Combinatorial enzyme-catalysis for bioproduction of ginsenoside Compound K

Wenhua Yang[1], Jianli Zhou[2], Qiuya Gu[1], Jean Damascene Harindintwali[3], Xiaobo Liu[4*] & Xiaobin Yu[1*]

[1]Key Laboratory of Industrial Biotechnology, Ministry of Education, School of Biotechnology, Jiangnan University, Wuxi 214000, Jiangsu, China.

[2]Guizhou Province Key Laboratory of Fermentation Engineering and Biopharmacy, School of Liquor and Food Engineering, Guizhou University, Guiyang 550003, Gui Zhou, China

[3]CAS Key Laboratory of Soil Environment and Pollution Remediation, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, Jiangsu, China

[4]School of Environmental and Biological Engineering, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, China.

*Corresponding author. Email: xbyu@jiangnan.edu.cn; xbliu@njust.edu.cn

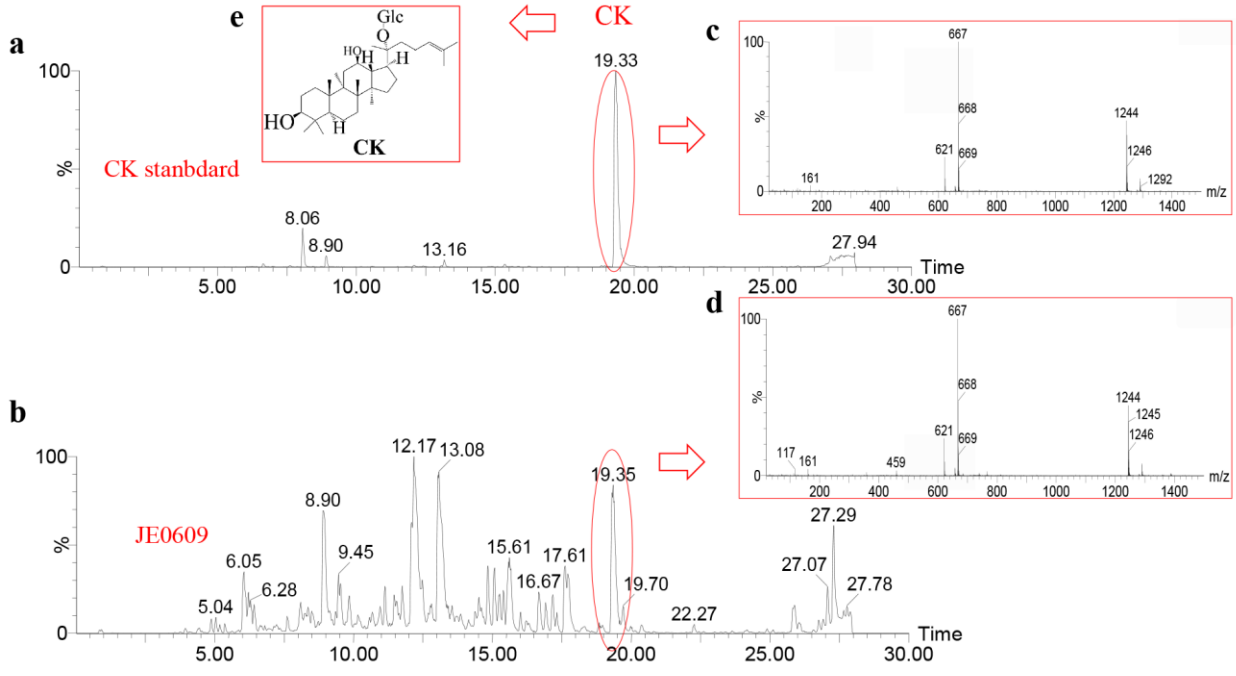
This materials include content:

Supplementary Figs. 1 to 15

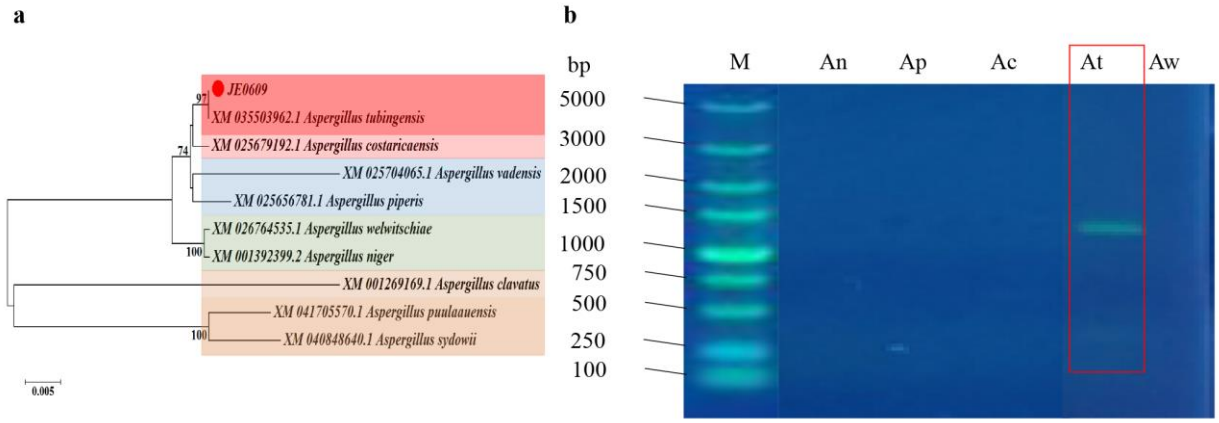
Supplementary Tables 1 to 14

Supplementary References (1 to 23)

