Supplementary materials

Insights into the molecular mechanism of a new efficient whole-cell biocatalyst *Enterobacter ludwigii* YYP3 in 5-hydroxymethylfurfural reduction

Xin Pan[a † *], Xue Wang [a †], Sihua Wu[c d †], Lei Xu[a], Leilei Zhang[a], Zhan Zhang[e], Bingfeng Li[b], Xuejun He[b], Siyuan Chang[b *]

[a] Central Laboratory, The Affiliated Hospital of Yangzhou University, Yangzhou University, Yangzhou 225000, Jiangsu, China

[b] College of life and health, Nanjing Polytechnic Institute, 625 Geguan Road, Nanjing 210048, Jiangsu, China

[c] State Key Laboratory of Biochemical Engineering, Institute of Process Engineering, Chinese Academy of Sciences, Beijing, P. R. China

[d] Division of Molecular Science, Graduate School of Science and Technology, Gunma University, Kiryu, Gunma, Japan

[e] School of Chemistry and Chemical Engineering, Yangzhou University, Yangzhou 225000, Jiangsu, China

[†]These authors contributed equally to this article

[*]Corresponding authors: Xin Pan (xinpanphd@163.com), Siyuan Chang (fenwuyi@163.com)

16S rDNA gene sequence 1506 bp

GAGTTTGATCATGGCTTCAGATTGAACGCTGGCGGCAGGCCTAACACATGCAAGTCGA
ACGGTAGCACAGAGAGCTTGCTCTCGGGTGACGAGTGGCGGACGGGTGAATAATGTCT
GGGAAACTGCCTGATGGAGGGGGATAACTACTGGAAACGGTAGCTAATACCGCATAAC
GTCGCAAGACCAAAGAGGGGGACCTTCGGGCCTCTTGCCATCAGATGTGCCCAGATGG
GATTAGCTAGTAGGTGGGGTAACGGCTCACCTAGGCGACGATCCCTAGCTGGTCTGAG
AGGATGACCAGCCACACTGGAACTGAGACACGGTCCAGACTCCTACGGGAGGCAGCA
GTGGGGAATATTGCACAATGGGCGCAAGCCTGATGCAGCCATGCCGCGTGTATGAAGA
AGGCCTTCGGGTTGTAAAGTACTTTCAGCGGGGAGGAAGGTGGTTGTGGTTAATAACC
GCAGCAATTGACGTTACCCGCAGAAGAAGCACCGGCTAACTCCGTGCCAGCAGCCGC
GGTAATACGGAGGGTGCAAGCGTTAATCGGAATTACTGGGCGTAAAGCGCACGCAGGC
GGTCTGTCAAGTCGGATGTGAAATCCCCGGGCTCAACCTGGGAACCGCATTCGAGACT
GGCAGGCTAGAGTCTTGTAGAGGGGGGTAGAATTCCAGGTGTAGCGGTGAAATGCGTA
GGGATCTGGAGGAATACCGGTGGCGAAGGCGGCCCCCTGGACAAAGACTGACGCTCA
GGTGCGAAAGCGTGGGGAGCAAACAGGATTAGATACCCTGGTAGTCCACGCCGTAAAC
GATGTCGACTTGGAGGTTGTGCCCTTGAGGCGTGGCTTCCGGAGCTAACGCGTTAAGT
CGACCGCCTGGGGGAGTACGGCCGCAAGGTTAAAACTCAAATGAATTGGCGGGGGCC
CGCACAAGCGGTGGAGCATGTGGTTTAATTCGATGCAACGCGAAGAACCTTACCTTAC
TCTTGACATCCAGAGAACTTTCCAGAAGATGGATTGGTGCCTTCGGGAACTCTGAGAA
CAGGTGCTGCATGGCTGTCGTCAGCTCGTGTTGTGAAATGTTGGGTTAAGTCCCGCAA
CGAGCGCAACCCTTATCCTTTGTTGCCAGCGGTCCGGCCGGGAACTCAAAGGAGACTG
CCAGTGATAAACTGGAGGAAGGTGGGGATGACGTCAAGTCATCATGGCCCTTACGAGT
AGGGCTACACACGTGCTACAATGGCGCATACAAAGAGAAGCGAACTCGCGAGAGCAA
GCGGACCTCATAAAGTGCGTCGTAGTCCGGATTGGAGTCTGCAACTCGACTCCATGAA
GTCGGAATCGCTAGTAATCGTAGATCAGAATGCTACGGTGAATACGTTCCCGGGCCTTG
TACACACCGCCCGTCACACCATGGGAGTGGGTTGCAAAAGAAGTAGGTAGCTTAACCT
TCGGGAGGGCGCTTACCACTTTGTGATTCATGACTGGGGTGAAGTCGTAACAAGGTAA
CC

The GenBank accession number for the 16S rDNA gene sequence of the strain *E. ludwigii* YYP3 is MK968765.

Full-length gene sequence of SDR family oxidoreductase *El*SDR-ykvO

ATGAGCCGTCTGCAGGGTAAACGTGCGCTGATCACCGGTGGTACCTCTGGTATCGGTCT
GGCGACCGCGAAACTGTTCGTTGCTGAAGGCGCGCGTGTTATGGTTACCGGCATCAAC
CCGGATTCTGTTGCGAAAGCGAAACTGGAACTGGGTAAAGATGTTGTTGTTGTTAGCG
CGGATTCTGCGGATGTTAACGCGCAGAAAGCGCTGGCGCAGACCGTGCAGGAACACTT
CGGTGAACTGGACATCGCGTTCCTGAACGCGGGTATCAGCATGTACATGCCGATCGAA
GCGTGGACCGAAGAAATGTTCGATCGTATCTACGATATCAACGTTAAAGGTCCGTACTT
CCTGATGCAGGCGCTGCTGCCGGTTTTCGCGTCTAGCGCGAGCGTTGTTTTCAACACCA
GCATCAACGCGCACACCGGTCCGGTTAACAGCTCTGTTTACGGTAGCACCAAAGCGGC
TCTGCTGAACATGTCTAAAACCCTGTCTAACGAACTGCTGAGCCGTGGTATCCGTATCA
ACGCGGTTTCTCCGGGTCCGGTTGATACCCCGCTGTACGATAAAGCGGGTATCCCGGTT
GAATACCACGATCAGGTTATGAAAGATATCGTTGCTACCATCCCGGCGGGTCGTTTCGG
CAAACCGCAGGAAGTTGCGCAGGCGGTTCTGTACTTCGCAAGCGATGCTAGCGCGTGG
ACCGTTGGTAGCGAAATCATCATCGATGGTGGTGTTAGCATCTAA

Full-length gene sequence of SDR family oxidoreductase *El*SDR-SSP1627

ATGACCGATAACATCATCGGTAAAGTTATCGTTATCACCGGTGCGAGCAGCGGCATGGG
TGAAGCGGCGGCGCGTTACCTGGCGGAAAAAGGCGCGAAAGTTGTTATGGCGGCGCG
TCGTATCGATCGTATCGAAGCGATTGCTTCTGAACTGCAGAAACAGAACAAAGAAGCG
ATCGCGGTTGCGACCGATGTTACCAAACTGGATGATGTTAACAACCTGATCGAAACCGC
GGTGAACAAATTCGGTCGTGTTGATGTTCTGATCAACAACGCGGGTCTGATGCCGCTG
AGCCGTCTGGAACAGGGCAACGTTGATGAATGGAACCAGATGATTGATGTTAACCTGC
GTGGCGTTCTGCACGGTATTGCAGCGGTTCTGCCGTATATGAAAAGCCAGAAAACCGG
CCACATTATCAACACCGCTAGCGTTGCGGCGCACCTGGTTTTCCAGAGCAGCGCGGTTT
ACTCTGCGACCAAATTCGCGGTTCGTGCGCTGACCGATGGCCTGCGTCAGGAAATGGC
GGCGCACAACATCCGTGTTACCCTGGTTTCTCCGGGCGCGGTTAAAACCGAACTGCTG
GAACACATTACCGATAAAGATGTTAAAGCGGCGAACCAGGATTACGTTGAAAACATCG
GCGTTCCGGCGGAAACCTTCGCGCGTATGGTTGCGTTCGCGATCAACGAACCGGAAGA
TGTTGGTGTTAGCGAAATCATCTTCCGTCCGACCGCGCAGGAACTGTAA

Table S1 Primers used for mutation analysis.

Mutant		Primer sequence (5'-3')
S143A	F	CACATTATCAACACCGCTGCCGTTGCGGCGCACC
	R	GGTGCGCCGCAACGGCAGCGGTGTTGATAATGTG
Y156A	F	TCCAGAGCAGCGCGGTTGCCTCTGCGACCAAATTC
	R	GAATTTGGTCGCAGAGGCAACCGCGCTGCTCTGGA
K160A	F	GGTTTACTCTGCGACCGCATTCGCGGTTCGTGCG
	R	CGCACGAACCGCGAATGCGGTCGCAGAGTAAACC

Table S2 The templates used to build the conformation of enzymes analyzed in this study.

