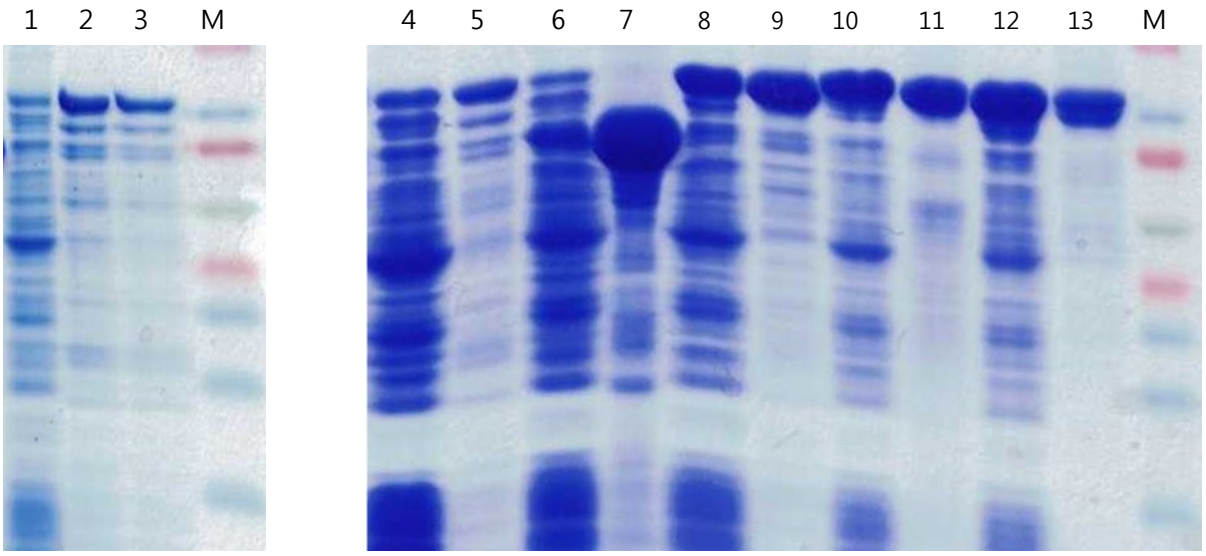
[a] The numbers in parentheses are the statistics from the highest resolution shell.

Supplementary scheme



Scheme S1. Carboligation of two acetaldehyde molecules to *(R)*-acetoin.

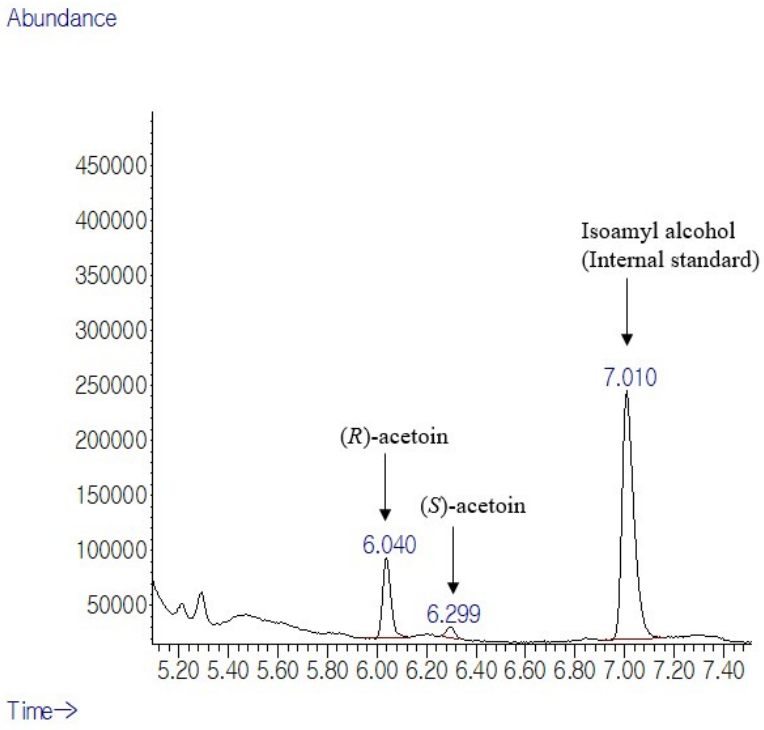
Supplementary figures



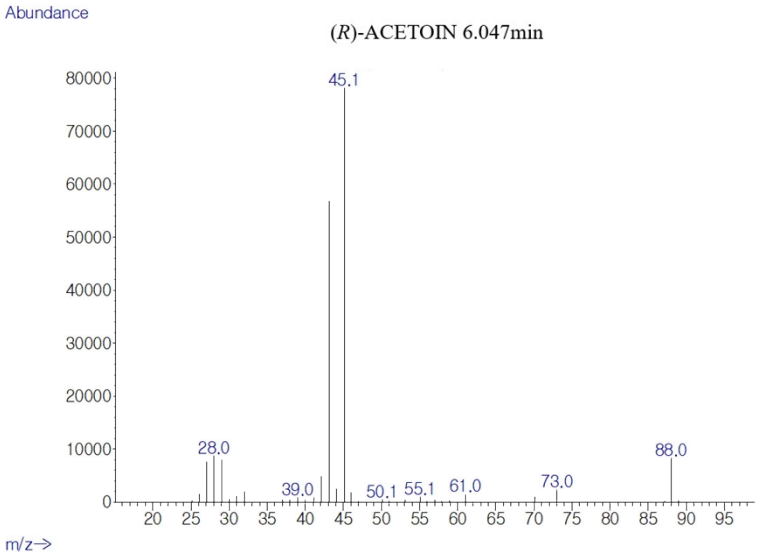
(A) (B)