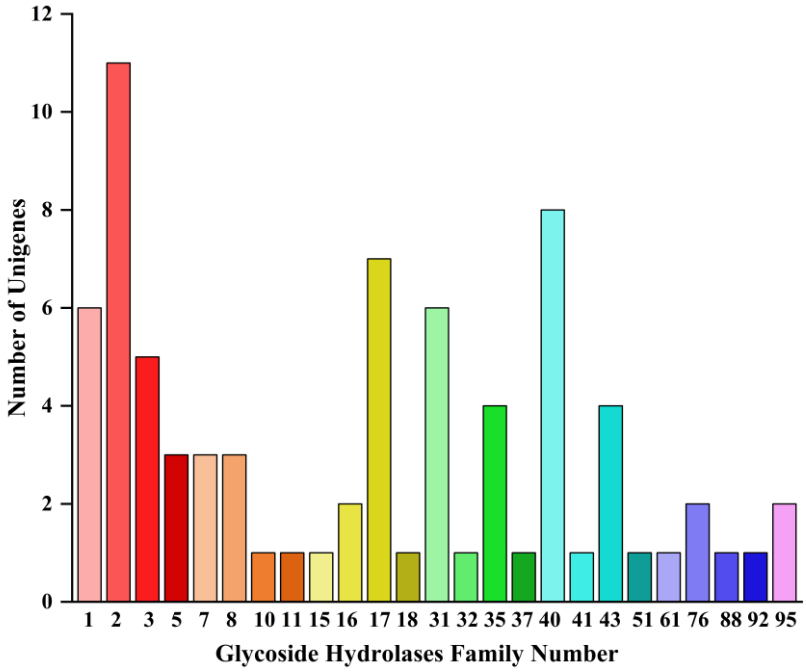
Supplementary Figures



Supplementary Fig. 1 UPLC-Q-TOF-MS/MS analysis of CK standard and the sample of JE0609 extracellular enzymes catalyzed ginseng extracts. a Ion flow profiles of CK standard. **b** Ion flow profiles of samples of JE0609 extracellular enzyme-catalyzed ginseng extracts. **c** Mass spectra of CK standard at 19.33 min. **d** Mass spectra of samples of ginseng extract catalyzed by JE0609 extracellular enzyme at 19.35 min. **e** The chemical structure formula of CK.



Supplementary Fig. 2 Identification of JE0609. a Evolutionary trees were constructed based on ITS sequencing results. MEGA 6.0 was used for the construction of evolutionary trees. The scale bars represent 0.005 substitutions per site. The tree was constructed using a neighbor-joining method with bootstrap values of 1000 replications. **b** Nucleic acid electropherogram of the β-tubulin genes of different *Aspergillus* sp. strains obtained by PCR using the cDNA of JE0609 as a template and according to the primers of Table S2.



Supplementary Fig. 3 Statistics of glycoside hydrolases family in JE0609 transcriptome. The results showed that 76 of the 73045 annotated genes were classified as Glycoside Hydrolases.

706-130kDa

gi|2077896 Mass: 93646 Score: 277 Matches: 6(0) Sequences: 6(0)

beta-D-glucosidase [Aspergillus kawachii]

Check to include this hit in error tolerant search or archive report

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
9	1392.7445	1391.7372	1391.7521	-10.73	0	31	81	5	U	K.GADIQLGPAAGPLGR.S
17	1492.6939	1491.6867	1491.6895	-1.89	0	40	8.4	1	U	K.HYIAYEQEHFR.Q
35	1862.8950	1861.8877	1861.9185	-16.52	0	55	0.24	1	U	K.TMHELYLWPFADAIR.A
43	2069.0360	2068.0287	2068.0477	-9.16	0	53	0.31	1	U	K.VAGDEVPQLYVSLGGPNEPK.I
60	2386.1562	2385.1490	2385.1747	-10.79	1	49	0.68	1	U	K.AMGQEFSDKGADIQLGPAAGPLGR.S
62	2769.3043	2768.2970	2768.3253	-10.23	0	49	0.65	1	U	K.TASVSLVFVNADSGEGYINVDGNLGDR.K

803-110kDa

AMYG ASPAW Mass: 68267 Score: 366 Matches: 6(6) Sequences: 6(6)