Enzyme	Template (PDB)	Resolution (Å)	Sequence Identity (%)	Ligands in active center
reductive aminase *Asp*RedAm	5G6R	1.82	100	NADPH
alcohol dehydrogenase *Mg*AAD1669	3UYI	2.31	35.93	None
NADH-dependent reductase CmCR	3GDF	2.50	50.19	None
*El*SDR-ykvO	4FGS	1.8	43.21	None
*El*SDR-SSP1627	3TFO	2.1	50.84	None
	6IHI	1.8	34.48	NADPH, (13R, 17S)-ethyl secol

Table S3 Representative chemical processes for the reduction of HMF to BHMF.

Catalyst	Solvent	H-donor	HMF (mM)	Reaction conditions	Yield (%)	Selectivity (%)	Time (h)	Space time yield ($g\ L^{-1}\cdot h^{-1}$)	Ref.
Noble metal catalyst									
Pd/C	Tetrahydrofuran	H_2	238	80 °C, 100 bar	97	82	20	1.2	[1]
Pt/MCM-41	Water	H_2	2000	35 °C, 8 bar	100	98.9	2	126.7	[2]
$Pt/CeO_2–ZrO_2$	Ethanol	H_2	206	170 °C, 10 bar	97.0	100	8	3.2	[3]
$Ru(OH)_x/ZrO_2$	1-butanol	H_2	323	120 °C, 15 bar	99.0	99.0	6	6.8	[4]
$Ir–ReO_x/SiO_2$	Water	H_2	1000	30 °C, 8 bar	>99	>99.0	6	20.9	[5]
Au/Al_2O_3	Water	H_2	100	120 °C, 65 bar	100	96.0	2	6.2	[6]
Non-noble metal catalyst									
Cu-ZnO	1,4-dioxane	H_2	340	100 °C, 15 bar	100	99.1	2	21.6	[7]
Cu/SiO_2	1,4-dioxane	H_2	158	100 °C, 50 bar	99.6	97.5	2	9.8	[8]
Cu/Al_2O_3	Methanol	H_2	50	130 °C, 30 bar	>99	93.0	1	5.9	[9]
Ni-Fe/CNT	1-butanol	H_2	200	110 °C, 30 bar	100	96.1	18	1.4	[10]
Co-400	Methanol	H_2	100	90 °C, 20 bar	94	93.0	1	11.2	[11]
Acid-base catalyst									
$ZrO(OH)_2$	Ethanol	Ethanol	162	150 °C, 1 bar	94.1	83.7	2.5	6.5	[12]
ZrPN	Isopropanol	Isopropanol	250	140 °C, 1 bar	99	98	2	15.5	[13]
Hf-DTMP	sec-butanol	sec-butanol	132	130 °C, 1 bar	99.1	96.8	3	5.4	[14]
Hf–LigS	2-propanol	2-propanol	100	100 °C, 1 bar	97.3	92.2	2	5.7	[15]

Table S4 Scale-up production of BHMF by fed-batch via various whole-cell biocatalysts.

Whole-cell biocatalyst	HMF (mM)	Fed-batch times	Time (h)	BHMF (mM)	Ref.
M. guilliermondii SC1103	200	4	24.5	191	[16]
Immobilized *M. guilliermondii* SC1103	200	2	7	176	[17]
A. subglaciale F134	500	5	15	430	[18]
B. contaminans NJPI-15	700	7	48	656	[19]
Recombinant *S. cerevisiae*[a]	450	3	23	345	[20]
E. ludwigii YYP3	300	3	9	290	This study

[a] Recombinant *S. cerevisiae* containing the overexpressed alcohol dehydrogenase *Mg*AAD1669.

Table S5 Genome characteristics of strain *E. ludwigii* YYP3.

Characteristics	**Value**
Raw reads size (bp) in Illumina platform	1,298,455,500
Clean reads size (bp)	1,261,843,111
Sequencing depth	259×
Total sequence length (bp) in Oxford Nanopore ONT platform	1,311,836,297
GC content (%)	54.37
Chromosome size (bp)	4,854,702
CDS in chromosome	4551
Total size (bp)	4,272,300
Mean length (bp)	938.76
Gene GC content (%)	46.22
5S rRNAs	9
16S rRNAs	8
23S rRNAs	8
tRNAs	84
Other ncRNAs	107
VFDB numbers	127
Antibiotic Resistance Genes	100
Protein coding genes	4551
nr annotation (genome)	4499
Swiss-Prot annotation (genome)	3941
COGs annotation (genome)	4267
GO annotation (genome)	3641
KEGG annotation (genome)	3031
All annotated genes	4502

Table S6 The significantly enriched top 20 GO terms of biological processes and molecular functions.

Category	GO.ID	Term	Up	Down	DEG	Total	P-value
Nucleic acid binding							
MF	GO:0003677	DNA binding	49	6	55	475	1.81819e-06
MF	GO:0003676	nucleic acid binding	52	11	63	634	5.04099e-05
MF	GO:0043565	sequence-specific DNA binding	13	2	15	80	0.000103422
Biofilm formation							
BP	GO:0042710	biofilm formation	10	0	10	25	1.83723e-06
BP	GO:0044010	single-species biofilm formation	7	0	7	13	6.29173e-06
Cellular response to stimulus							
BP	GO:0070887	cellular response to chemical stimulus	7	1	8	22	4.71408e-05
MF	GO:0004364	glutathione transferase activity	5	1	6	13	6.9897e-05
BP	GO:0046677	response to antibiotic	8	0	8	24	9.65017e-05
Amino acid biosynthetic/metabolic process							
BP	GO:0009082	branched-chain amino acid biosynthetic process	0	10	10	23	7.24067e-07
BP	GO:0009081	branched-chain amino acid metabolic process	0	10	10	26	2.81357e-06
BP	GO:0000105	histidine biosynthetic process	0	6	6	9	5.67821e-06
BP	GO:0006551	leucine metabolic process	0	5	5	6	7.00398e-06
BP	GO:0009098	leucine biosynthetic process	0	5	5	6	7.00398e-06
BP	GO:1901607	alpha-amino acid biosynthetic process	2	22	24	144	1.69594e-05
BP	GO:0008652	cellular amino acid biosynthetic process	3	23	26	169	3.22802e-05
Other process							
BP	GO:0098630	aggregation of unicellular organisms	10	0	10	25	1.83723e-06
BP	GO:0098743	cell aggregation	10	0	10	25	1.83723e-06
BP	GO:0051704	multi-organism process	12	0	12	38	3.15336e-06
BP	GO:0051703	intraspecies interaction between organisms	7	0	7	13	6.29173e-06
BP	GO:0044764	multi-organism cellular process	9	0	9	26	2.45817e-05

MF: molecular function

BP: biological process

DEG: differentially expressed gene

Table S7 Several potential defense mechanisms of microorganisms for the toxic furanic aldehydes.

Microorganism	Defense process	Related Gene (Protein name)	Ref.
Saccharomyces cerevisiae	Enhance pentose phosphate pathway	*ZWF1*, *GND1*, *GND2* (6-phosphogluconate dehydrogenase)	21-24
	Maintain redox homeostasis	*HYR1* (thiol peroxidase), *TRX1*, *TRR1* (thioredoxin), *SOD1*, *SOD2* (superoxide dismutase), *CTA1*, *CTT1* (catalase), *GLR1* (glutathione reductase), *GPX1*, *GPX2* (glutathione peroxidase), *GTT2*, *GTO1*, *ECM4* (glutathione transferase)	
	Enhance glycolysis and TCA pathways	*PYK2* (pyruvate kinase), *CIT1* (citrate synthase), *ACO1* (aconitase), *FUM1* (fumarase)	
	Enhance cell membrane adaptation and biosynthesis	*INO1* (inositol-3-phosphate synthase), *RSB1* (specific transporter ATPase gene), *ICT1* (acyl-CoA-dependent lysophosphatidic acid acyltransferase)	
	Promote DNA replication and repair	*FMP16* (Found in mitochondrial proteome protein 16)	
	Increase stress-response protein expression	*HSP26, HSP82, HSP104, SSA4* (heat-shock proteins)	
Rhizopus oryzae	Repress protein synthesis Reduce aerobic respiration pathway and xylose metabolism Strengthen the stabilization of phosphatidylcholine bilayers	-	25
Zymomonas mobilis	Alter or modify the composition and structure of the cell membrane	*fliC* (flagellin domain-containing protein), *MreC* (rod shape-determining protein), *MscS* (ion channel protein), *ostA* (organic solvent tolerance protein), *lgt* (lipoprotein)	26
	Repress protein synthesis	*rpsD*, *rpsF*, *rplI*, *rbsR*, *frr*, *rbfA* (ribosomal proteins), *proS*, *alaS*, *leuS*, *glyS*, *pheT*, *valS* (tRNA synthetases), *glnA*, *trpA*, *trpB*, *argG*, *gltB*, *ilvE*, *glnB*, *serA*, *serC* (amino acid metabolism-related genes)	
	Repress terpenoid biosynthesis	*hpnD*, *hpnE*, *dxs*, *dxr* (terpenoid biosynthesis-related proteins)	
	Promote DNA replication and repair	*addA* (double-strand break repair helicase), *addB* (double-strand break repair protein), *ung* (uracil-DNA glycosylase), *radC* (DNA repair protein), *mutL* (DNA mismatch repair enzyme)	
	Increase universal stress gene expression	*dnaJ* (haperone protein), *lon* (ATP-dependent protease)	
	Upregulate transcriptional regulator	*Lys*R family, *Lyt*R family, GntR family, *Tet*R family, *Lac*I family, *rpoD*	
	Upregulate putative respiratory gene	*rnfA*, *rnfB* (putative NADH/ubiquinone oxidoreductase subunit)	

Table S8 The downregulated genes in the nucleic acid, biofilm, cellular response, and amino acid process.