Fig. S1. SDS-PAGE analysis of the bacterial SucAs expressed in *E. coli* and purified via affinity chromatography. The recombinant *E. coli* were cultivated at 37°C with shaking at 250 rpm. The cultivation temperature was shifted to 20°C just after induction of expression of the target enzymes. The enzymes were purified via affinity chromatography on a Ni-NTA gel matrix. M, Protein ladder; lane 1, a soluble fraction of MaSucA expressed in *E. coli*; lane 2, elution fraction of MaSucA (A); lane 3, elution fraction of MaSucA; lane 4, a soluble fraction of MtSucA in *E. coli*; lane 5, elution fraction of MtSucA; lane 6, a soluble fraction of MtputSucA in *E. coli*; lane 7, elution fraction of MtputSucA; lane 8, a soluble fraction of MbSucA in *E. coli*; lane 9, elution fraction of MbSucA; lane 10, a soluble fraction of VvSucA in *E. coli*; lane 11, elution fraction of VvSucA; lane 12, a soluble fraction of EcSucA in *E. coli*; lane 13, elution fraction of EcSucA (B).

(A)


Abundance
450000
400000
350000
300000
250000
200000
150000
100000
50000
Isoamyl alcohol
(Internal standard)
7.010
(R)-acetoin
(S)-acetoin
6.040
6.299
5.20 5.40 5.60 5.80 6.00 6.20 6.40 6.60 6.80 7.00 7.20 7.40
Time→

(B)


Abundance
(R)-ACETOIN 6.047min
80000
70000
60000
50000
40000
30000
20000
10000
0
45.1
28.0
39.0
50.1
55.1
61.0
73.0
88.0
20 25 30 35 40 45 50 55 60 65 70 75 80 85 90 95
m/z→

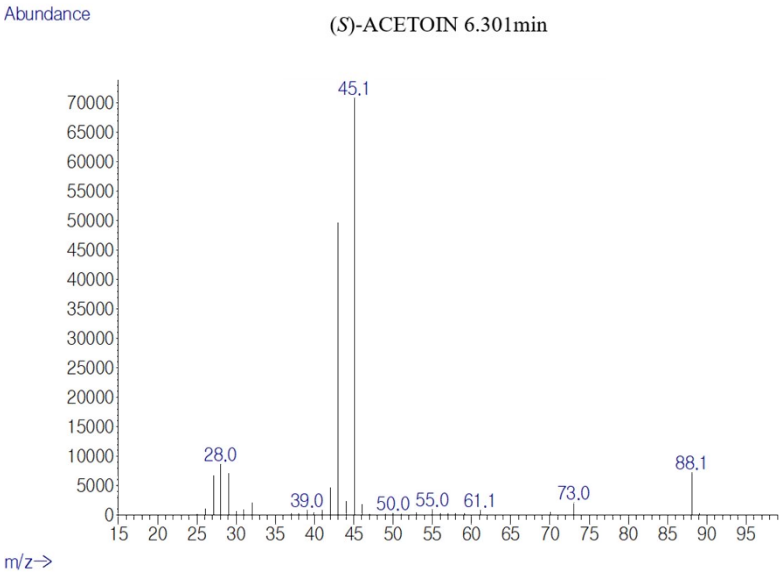


Fig. S2. GC/MS analysis of the reaction products by EcSucA (A) and mass spectrum of *(R)*- and *(S)*-acetoin (B). The reaction was initiated by adding the reaction substrate (i.e., 25 mM acetaldehyde) into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgSO_4$, and 1 mg/mL of EcSucA. The reaction mixture was incubated at 30°C and 300 rpm for 10 h.

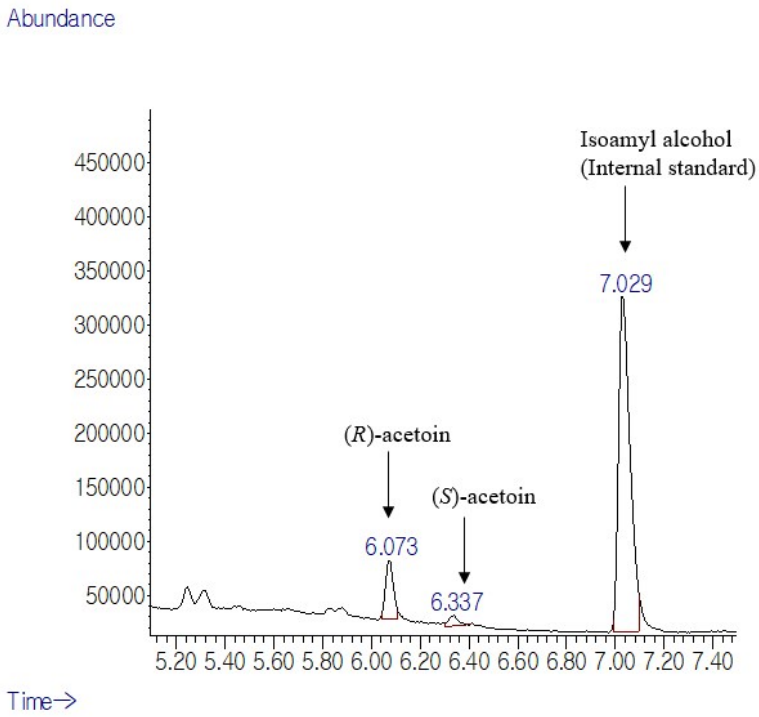


Fig. S3. GC/MS analysis of the reaction products by MbSucA. The reaction was initiated by adding 25 mM acetaldehyde into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgSO_4$, and 1 mg/mL of MbSucA. The reaction mixture was incubated at 30°C and 300 rpm for 10 h.