Glucoamylase OS=Aspergillus awamori OX=105351 GN=GLAA PE=1 SV=1

Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
1	863.4895	862.4822	862.4913	-10.47	0	45	0.032	1	U	K.TLVDLFR.N
9	1485.7626	1484.7554	1484.7736	-12.26	1	56	0.0023	1	U	R.ALANHKEVVDSFR.S
14	1633.8017	1632.7944	1632.8107	-9.99	0	87	1.6e-06	1	U	R.ATLDSWLSNEATVAR.T
17	1864.9379	1863.9306	1863.9591	-15.28	1	45	0.02	1	U	R.DLTWSTAALLTANNRR.N
20	1966.9458	1965.9386	1965.9643	-13.11	0	48	0.01	1	U	R.SIYTLNDGLSDSEAVAVGR.Y
22	1983.9063	1982.8990	1982.9235	-12.34	0	84	2.3e-06	1	U	K.FNVDETAYTGSWGRPQR.D

803-65kDa

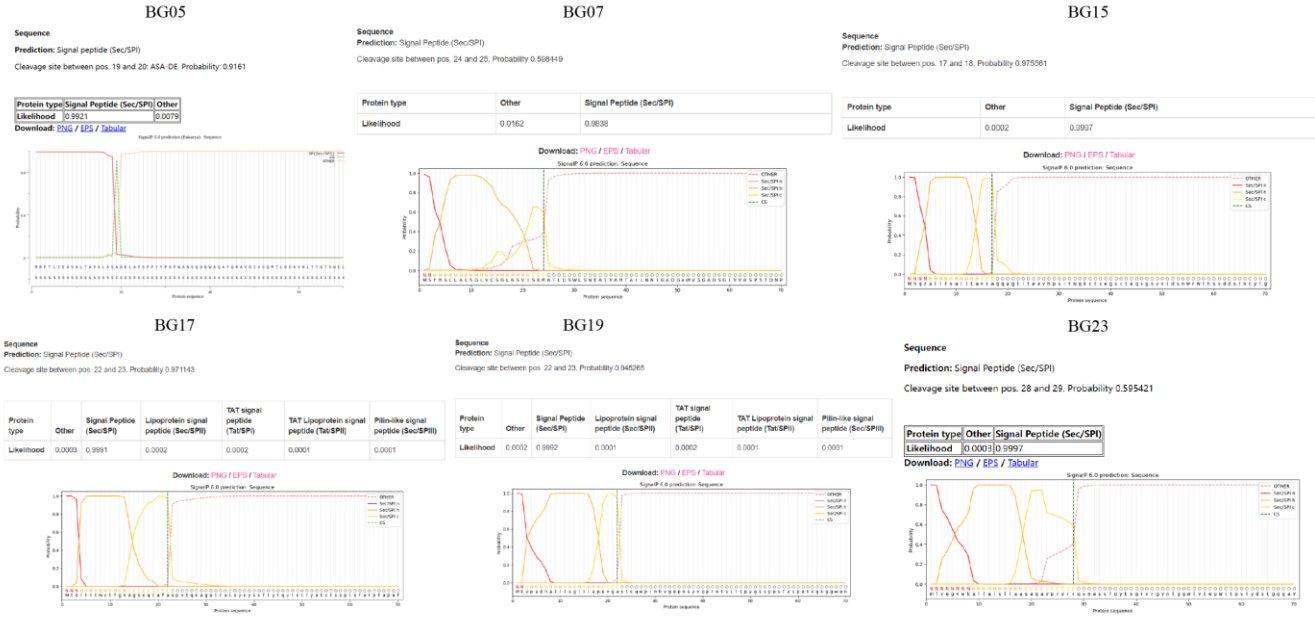
CBHA ASPNC Mass: 48216 Score: 48 Matches: 1(1) Sequences: 1(1)

Probable 1,4-beta-D-glucan cellobiohydrolase A OS=Aspergillus niger (strain CBS 513.88 / FGSC A1513) OX=425011 GN=cbhA PE=3 SV=1