Gene_ID	Gene	Protein name	Gene_ID	Gene	Protein name
Nucleic acid binding					
gene1246	*seqA*	replication initiation negative regulator SeqA	gene1510	*rpsA*	30S ribosomal protein S1
gene1492	*lrp*	leucine-responsive transcriptional regulator Lrp	gene1707	*rne*	flagellar basal body P-ring protein FlgI
gene1511	*ihfB*	integration host factor subunit beta	gene1842	*pheS*	phenylalanine--tRNA ligase subunit alpha
gene213	*hupA*	YdeI family stress tolerance OB fold protein	gene3069	*rplY*	50S ribosomal protein L25
gene2518	*narL*	two-component system response regulator NarL	gene753	*tsf*	translation elongation factor Ts
gene4085	*fis*	DNA-binding transcriptional regulator Fis			
Cellular response to stimulus					
gene3956	*garL*	2-dehydro-3-deoxyglucarate aldolase			
gene1944	*gstA*	glutathione transferase GstA			
Amino acid biosynthetic/metabolic process					
gene31	*ilvN*	acetolactate synthase small subunit	gene2929	*hisC*	histidinol-phosphate transaminase bifunctional
gene4224	*asd*	aspartate-semialdehyde dehydrogenase	gene2930	*hisB*	histidinol-phosphatase/imidazoleglycerol-phosphate dehydratase HisB
gene4398	*ilvC*	ketol-acid reductoisomerase	gene2933	*hisF*	imidazole glycerol phosphate synthase subunit HisF
gene4402	*ilvE*	branched-chain-amino-acid transaminase	gene2934	*hisI*	bifunctional phosphoribosyl-AMP cyclohydrolase/phosphoribosyl-ATP diphosphatase HisIE
gene4403	*ilvM*	acetolactate synthase 2 small subunit	gene2418	*trpC*	bifunctional indole-3-glycerol-phosphate synthase TrpC/phosphoribosylanthranilate isomerase TrpF
gene4404	*ilvG*	acetolactate synthase 2 catalytic subunit	gene3416	*glyA*	serine hydroxymethyl transferase
gene664	*leuD*	3-isopropylmalate dehydratase small subunit	gene4522	*asnA*	aspartate--ammonia ligase
gene665	*leuC*	3-isopropylmalate dehydratase large subunit	gene642	*folA*	type 3 dihydrofolate reductase
gene666	*leuB*	3-isopropylmalate dehydrogenase	gene724	*panD*	aspartate 1-decarboxylase
gene667	*leuA*	2-isopropylmalate synthase	gene749	*dapD*	2,3,4,5-tetrahydropyridine-2,6-dicarboxylate N-succinyltransferase
gene2927	*hisG*	ATP phosphoribosyltransferase	gene725	*panC*	pantoate--beta-alanine ligase
gene2928	*hisD*	bifunctional histidinal dehydrogenase/ histidinol dehydrogenase			

Table S9 The upregulated genes in the nucleic acid, biofilm, cellular response, and amino acid process.

Gene_ID	Gene	Protein name	Gene_ID	Gene	Protein name
Nucleic acid binding					
gene1047		TetR Family Transcriptional Regulator	gene2439	*anr*	Crp/Fnr family transcriptional regulator
gene1049	*ramA*	transcriptional regulator	gene2448		AraC family transcriptional regulator
gene1066	*bvgA*	response regulator	gene270	*soxS*	superoxide response transcriptional regulator SoxS
gene1085	*marA*	AraC family transcriptional regulator	gene271	*soxR*	redox-sensitive transcriptional activator SoxR
gene1188	*ybdO*	LysR family transcriptional regulator	gene2978		winged helix-turn-helix domain-containing protein
gene1207	*ybeF*	YbeF family transcriptional regulator	gene3188	*int*	tyrosine-type recombinase/integrase
gene1290	*ybgS*	YbgS	gene3200	*glxR*	Crp/Fnr family transcriptional regulator
gene1304	*ybcK*	recombinase family protein	gene3201	*decR*	Lrp/AsnC family transcriptional regulator
gene138	*yiaU*	DNA-binding transcriptional LysR family regulator	gene3260	*mocR*	PLP-dependent aminotransferase family protein
gene1663	*csgD*	transcriptional regulator CsgD	gene3265		hypothetical protein F917_01746
gene1883	*ariR*	two-component-system connector protein AriR	gene3266	*yoeC*	tyrosine-type recombinase/integrase
gene1887	*bluR*	MerR family transcriptional regulator	gene332	*ecnR*	LuxR C-terminal-related transcriptional regulator
gene1954	*ydgT*	transcription modulator YdgT	gene3485	*intA*	integrase
gene1980	*abgR*	LysR family transcriptional regulator	gene3525	*stpA*	DNA-binding protein stpA
gene2043	*yafC*	LysR family transcriptional regulator	gene3644	*gcvA*	glycine cleavage system transcriptional regulator GcvA
gene2092		TetR/AcrR family transcriptional regulator	gene4056	*argR*	transcriptional regulator ArgR
gene2158	*tehB*	tellurite resistance methyltransferase TehB	gene459	*intB*	integrase arm-type DNA-binding domain-containing protein
gene2182	*marA*	Multiple antibiotic resistance protein marA	gene461		restriction endonuclease subunit S
gene2183	*marR*	Multiple antibiotic resistance protein marR	gene463	*MJ0130*	restriction endonuclease subunit S
gene2192	*yneJ*	LysR family transcriptional regulator	gene606	*arcA*	two-component system response regulator ArcA
gene2287		TetR family transcriptional regulator	gene857	*yqeI*	winged helix-turn-helix domain-containing protein
gene2296	*yczG*	helix-turn-helix domain-containing protein	gene975	*hha*	hemolysin expression modulator Hha
gene2312	*mcbR*	GntR family transcriptional regulator	gene976	*tomB*	Hha toxicity modulator TomB
gene2343		DUF1294 domain-containing protein	gene2398	*yciH*	stress response translation initiation inhibitor YciH

gene2385	*rob*	helix-turn-helix domain-containing protein	gene3436	*lepA*	translation elongation factor 4
gene2389	*ykgA*	helix-turn-helix domain-containing protein	gene3546	*csrA*	carbon storage regulator CsrA
Biofilm formation					
gene1662	*csgE*	curli production assembly/transport protein CsgE	gene1885	*ycgZ*	hypothetical protein
gene1663	*csgD*	transcriptional regulator CsgD	gene2136	*ydeI*	YdeI family stress tolerance OB fold protein
gene1665	*csgB*	Minor curlin subunit	gene2543	*ychH*	stress-induced protein YchH
gene1680	*yceO*	YceO family protein	gene4057	*yhcN*	peroxide/acid stress response protein YhcN
gene1883	*ariR*	two-component-system connector protein AriR	gene976	*tomB*	Hha toxicity modulator TomB
Cellular response to stimulus					
gene1883	*ariR*	two-component-system connector protein AriR	gene1417	*gstB*	glutathione S-transferase family protein
gene1885	*ycgZ*	hypothetical protein	gene2045	*GSTO1*	glutathione S-transferase family protein
gene2136	*ydeI*	YdeI family stress tolerance OB fold protein	gene2298	*yncG*	glutathione S-transferase family protein
gene2183	*marR*	Multiple antibiotic resistance protein marR	gene272	*gst3*	glutathione S-transferase
gene2543	*ychH*	stress-induced protein YchH	gene2026	*cat*	type A chloramphenicol O-acetyltransferase
gene4057	*yhcN*	peroxide/acid stress response protein YhcN	gene2181	*marB*	multiple antibiotic resistance protein MarB
gene4308	*treF*	alpha,alpha-trehalase	gene2182	*marA*	multiple antibiotic resistance protein marA
gene1102	*gstB*	glutathione transferase GstA			
Amino acid biosynthetic/metabolic process					
gene1887	*bluR*	MerR family transcriptional regulator			
gene4056	*argR*	transcriptional regulator ArgR			
gene892	*aroM*	protein AroM			

Table S10 The upregulated enzymes with redox activity in the oxidation-reduction process of GO terms

Gene_ID	Gene	Protein name	FPKM value					
			Control1	Control2	Control3	HMF1	HMF2	HMF3
gene2390	*FabI*	enoyl-ACP reductase FabI	318.59	269.68	259.98	345.77	445.31	616.92
gene1476	*hcp*	Hydroxylamine reductase	71.84	198.51	194.41	282.41	268.68	230.24
gene2044	*ykvO*	SDR family oxidoreductase	50.89	101.48	115.14	125.31	138.49	137.82
gene3827	*yqhD*	alcohol dehydrogenase	84.72	213.18	225.92	329.62	316.49	340.01
gene347	*queG*	tRNA epoxyqueuosine(34) reductase QueG	65.50	154.40	174.96	286.96	356.10	298.69
gene429	*nrdG*	anaerobic ribonucleoside-triphosphate reductase-activating protein	131.36	223.44	230.24	350.92	412.32	343.26
gene1131	*Molybdopterin*	molybdopterin-dependent oxidoreductase	21.88	48.23	58.66	73.93	121.91	150.09
gene2046	*SSP1627*	SDR family oxidoreductase	54.30	115.22	125.86	173.44	261.85	285.80
gene1189	*ahpC*	alkyl hydroperoxide reductase subunit C	100.68	158.08	168.13	243.93	243.69	292.67
gene1416	*yliI*	PQQ-dependent sugar dehydrogenase	44.54	178.13	160.96	286.47	308.94	445.98

Table S11 Enzyme activities and steady-state kinetic parameters of *El*SDR-ykvO, *El*SDR-SSP1627, and *El*SDR-SSP1627 mutants in the HMF reduction.