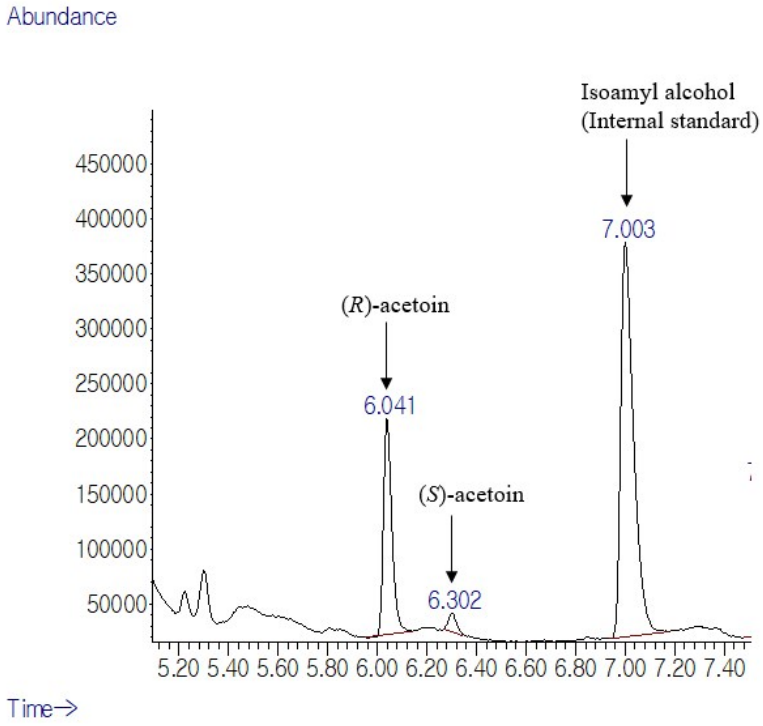


Fig. S4. GC/MS analysis of the reaction products by VvSucA. The reaction was initiated by adding 25 mM acetaldehyde into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgSO_4$, and 1 mg/mL of VvSucA. The reaction mixture was incubated at 30°C and 300 rpm for 10 h.

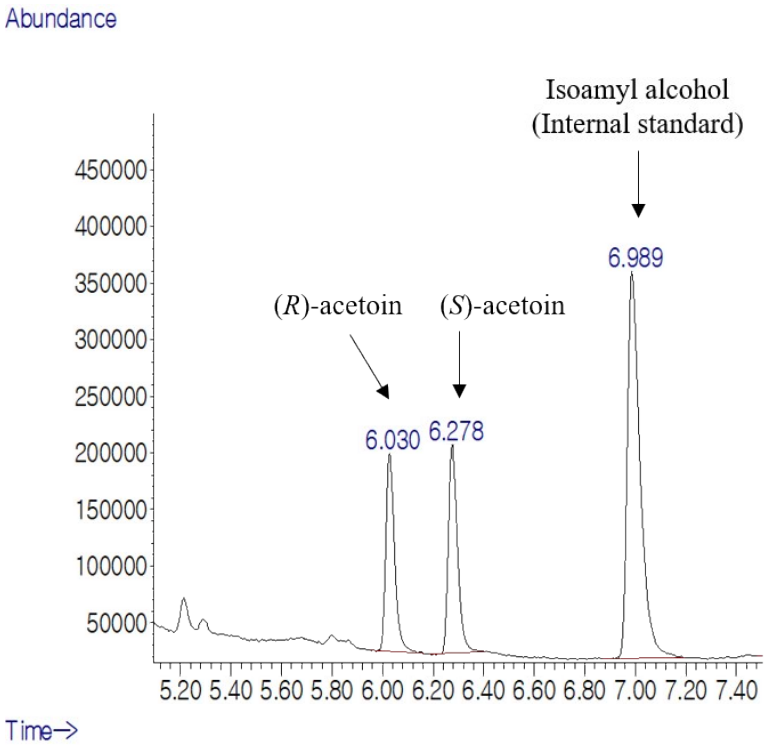


Fig. S5. GC/MS analysis of the reaction products by formolase. The reaction was initiated by adding 25 mM acetaldehyde into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgSO_4$, and 1 mg/mL of formaldehyde-ligating formolase. The reaction mixture was incubated at 30°C and 300 rpm for 10 h.