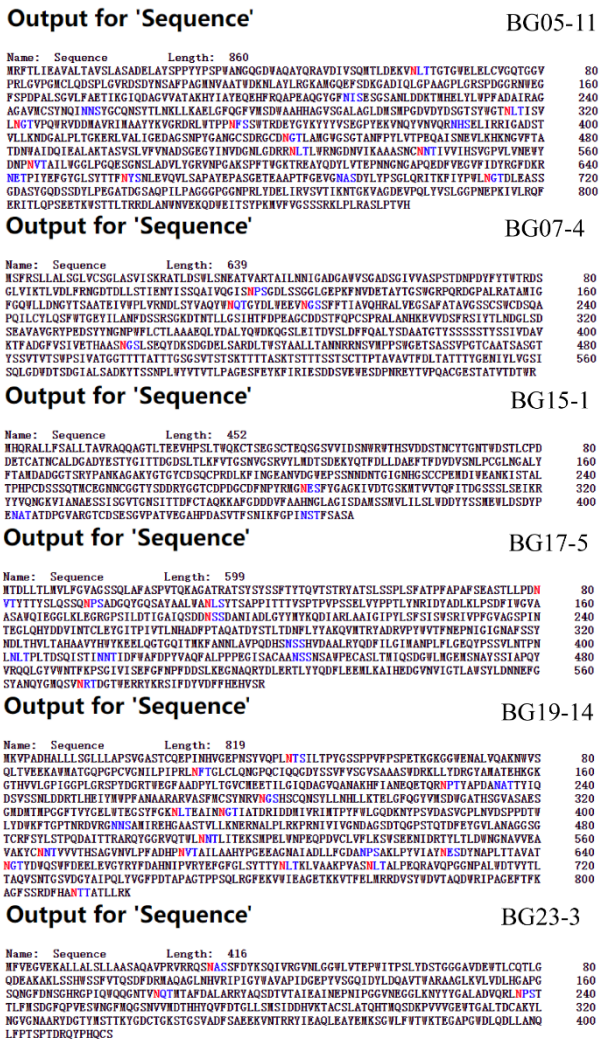
Check to include this hit in error tolerant search

Query	Observed	Mr(expt)	Mr(calc)	ppm	Miss	Score	Expect	Rank	Unique	Peptide
7	1023.5416	1022.5343	1022.5145	19.4	0	48	0.016	1	U	K.FVTGSNVGSR.L

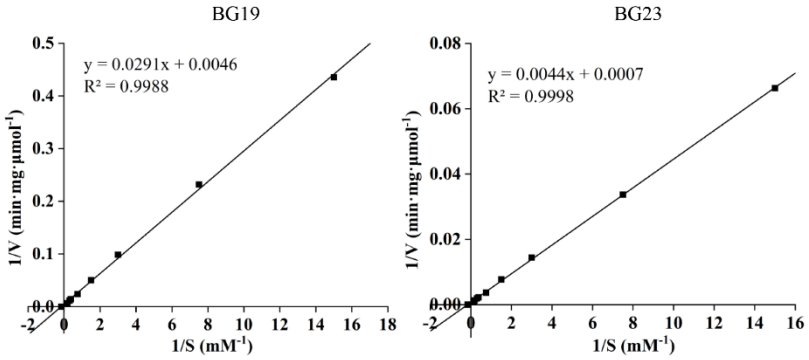
Supplementary Fig. 4 Raw data of peptide fingerprints mass associated with glycoside hydrolases in the protein bands of the enzyme solutions 706 and 803. Only the protein bands of 706-130 kDa, 803-110 kDa, and 803-65 kDa were associated with glycoside hydrolases.



Supplementary Fig. 5 Raw data of peptide fingerprints mass associated with glycoside hydrolases in the protein bands of the enzyme solutions 706 and 803. Only the protein bands of 706-130 kDa, 803-110 kDa, and 803-65 kDa were associated with glycoside hydrolases.

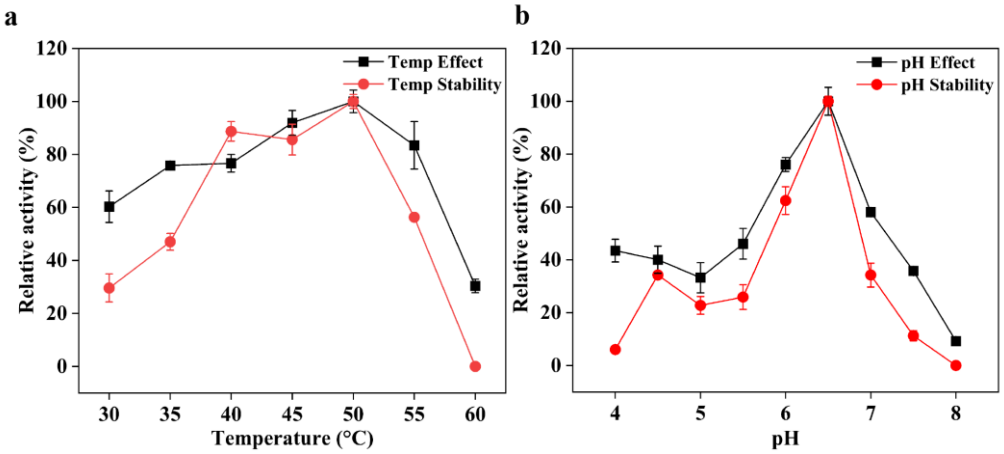


Supplementary Fig. 6 Protein N-glycosylation was predicted using NetNGlyc 1.0 server (https://services.healthtech.dtu.dk/service.php?NetNGlyc-1.0). Named by the protein name-N-glycosylation number. BG05-11 represents 11 N-glycosylation sites in BG05.