Enzyme	Enzyme activity[a] (mU mg^{-1})		Steady-state kinetic parameter[b]		
	NADPH	NADH	K_m (mM)	k_{cat} (s^{-1})	k_{cat}/K_m (mM^{-1} s^{-1})
*El*SDR-ykvO	298.56	12.22	0.1096	0.667	6.09
*El*SDR-SSP1627	466.87	7.56	0.0850	0.942	11.76
*El*SDR-SSP1627-S143A	NA[c]				
*El*SDR-SSP1627-Y156A	NA				
*El*SDR-SSP1627-K160A	NA				

[a]Reaction conditions: 0.4 mL Tris-HCl buffer (100 mM, pH 8), 5 mM HMF, 0.2 mM NADPH or NADH, 30°C, and 20 μg mL^{-1} purified enzyme. The values are the average of three independent experiments.

[b]Reaction conditions: 0.4 mL Tris-HCl buffer (100 mM, pH 8), 20 μM to 5 mM HMF, 0.2 mM NADPH, 30°C, and 20 μg mL^{-1} purified enzyme. The values are the average of three independent experiments.

[c]NA: not active.

Table S12 The residue identity and RMSD value of the enzymes analyzed in this study.

	*El*SDR-ykvO		*El*SDR-SSP1627	
	Identity (%)	RMSD (Å)	Identity (%)	RMSD (Å)
NADH-dependent reductase CmCR	27.9	0.88	22.9	0.97
alcohol dehydrogenase *Mg*AAD1669	13.5	16.72	8.6	15.05
reductive aminase *Asp*RedAm	12.9	15.65	12.5	14.14
*El*SDR-ykvO	-	-	25.0	1.056
*El*SDR-SSP1627	25.0	1.056	-	-

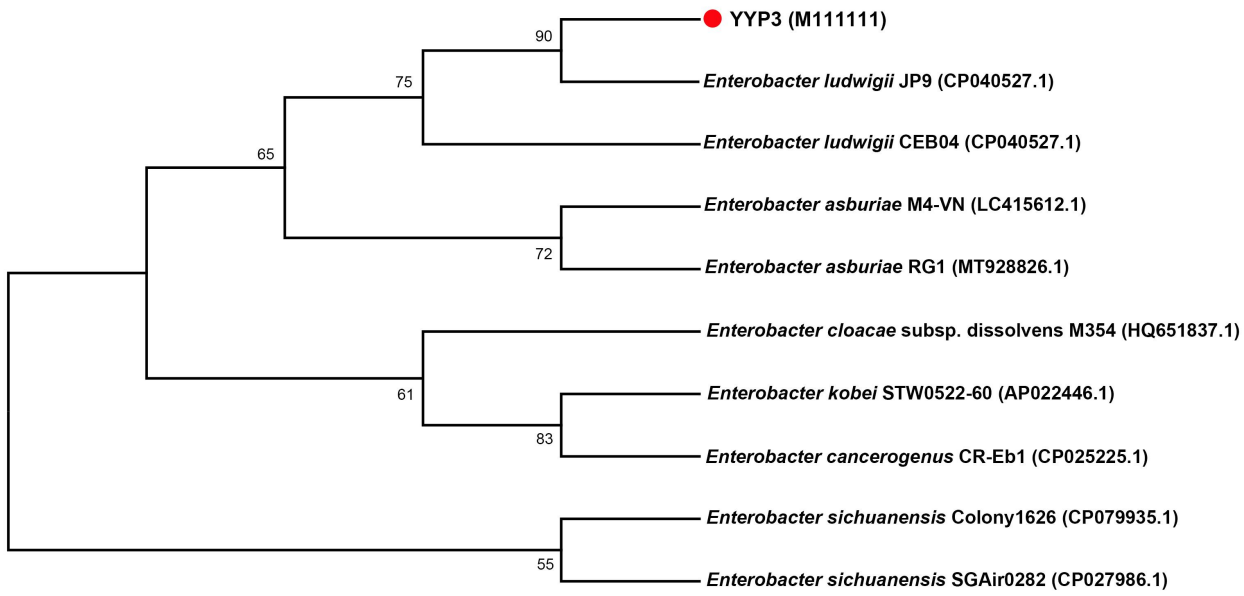

Fig. S1 Phylogenetic tree of YYP3 based on the 16S rDNA gene sequence.

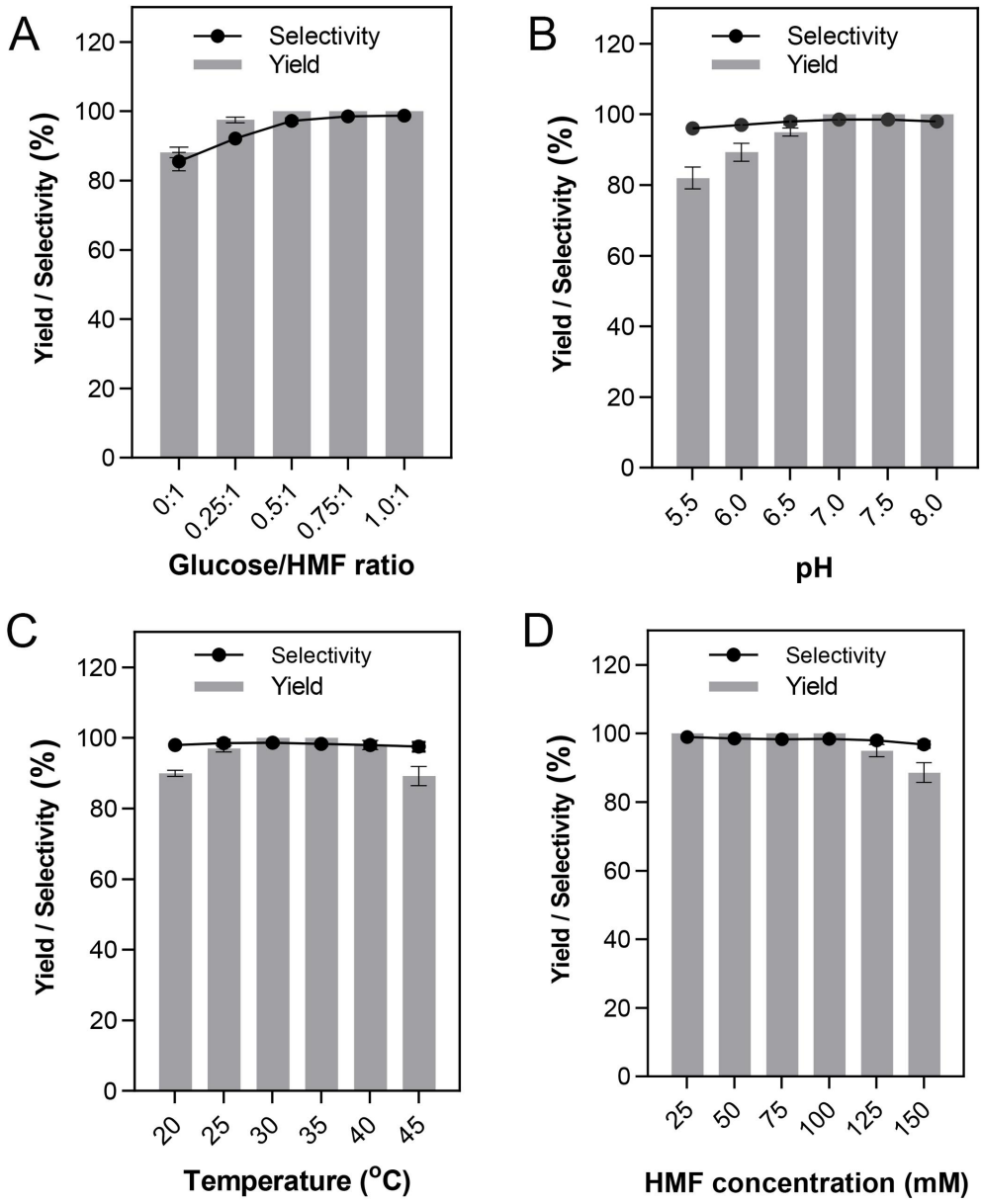

Fig. S2 Multiple steps of BHMF production optimization. (A) Ratio of glucose to HMF, (B) initial pH, (C) temperature, and (D) HMF concentration. The initial production conditions: pH 7.0, 30°C, 50 mM HMF, and 20 mg mL^{-1} (wet weight) of cells.