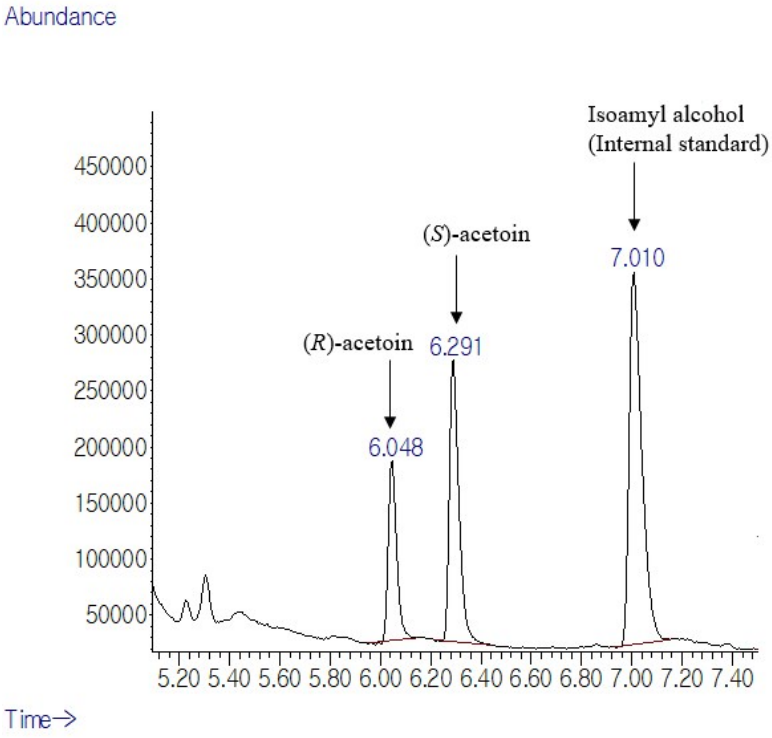
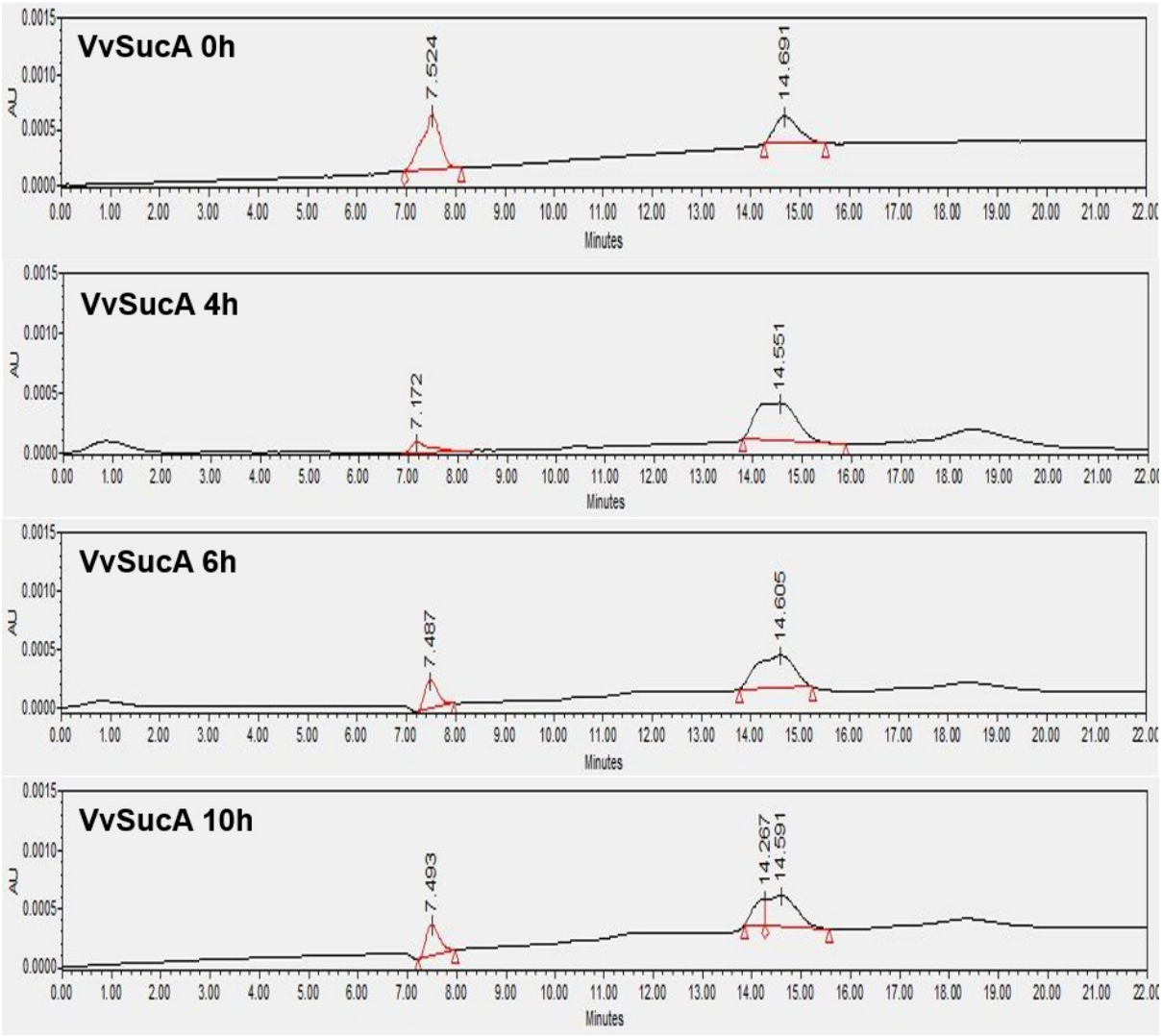
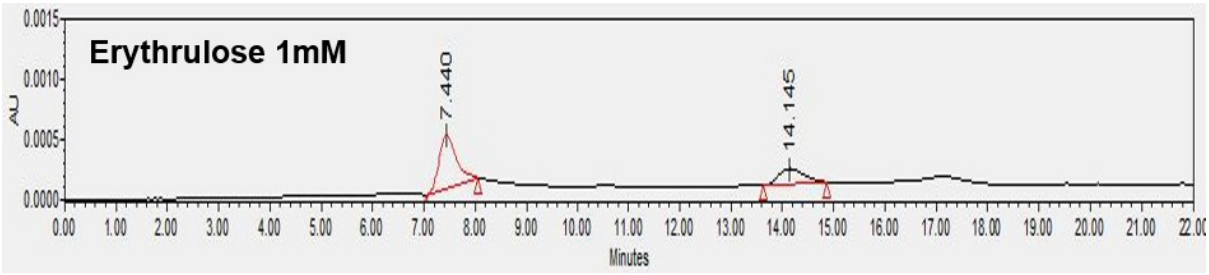