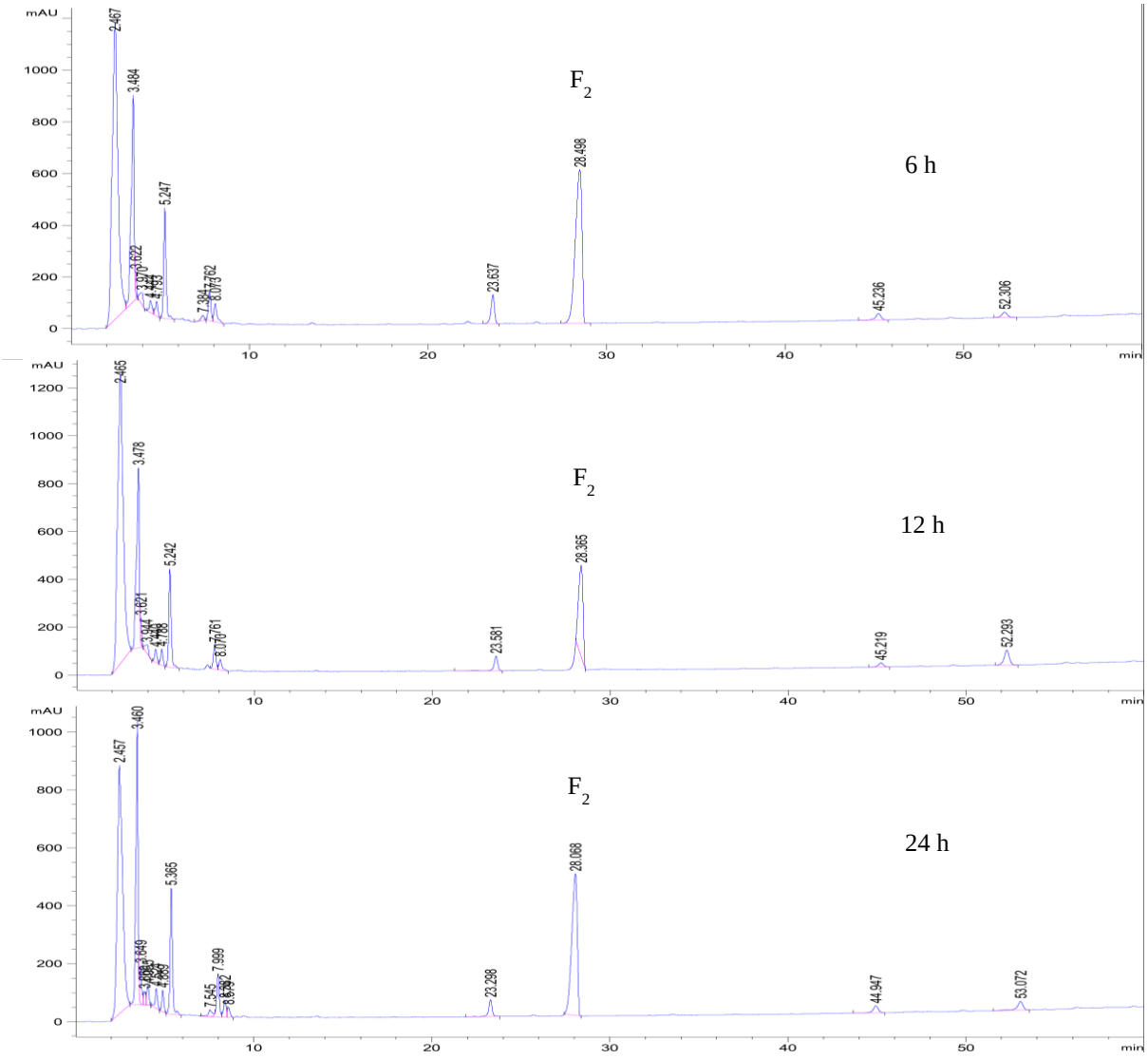


Protein	V_{max} (μmol·min^{-1}·mg^{-1})	K_m (mM)
BG19	217.39	6.33
BG23	1428.57	6.29

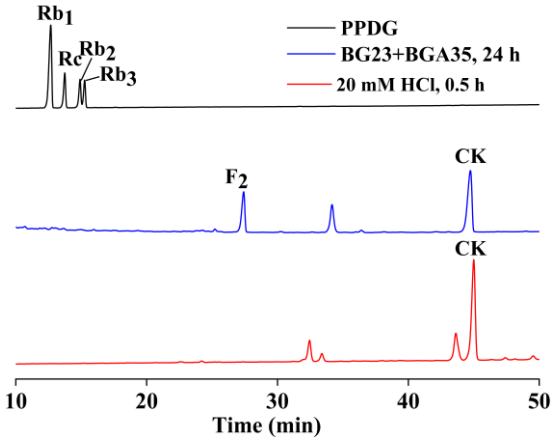
Supplementary Fig. 7 The enzyme kinetic parameters of purified BG19 and BG23.