Different molar ratios of glucose to HMF, from 0:1 to 1:1, were tested and the results showed that a small amount of glucose supplementation (0.25:1) led to a clearly increased yield (97.5%) and selectivity (92.2%). When the glucose concentration was 37.5 mM (the molar ratio of glucose to HMF was 0.75:1), BHMF was produced with a yield >99% and 98.5% selectivity within 1.5 h (A). The BHMF yields increased continually from 82.1% to >99% with a variation of pH from 5.5 to 7.0, while the highest yields (>99%) were maintained at pH 7.0–8.0. Notably, excellent BHMF

selectivities (96.1%–98.5%) were observed across the entire pH range (B). The impact of reaction temperature on the biosynthesis of BHMF was examined from 20°C to 45°C. *E. ludwigii* YYP3 displayed the highest yield and selectivity at 30°C–35°C, but when the temperature exceeded 40°C, the catalytic efficiency decreased significantly (C). When the HMF concentration increased from 25 mM to 100 mM, *E. ludwigii* YYP3 retained high yield (>99%) and selectivity (98.5 %). When the HMF concentration increased more than 100 mM, the decreased yield and selectivity were observed (D).

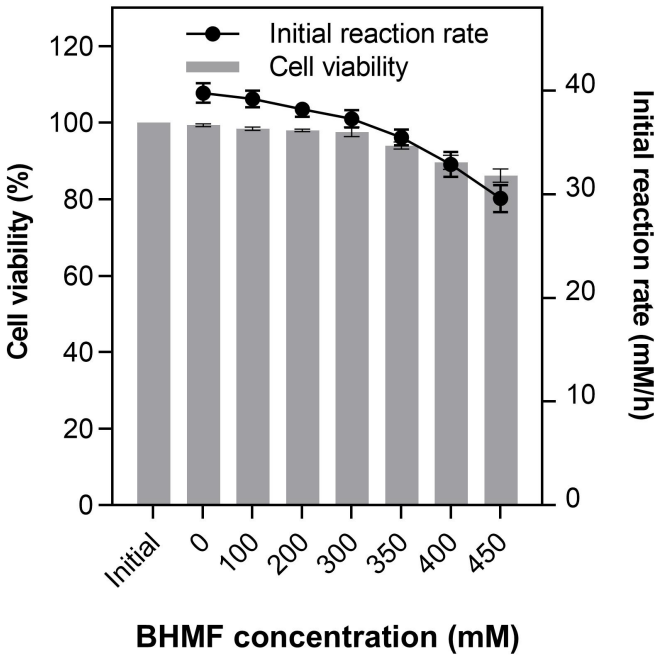

Fig. S3 Effects of various concentrations of BHMF on the whole-cell conversion of HMF. Conditions: 50 mM HMF, 0–450 mM BHMF, glucose/HMF ratio of 0.75:1, pH 7.0, 30°C, and 20 mg mL^{-1} (wet weight) of cells. Cell viability was measured under the above conditions without HMF.

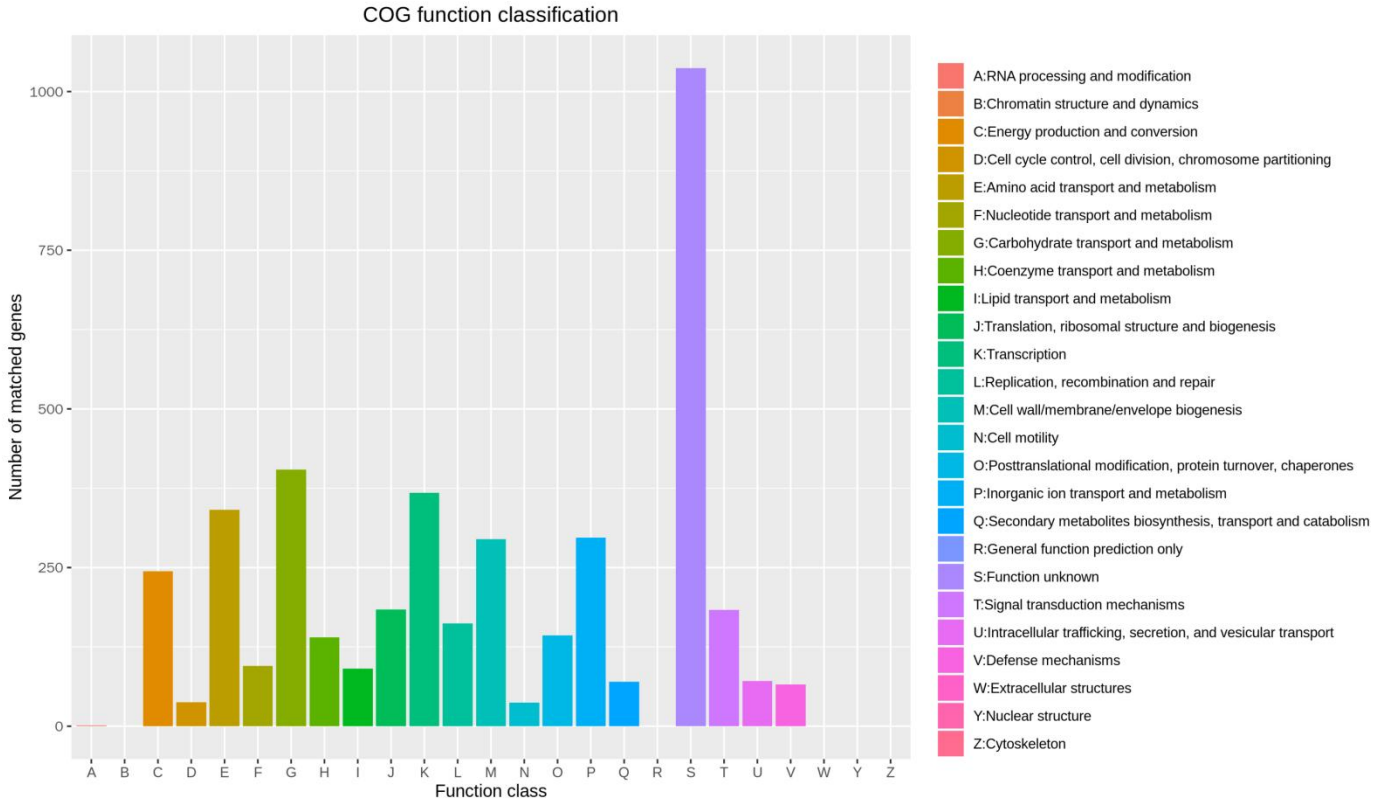

Fig. S4 Genome functional annotation of the *E. ludwigii* YYP3 chromosome against the COG database.