Fig. S6. GC/MS analysis of the reaction products by ZmPDC. The reaction was initiated by adding 25 mM acetaldehyde into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgSO_4$, and 1 mg/mL of ZmPDC. The reaction mixture was incubated at 30°C and 300 rpm for 10 h.

(A)


VvSucA 0h
7.524
14.691
VvSucA 4h
7.172
14.551
VvSucA 6h
7.487
14.605
VvSucA 10h
7.493
14.267
14.591
AU
Minutes

(B)


Erythrulose 1mM
7.440
14.145
AU
Minutes

(C)

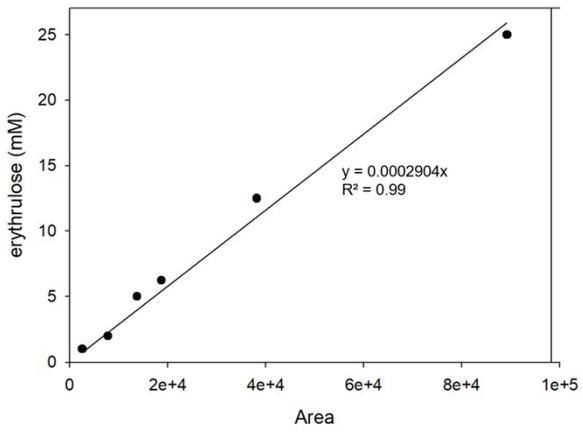


Fig. S7. Biotransformation of glycolaldehyde into erythrulose by VvSucA. HPLC analysis of the reaction products (A) and 1mM erythrulose (B). The standard curve used for determination of erythrulose concentrations (C). The reaction was initiated by adding the reaction substrate (i.e., 25 mM glycolaldehyde) into 50 mM phosphate buffer (pH 8.0) containing 0.1 mM ThDP, 2 mM $MgSO_4$, and 1 mg/mL of VvSucA. The reaction mixture was incubated at 30^{o}C and 300 rpm for 10 h.