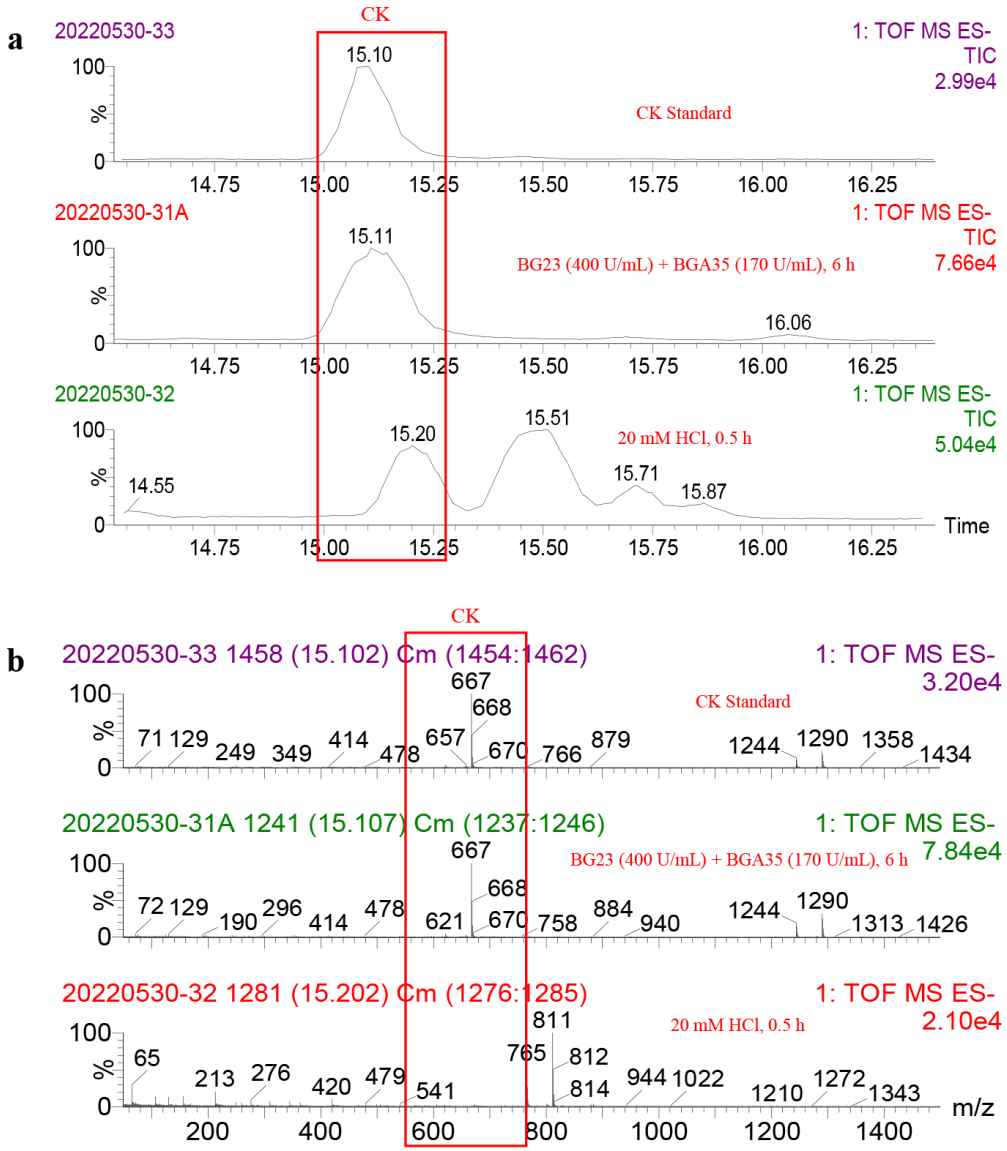
Supplementary Fig. 8 Effects of temperature and pH on the activity and stability of the BGA35.