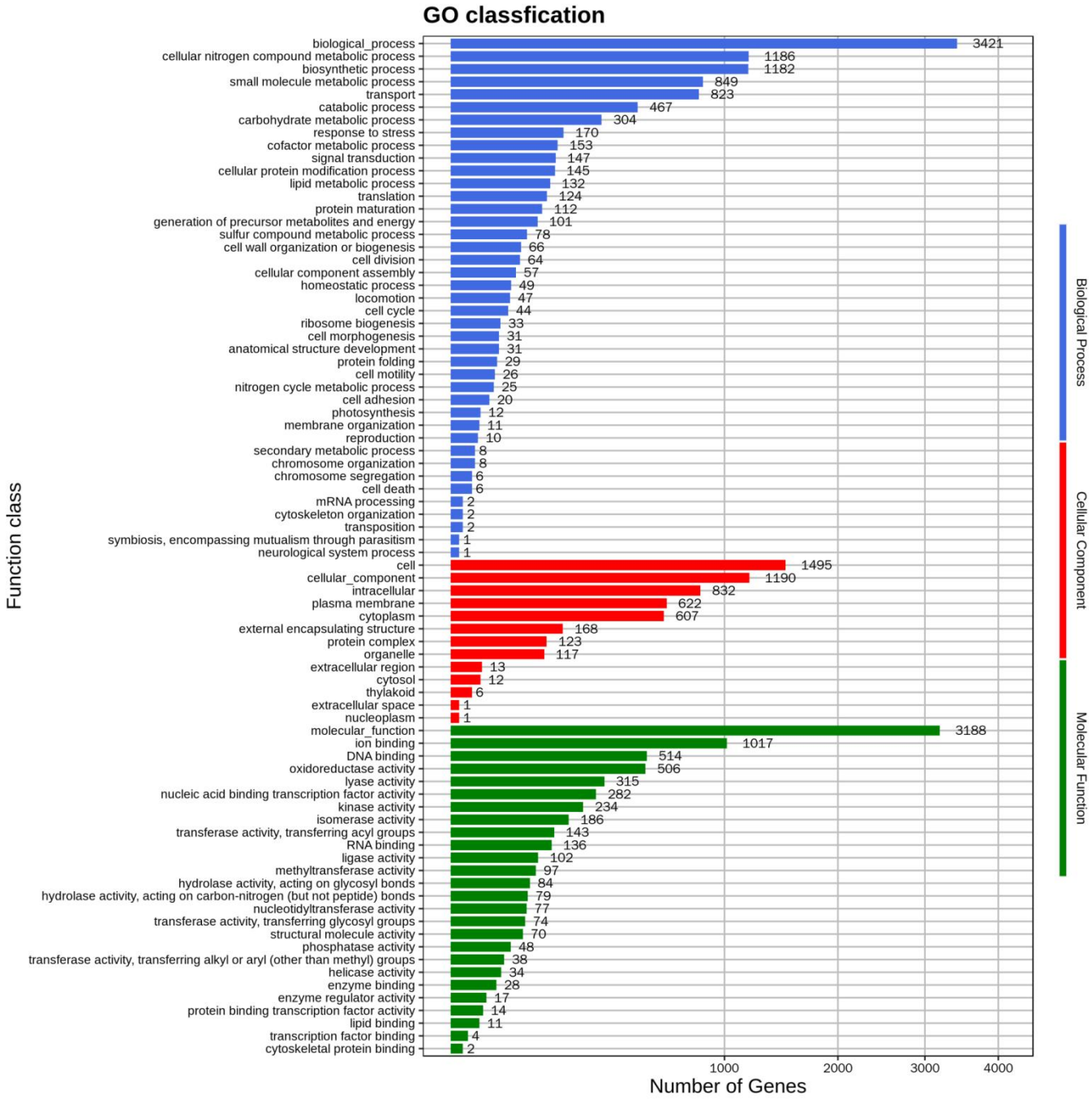

Fig. S5 Genome functional annotation of the *E. ludwigii* YYP3 chromosome against the GO database.

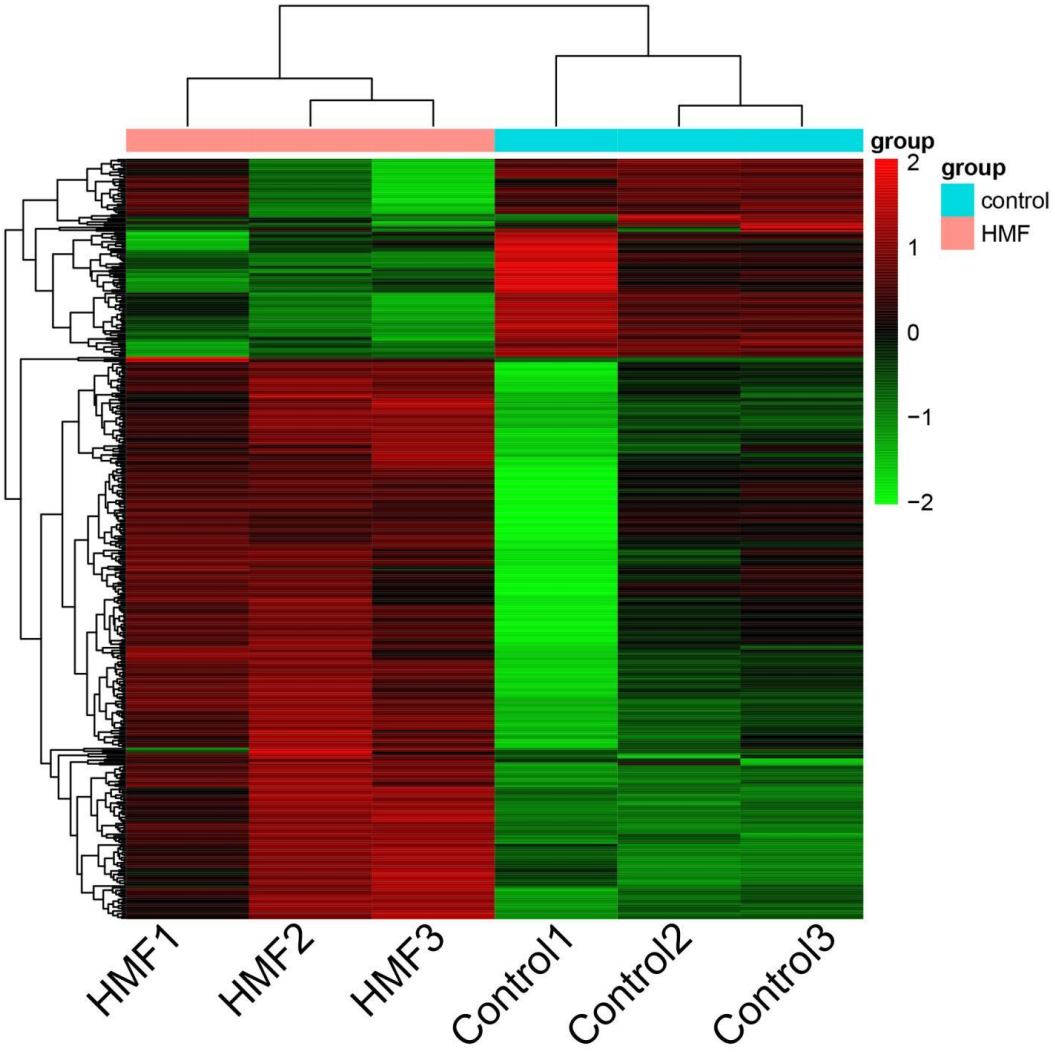

Fig. S6 Hierarchical cluster heat map of gene expression in *E. ludwigii* YYP3. HMF groups were incubated for 1h under the optimum reaction conditions with 100 mM HMF; control groups were incubated for 1h under the optimum reaction conditions without HMF.

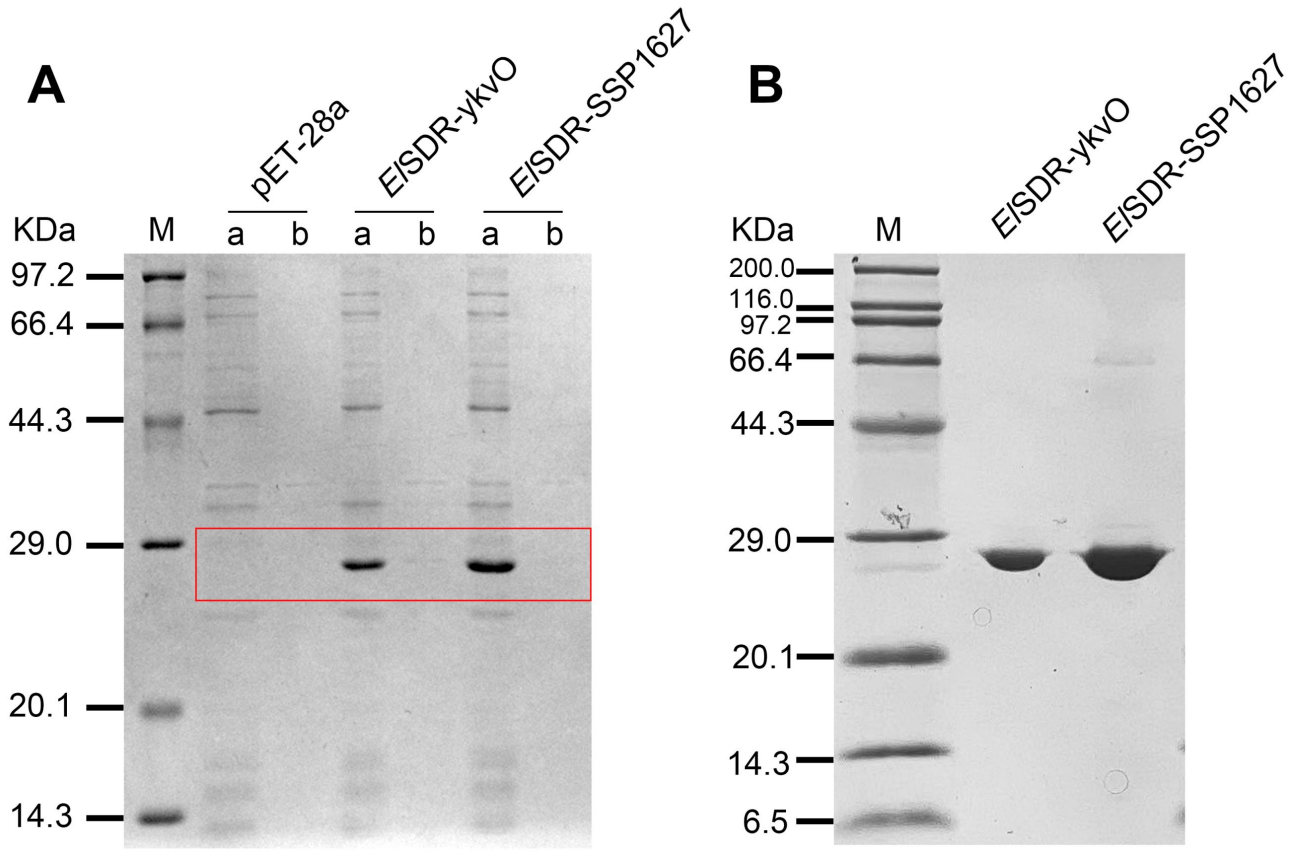

Fig. S7 (A) The heterologous expression of *El*SDR-SSP1627 and *El*SDR-ykvO in *E. coli* BL21 (DE3). M, protein marker; "a" indicates the soluble fractions and "b" indicates the insoluble fractions. The *E. coli* BL21 (DE3) harboring the empty vector pET28a was used as the control. (B) The *El*SDR-SSP1627 and *El*SDR-ykvO were purified by His-tag affinity chromatography.