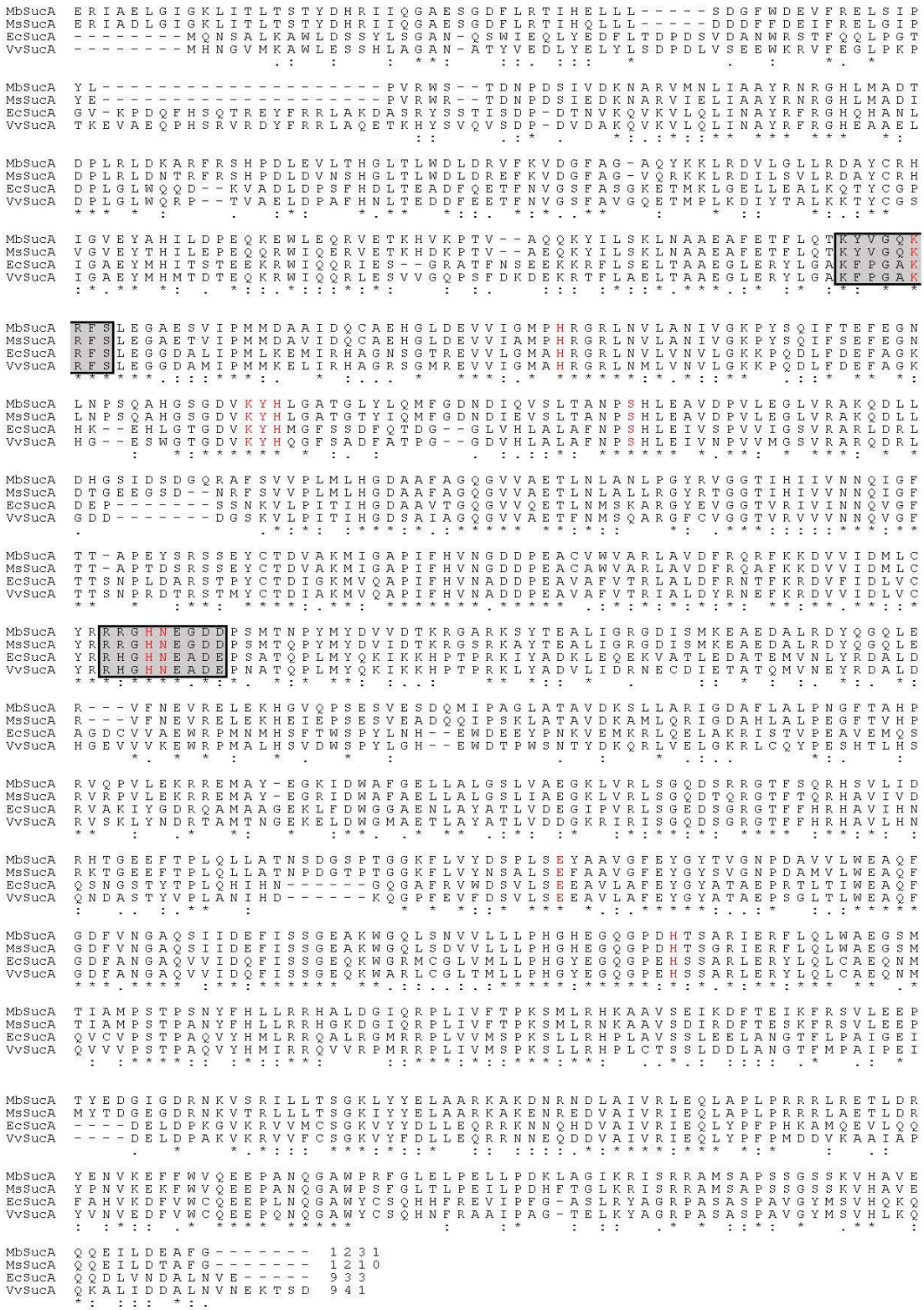

MbSucA ERIAELGIGKLITLTSTYDHRIIQGAESGDFLRTIHELLL-----SDGFWDEVFRELSIP 351
MsSucA ERIADLGIGKLITLTSTYDHRIIQGAESGDFLRTIHQLLL-----DDDFFDEIFRELGIP 332
EcSucA ---------MQNSALKAWLDSSYLSGAN-QSWIEQLYEDFLTDPDSVDANWRSTFQQLPGT 51
VvSucA ---------MHNGVMKAWLESSHLAGAN-ATYVEDLYELYLSDPDLVSEEWKRVFEGLPKP 51
MbSucA YL-------------------PVRWS--TDNPDSIVDKNARVMNLIAAYRNRGHLMADT 389
MsSucA YE-------------------PVRWR--TDNPDSIEDKNARVIELIAAYRNRGHLMADI 370
EcSucA GV-KPDQFHSQTREYFRRLAKDASRYSSTISDP-DTNVKQVKVLQLINAYRFRGHQHANL 109
VvSucA TKEVAEQPHSRVRDYFRRLAQETKHYSVQVSDP-DVDAKQVKVLQLINAYRFRGHEAAEL 110
MbSucA DPLRLDKARFRSHPDLEVLTHGLTLWDLDRVFKVDGFAG-AQYKKLRDVLGLLRDAYCRH 448
MsSucA DPLRLDNTRFRSHPDLDVNSHGLTLWDLDREFKVDGFAG-VQRKKLRDILSVLRDAYCRH 429
EcSucA DPLGLWQQD--KVADLDPSFHDLTEADFQETFNVGSFASGKETMKLGELLEALKQTYCGP 167
VvSucA DPLGLWQRP--TVAELDPAFHNLTEDDFEETFNVGSFAVGQETMPLKDIYTALKKTYCGS 168
MbSucA IGVEYAHILDPEQKEWLEQRVETKHVKPTV--AQQKYILSKLNAAEAFETFLQTKYVGQK 506
MsSucA VGVEYTHILEPEQQRWIQERVETKHDKPTV--AEQKYILSKLNAAEAFETFLQTKYVGQK 487
EcSucA IGAEYMHITSTEEKRWIQQRIES--GRATFNSEEKKRFLSELTAAEGLERYLGAKFPGAK 225
VvSucA IGAEYMHMTDTEQKRWIQQRLESVVGQPSFDKDEKRTFLAELTAAEGLERYLGAKFPGAK 228
MbSucA RFSLEGAESVIPMMDAAIDQCAEHGLDEVVIGMPHRGRLNVLANIVGKPYSQIFTEFEGN 566
MsSucA RFSLEGAETVIPMMDAVIDQCAEHGLDEVVIAMPHRGRLNVLANIVGKPYSQIFSEFEGN 547
EcSucA RFSLEGGDALIPMLKEMIRHAGNSGTREVVLGMAHRGRLNVLVNVLGKKPQDLFDEFAGK 285
VvSucA RFSLEGGDAMIPMMKELIRHAGRSGMREVVIGMAHRGRLNMLVNVLGKKPQDLFDEFAGK 288
MbSucA LNPSQAHGSGDVKYHLGATGLYLQMFGDNDIQVSLTANPSHLEAVDPVLEGLVRAKQDLLL 626
MsSucA LNPSQAHGSGDVKYHLGATGTYIQMFGDNDIEVSLTANPSHLEAVDPVLEGLVRAKQDLLL 607
EcSucA HK--EHLGTGDVKYHMGFSSDFQTDG--GLVHLALAFNPSHLEIVSPVVIGSVRARLDRL 341
VvSucA HG--ESWGTGDVKYHQGFSADFATPG--GDVHLALAFNPSHLEIVNPVVMGSVRARQDRL 344
MbSucA DHGSIDSDGQRAFSVVPLMLHGDAAFAGQGVVAETLNLANLPGYRVGGTIHIIVNNQIGF 686
MsSucA DTGEEGSD--NRFSVVPLMLHGDAAFAGQGVVAETLNLALLRGYRTGGTIHIVVNNQIGF 665
EcSucA DEP-------SSNKVLPITIHGDAAVTGQGVVQETLNMSKARGYEVGGTVRIVINNQVGF 394
VvSucA GDD-------DGSKVLPITIHGDSAIAGQGVVAETFNMSQARGFCVGGTVRVVVNNQVGF 397
MbSucA TT-APEYSRSSEYCTDVAKMIGAPIFHVNGDDPEACVWVARLAVDFRQRFKKDVVIDMLC 745
MsSucA TT-APTDSRSSEYCTDVAKMIGAPIFHVNGDDPEACAWVARLAVDFRQAFKKDVVIDMLC 724
EcSucA TTSNPLDARSTPYCTDIGKMVQAPIFHVNADDPEAVAFVTRLALDFRNTFKRDVFIDLVC 454
VvSucA TTSNPRDTRSTMYCTDIAKMVQAPIFHVNADDPEAVAFVTRIALDYRNEFKRDVVIDLVC 457
MbSucA YRRRGHNEGDDPSMTNPYMYDVVDTKRGARKSYTEALIGRGDISMKEAEDALRDYQGQLE 805