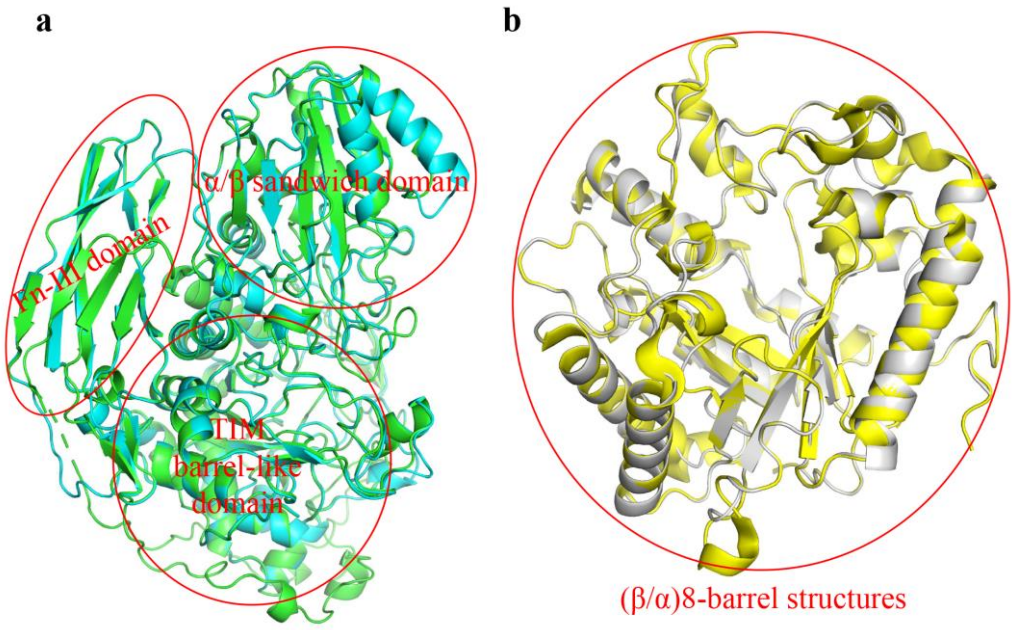
Supplementary Fig. 9 HPLC analysis of the samples of the graded combinatorial enzyme catalytic strategy were sampled at 6 h, 12 h, and 24 h of the second stage.



Supplementary Fig. 10 HPLC analysis of samples prepared by PPDG, one-pot combinatorial enzyme-catalyzed strategy, and acid hydrolysis process. PPDG contains Rb_1 with a final concentration of 2 mg/mL, and Rc, Rb_2 and Rb_3 with 1 mg/mL. BG23+BGA35, 6 h represents sample prepared by 400 U/mL BG23 combined with 170 U/mL BGA35 reacted with PPDG in 20 mM, pH 5.5 acetate buffer at 50 °C for 6 h. 20 mM HCl, 0.5 h represents a sample of PPDG hydrolyzed in 20 mM HCl for 0.5 h.



Supplementary Fig. 11 UPLC-Q-TOF-MS/MS analysis of samples prepared by CK standard, a one-pot combinatorial enzyme-catalyzed strategy, and acid hydrolysis process. a Ion flow mass spectra of samples prepared by CK standard, a one-step combination of enzyme-catalyzed strategy and acid hydrolysis process. **b** Mass spectra of ion fragments of samples prepared by CK standard, one-step combined enzyme-catalyzed strategy, and acid hydrolysis process at 15-15.25 min.

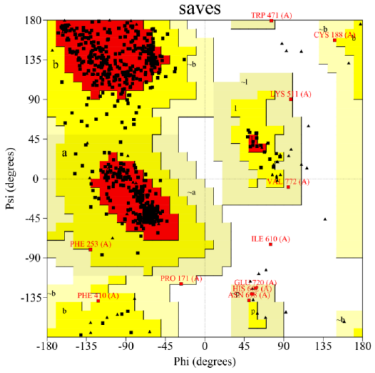


Supplementary Fig. 12 Structural matching of the SWISS-MODEL predicted protein model to the protein template. a Structural matching of the BG19 model with template 4iib. **b** Structural matching of the BG23 model with template 2pc8.

a

Protein	GMQE	QMEAN	Verify3D	ERRAT
BG19	0.72	0.75±0.05	87.20%	88.21%
BG23	0.79	0.81±0.05	97.39%	92.51%

b

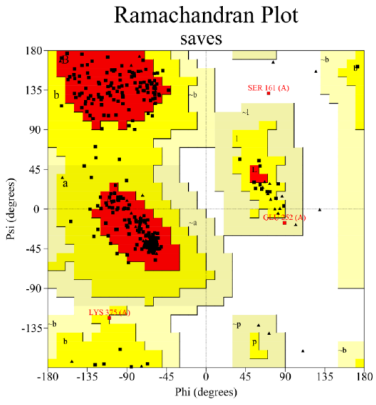


Plot statistics

Residues in most favoured regions [A,B,L]	546	86.9%
Residues in additional allowed regions [a,b,l,p]	72	11.5%
Residues in generously allowed regions [~a,~b,~l,~p]	9	1.4%
Residues in disallowed regions	1	0.2%
Number of non-glycine and non-proline residues	628	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	66	
Number of proline residues	46	
Total number of residues	742	