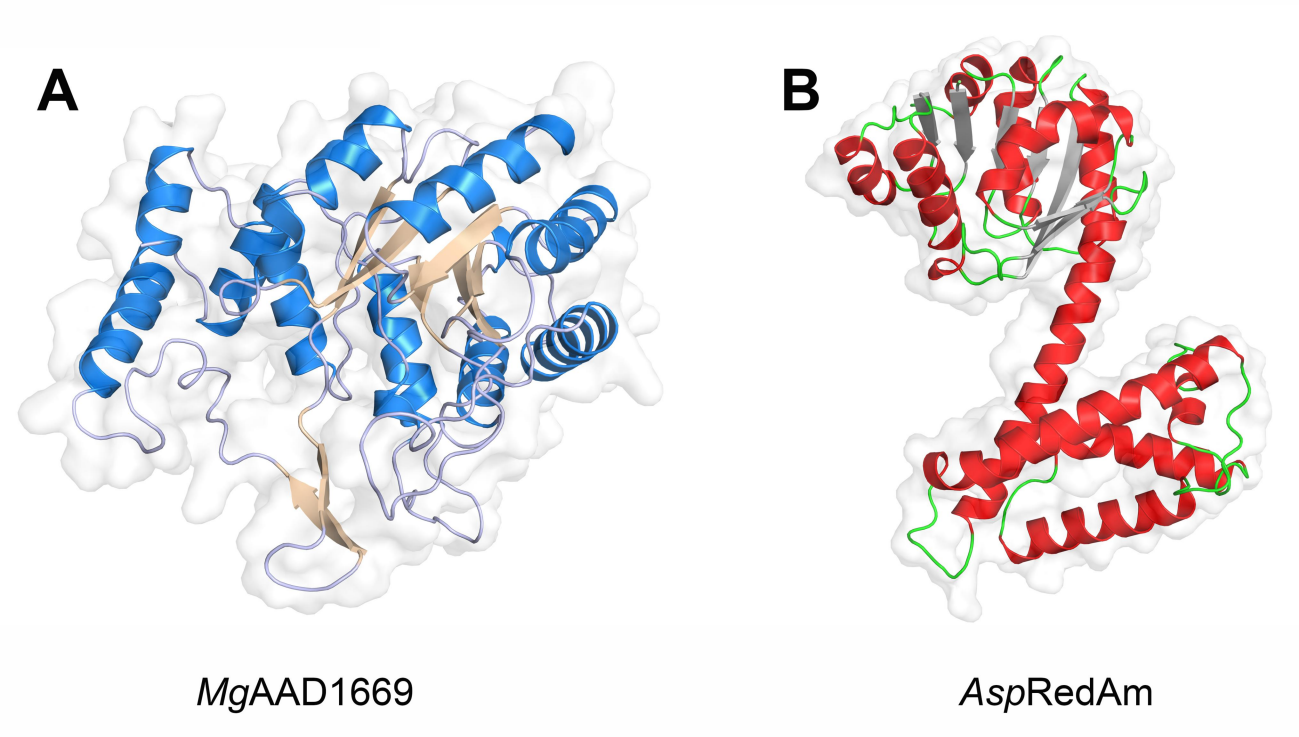

Fig. S8 (A) The predicted structure of *Mg*AAD1669 by homology modeling. (B) The predicted structure of *Asp*RedAm by homology modeling.

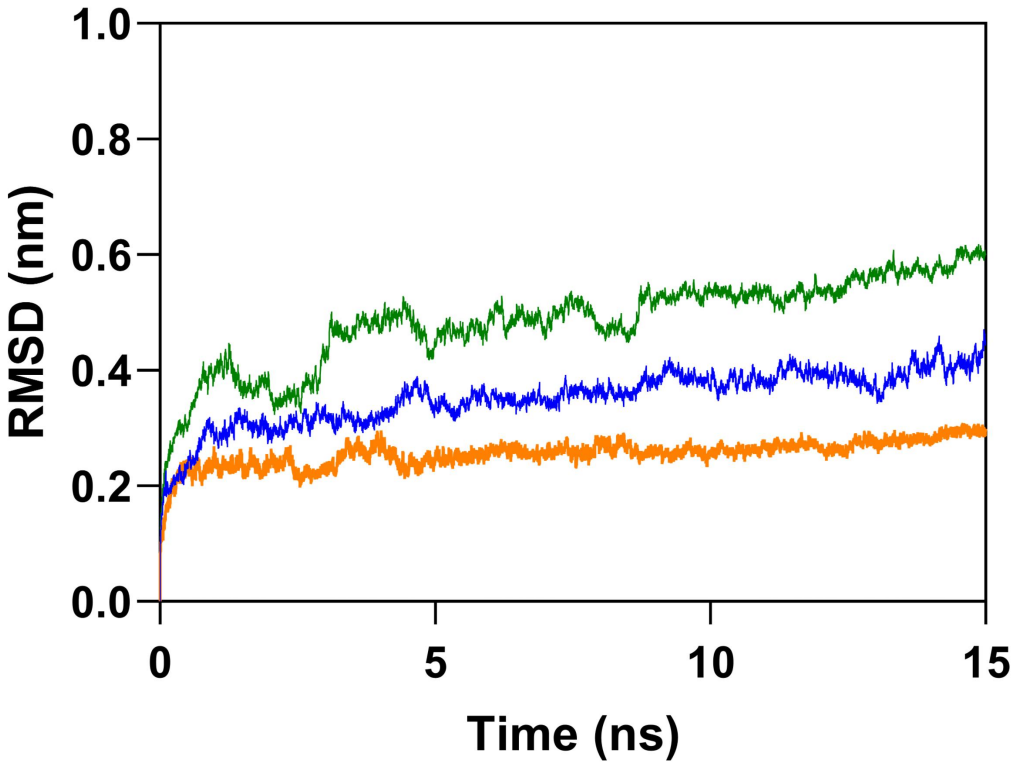

Fig. S9 The RMSD values of *El*SDR-SSP1627, *El*SDR-ykvO and CmCR by molecular dynamic simulation.

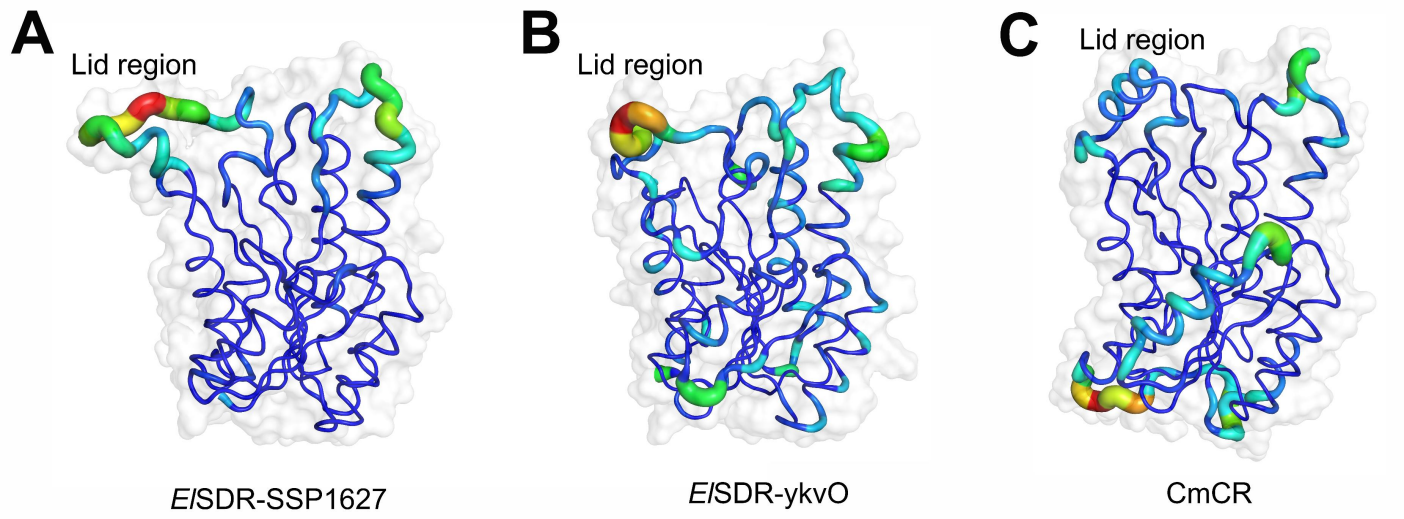

Fig. S10 Structures of *El*SDR-SSP1627 (A), *El*SDR-ykvO (B), and CmCR (C) colored by B-factor (low B-factor in blue, high B-factor in red). The closer to red, the greater the flexibility.