MsSucA YRRRGHNEGDDPSMTQPYMYDVIDTKRGSRKAYTEALIGRGDISMKEAEDALRDYQGQLE 784
EcSucA YRRHGHNEADEPSATQPLMYQKIKKHPTPRKIYADKLEQEKVATLEDATEMVNLYRDALD 514
VvSucA YRRHGHNEADEPNATQPLMYQKIKKHPTPRKLYADVLIDRNECDIETATQMVNEYRDALD 517
MbSucA R---VFNEVRELEKHGVQPSESVESDQMIPAGLATAVDKSLLARIGDAFLALPNGFTAHP 862
MsSucA R---VFNEVRELEKHEIEPSESVEADQQIPSKLATAVDKAMLQRIGDAHLALPEGFTVHP 841
EcSucA AGDCVVAEWRPMNMHSFTWSPYLNH--EWDEEYPNKVEMKRLQELAKRISTVPEAVEMQS 572
VvSucA HGEVVVKEWRPMALHSVDWSPYLGH--EWDTPWSNTYDKQRLVELGKRLCQYPESHTLHS 575
MbSucA RVQPVLEKRREMAY-EGKIDWAFGELLALGSLVAEGKLVRLSGQDSRRGTFSQRHSVLID 921
MsSucA RVRPVLEKRREMAY-EGRIDWAFAELLALGSLIAEGKLVRLSGQDTQRGTFTQRHAVIVD 900
EcSucA RVAKIYGDRQAMAAGEKLFDWGGAENLAYATLVDEGIPVRLSGEDSGRGTFFHRHAVIHN 632
VvSucA RVSKLYNDRTAMTNGEKELDWGMAETLAYATLVDDGKRIRISGQDSGRGTFFHRHAVLHN 635
MbSucA RHTGEEFTPLQLLATNSDGSPTGGKFLVYDSPLSEYAAVGFEYGYTVGNPDAVVLWEAQF 981
MsSucA RKTGEEFTPLQLLATNPDGTPTGGKFLVYNSALSEFAAVGFEYGYSVGNPDAMVLWEAQF 960
EcSucA QSNGSTYTPLQHIHN-------GQGAFRVWDSVLSEEAVLAFEYGYATAEPRTLTIWEAQF 686
VvSucA QNDASTYVPLANIHD-------KQGPFEVFDSVLSEEAVLAFEYGYATAEPSGLTLWEAQF 689
MbSucA GDFVNGAQSIIDEFISSGEAKWGQLSNVVLLLPHGHEGQGPDHTSARIERFLQLWAEGSM 1041
MsSucA GDFVNGAQSIIDEFISSGEAKWGQLSDVVLLLPHGHEGQGPDHTSGRIERFLQLWAEGSM 1020
EcSucA GDFANGAQVVIDQFISSGEQKWGRMCGLVMLLPHGYEGQGPEHSSARLERYLQLCAEQNM 746
VvSucA GDFANGAQVVIDQFISSGEQKWARLCGLTMLLPHGYEGQGPEHSSARLERYLQLCAEQNM 749
MbSucA TIAMPSTPSNYFHLLRRHALDGIQRPLIVFTPKSMLRHKAAVSEIKDFTEIKFRSVLEEP 1101
MsSucA TIAMPSTPANYFHLLRRHGKDGIQRPLIVFTPKSMLRNKAAVSDIRDFTESKFRSVLEEP 1080
EcSucA QVCVPSTPAQVYHMLRRQALRGMRRPLVVMSPKSLLRHPLAVSSLEELANGTFLPAIGEI 806
VvSucA QVVVPSTPAQVYHMIRRQVVRPMRRPLIVMSPKSLLRHPLCTSSLDDLANGTFMPAIPEI 809
MbSucA TYEDGIGDRNKVSRILLTSGKLYYELAARKAKDNRNDLAIVRLEQLAPLPRRRLRETLDR 1161
MsSucA MYTDGEGDRNKVTRLLLTSGKIYYELAARKAKENREDVAIVRIEQLAPLPRRRLAETLDR 1140
EcSucA ----DELDPKGVKRVVMCSGKVYYDLLEQRRKNNQHDVAIVRIEQLYPFPHKAMQEVLQQ 862
VvSucA ----DELDPAKVKRVVFCSGKVYFDLLEQRRNNEQDDVAIVRIEQLYPFPMDDVKAAIAP 865
MbSucA YENVKEFFWVQEEPANQGAWPRFGLELPELLPDKLAGIKRISRRAMSAPSSGSSKVHAVE 1221
MsSucA YPNVKEKFWVQEEPANQGAWPSFGLTLPEILPDHFTGLKRISRRAMSAPSSGSSKVHAVE 1200
EcSucA FAHVKDFVWCQEEPLNQGAWYCSQHHFREVIPFG-ASLRYAGRPASASPAVGYMSVHQKQ 921
VvSucA YVNVEDFVWCQEEPQNQGAWYCSQHNFRAAIPAG-TELKYAGRPASASPAVGYMSVHLKQ 924
MbSucA QQEILDEAFG-------- 1231
MsSucA QQEILDTAFG-------- 1210
EcSucA QQDLVNDALNVE----- 933
VvSucA QKALIDDALNVNEKTSD 941