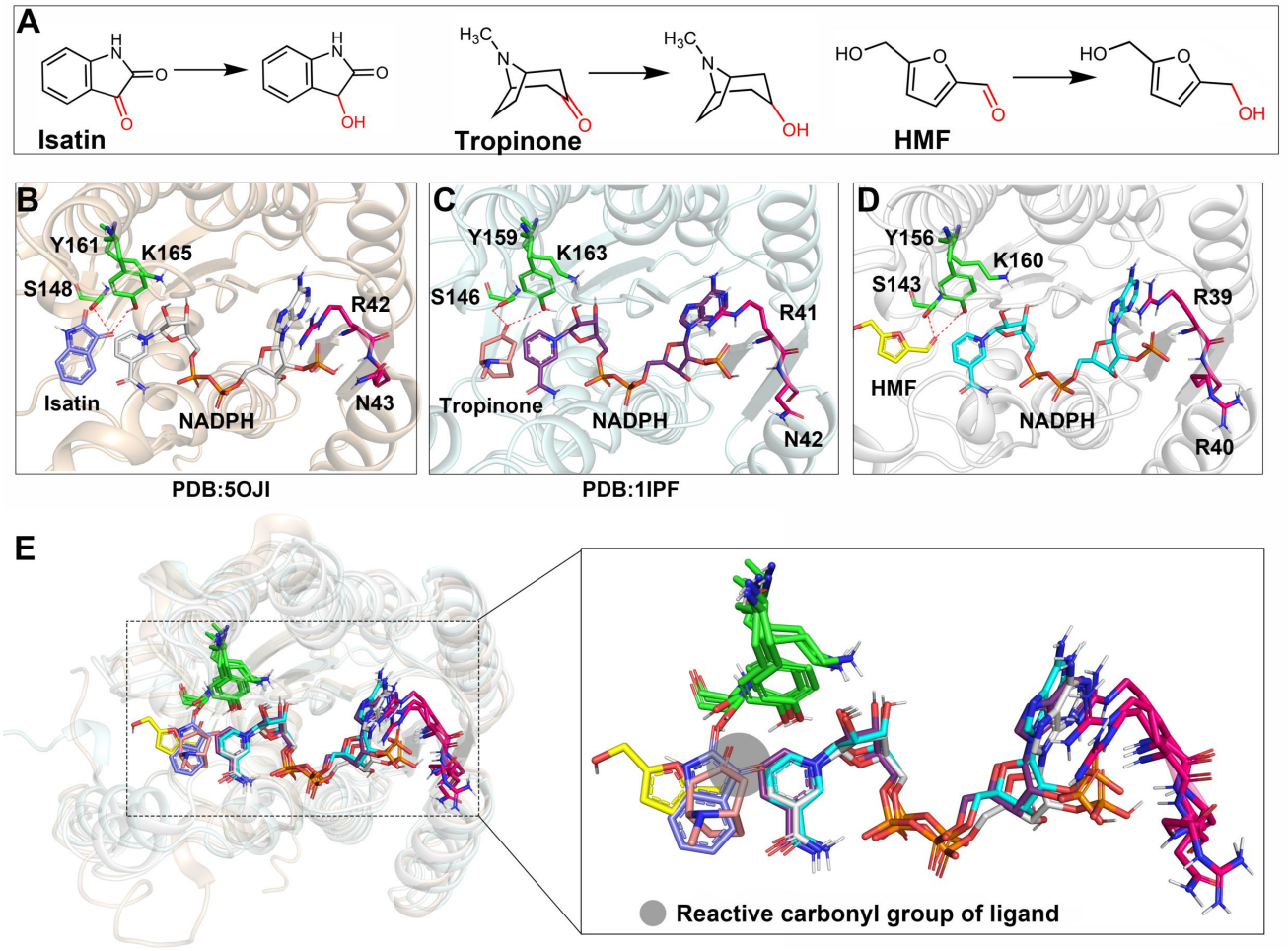

Fig. S11 (A) The reduction substrates of DHRS4, TR-II, and *El*SDR-SSP1627. (B) Binding mode analysis of the DHRS4 complex crystal structure (PDB ID: 5OJI). (C) Binding mode analysis of the TR-II complex crystal structure (PDB ID: 1IPF). (D) Docking results analysis of the *El*SDR-SSP1627 complex. (E) Superimposition of the complex structures of DHRS4, TR-II, and *El*SDR-SSP1627.

References

1. M. Audemar, K. D. O. Vigier, J. M. Clacens, F. De Campo and F. Jérôme, *Green Chem.*, 2014, **16**, 4110-4114.
2. M. Chatterjee, T. Ishizaka and H. Kawanami, *Green Chem.*, 2014, **16**, 4734-4739.
3. K. V. Vikanova, M. S. Chernova, E. A. Redina, G. I. Kapustin, O. P. Tkachenko and L. M. Kustov, *J. Chem. Technol. Biotechnol.*, 2021, **96**, 2421-2425.
4. J. Han, Y. H. Kim, H. S. Jang, S. Y. Hwang, J. Jegal, J. W. Kim and Y. S. Lee, *RSC Adv.*, 2016, **6**, 93394-93397.
5. M. Tamura, K. Tokonami, Y. Nakagawa and K. Tomishige, *Chem. Commun.*, 2013, **49**, 7034-7036.
6. J. Ohyama, A. Esaki, Y. Yamamoto, S. Arai and A. Satsuma, *RSC Adv.*, 2013, **3**, 1033-1036.
7. Y. Zhu, X. Kong, H. Zheng, G. Ding, Y. Zhu and Y. W. Li, *Catal. Sci. Technol.*, 2015, **5**, 4208-4217.
8. Y. Zhu, X. Kong, J. Yin, R. You, B. Zhang, H. Zheng, X. Wen, Y. Zhu and Y. W. Li, *J. Catal.*, 2017, **353**, 315-324.
9. K. T. V. Rao, Y. Hu, Z. Yuan, Y. Zhang and C. C. Xu, *Appl. Catal., A*, 2021, **609**, 117892.
10. L. Yu, L. He, J. Chen, J. Zheng, L. Ye, H. Lin and Y. Yuan, *ChemCatChem*, 2015, **7**, 1701-1707.
11. X. L. Li, K. Zhang, S. Y. Chen, C. Li, F. Li, H. J. Xu and Y. Fu, *Green Chem.*, 2018, **20**, 1095-1105.
12. W. Hao, W. Li, X. Tang, X. Zeng, Y. Sun, S. Liu and L. Lin, *Green Chem.*, 2016, **18**, 1080-1088.
13. H. Li, J. He, A. Riisager, S. Saravanamurugan, B. Song and S. Yang, *ACS Catal.*, 2016, **6**, 7722-7727.
14. L. Hu, X. Dai, N. Li, X. Tang and Y. Jiang, *Sustainable Energy Fuels*, 2019, **3**, 1033-1041.
15. S. Zhou, F. Dai, Y. Chen, C. Dang, C. Zhang, D. Liu and H. Qi, *Green Chem.*, 2019, **21**, 1421-1431.
16. Y. M. Li, X. Y. Zhang, N. Li, P. Xu, W. Y. Lou and M. H. Zong, *ChemSusChem*, 2017, **10**, 372-378.
17. Z. H. Xu, A. D. Cheng, X. P. Xing, M. H. Zong, Y. P. Bai and N. Li, *Bioresour. Technol.* ,2018, **262**, 177-183.
18. D. Chen, R. Cang, Z. D. Zhang, H. Huang, Z. G. Zhang and X. J. Ji, *Mol. Catal.*, 2021, **500**, 111341.
19. S. Chang, X. He, B. Li and X. Pan, *Front. Chem.*, 2021, **9**, 635191.
20. Z. H. Xia, M. H. Zong and N. Li, *Enzyme Microb. Technol.*, 2020, **134**, 109491.
21. Z. L. Liu, X. Huang, Q. Zhou and J. Xu, *Appl. Microbiol. Biotechnol.*, 2019, **103**, 5781-5796.
22. M. Ma and Z. L. Liu, *BMC Genomics*, 2010, **11**, 660.
23. P. Unrean, J. Gätgens, B. Klein, S. Noack and V. Champreda, *FEMS Yeast Res.*, 2018, **18**, foy106.
24. Z. L. Liu and M. Ma, *Appl. Microbiol. Biotechnol.*, 2020, **104**, 3473-3492.
25. X. Pan, H. Liu, J. Liu, C. Wang and J. Wen, *Bioresour. Technol.*, 2016, **222**, 24-32.
26. M. X. He, B. Wu, Z. X. Shui, Q. C. Hu, W. G. Wang, F. R. Tan, X. Y. Tang, Q. L. Zhu, K. Pan, Q. Li and X. H. Su, *Appl. Microbiol. Biotechnol.*, 2012, **95**, 189-199.