Fig. S8. Sequence alignment of SucAs whose structures are known. Four SucAs were aligned using Clustal Omega (http://www.ebi.ac.uk/Tools/msa/clustalo/). Some conserved residues at the entrance of and within the active site were colored in red. Two signature sequence motifs that has large structural variations are indicated with boxes in grey background. MbSucA, SucA from *M. bovis*; MsSucA, SucA from *M. smegmatis*; EcSucA, SucA from *E. coli*; VvSucA, SucA from *V. vulnificus* CMCP6.

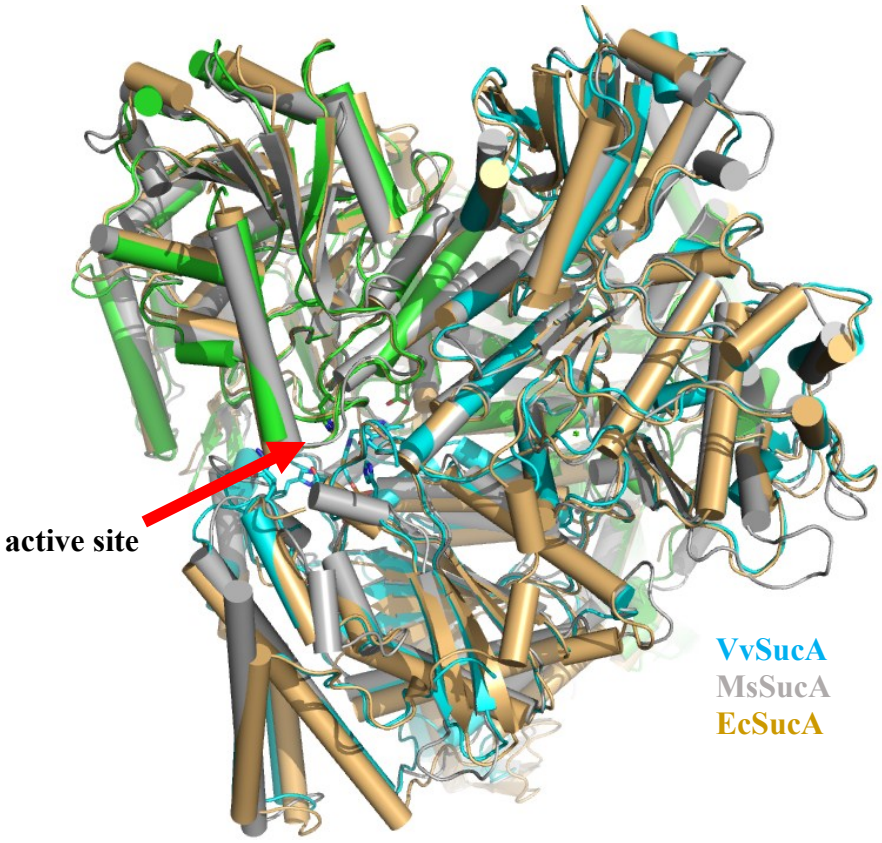


Fig. S9. Superposed SucA structures. Three SucA structures are differentiated by colors. The α-helices and β-strands are drawn with cylinders and some residues at the active sites are displayed with stick models.

References

1. C. Aslanidis and P. J. De Jong, *Nucleic Acids Res.*, **1990**, 18, 6069-6074.
2. S.N. Ho, H.D. Hunt, R.M. Horton, J.K. Pullen and L.R. Pease, *Gene* **1989**, 77, 51-59.
3. G. Bertani, *J. Bacteriol*., 1951, **62**, 293-300.
4. Z. Otwinowski and W. Minor, *Methods. Enzymol*., 1997, **276**, 307-326.
5. P. D. Adams, P. V. Afonine, G. Bunkóczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Head, L.-W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger and P. H. Zwart, *Acta Crystallogr. D*, 2010, **66**, 213-221.
6. P. Emsley and K. Cowtan, *Acta Crystallogr. D*. 2004, **60**, 2126-2132.
7. T. Gerhards, U. Mackfeld, M. Bocola, E. Lieres, W. Wiechert, M. Pohl and D. Rother, *Adv. Synth. Catal.* 2012, **354**, 2805-2820.
8. L. Zhang, R. Singh, S. D, Z. Guo, J. Li, F. Chen, Y. He, X. Guan, Y.C. Kang and J.K. Lee, *Green chem*., 2018, **20**, 230-242.
9. J.B. Siegel, A.L. Smith, S. Poust, A.J. Wargacki, A. Bar-Even, C. Louw, B.W. Shen, C.B. Eiben, H.M. Tran, E. Noor, J.L. Gallaher, J. Bale, Y. Yoshikuni, M.H. Gelb, J.D. Keasling, B.L. Stoddard, M.E. Lidstrom and D. Baker, *PNAS* 2015, **112**, 3704-3709.