BG19

c

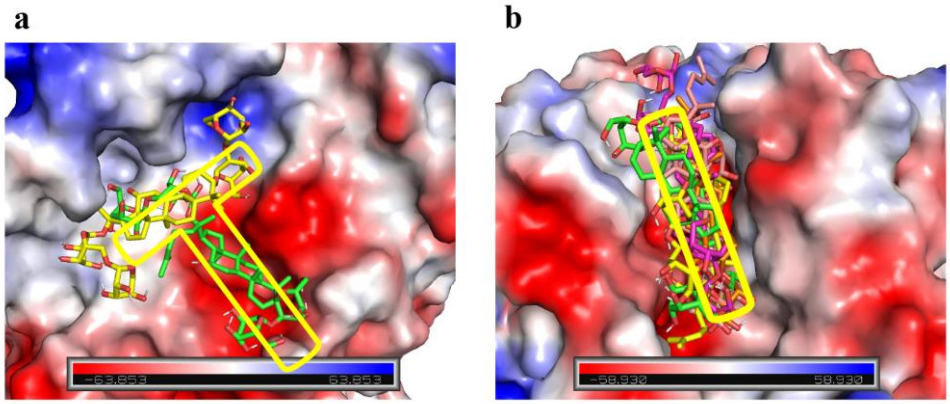


Plot statistics

Residues in most favoured regions [A,B,L]	284	86.9%
Residues in additional allowed regions [a,b,l,p]	40	12.2%
Residues in generously allowed regions [~a,~b,~l,~p]	2	0.6%
Residues in disallowed regions	1	0.3%
Number of non-glycine and non-proline residues	327	100.0%
Number of end-residues (excl. Gly and Pro)	2	
Number of glycine residues (shown as triangles)	38	
Number of proline residues	16	
Total number of residues	383	

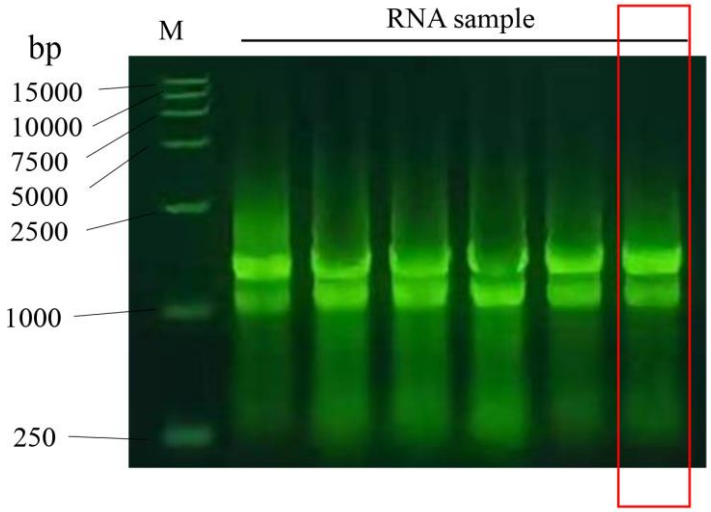
BG23

Supplementary Fig. 13 Quality assessment of the Swiss-Model predictive protein model. **a** Evaluation of the BG19 and BG23 models using SAVES v6.0 (https://saves.mbi.ucla.edu/). GMQE: 0-1, the larger the value the better the quality. QMEAN: 4-0, the closer to 0 the better the match. Verify3D > 80% is good. ERRAT around 95% or higher is good. **b** Ramachandran plot of the BG19 model. **c** Ramachandran plot of the BG23 model.



BG19 BG23

Supplementary Fig. 14 Surface electrostatic potential plots of BG19 and BG23 bound to substrates. a Surface electrostatic potential plot at the active site of BG19 receptor. **b** Surface electrostatic potential plot at the active site of BG23 receptor.



Supplementary Fig. 15 Nucleic acid electropherogram of RNA extracted from JE0609.

Supplementary Table 1. Changes in the content of ginsenosides in unfermented ginseng extracts and ginseng extracts catalyzed by JE0609 extracellular enzyme.

Concentration（mg/mL）	Ro	Re	Rg_1	Rb_1	Rc	Rb_2	Rb_3	Rd	F_2	Rg_3	CK
UGE*	12.66	18.97	0	0.96	4.42	2.04	3.83	12.85	3.00	0.43	0
JE0609†	8.85	0	12.34	0	0	1.37	0	7.37	1.71	0.35	5.22‡

*UGE is 10 mg/mL of uncatalyzed ginseng extract. †JE0609 is a sample of JE0609 extracellular enzyme reacted with 10 mg/mL of ginseng extract in 20 mM, pH 5.5 acetate buffer for 24 h at 45℃. ‡The molar conversion rate formula for CK, $\text{Molar conversion rate}\ (\%) = \frac{m_{CK}/M_{CK}}{m_{Rb1}/M_{Rb1}+m_{Rc}/M_{Rc}+m_{Rb2}/M_{Rb2}+m_{Rb3}/M_{Rb3}+m_{Rd}/M_{Rd}+m_{F2}/M_{F2}}$. The molar conversion rate of CK was 79.6%.

Supplementary Table 2. The features of raw sequencing data and improved one of JE0609.

Type	Raw sequencing data	Improved raw sequencing data
Total Reads Count (bp)	71272506	71272506
Total Bases (bp)	10690875900	10536322811
Average Read Length (bp)	150.00	147.83
Q20 (%)	98.63	98.68
Q30 (%)	95.92	96.01
GC (%)	53.19	53.29
N (ppm)	5.80	5.84

Supplementary Table 3. Information of assembled transcriptome sequence of JE0609.

Type	Contig	Unigene
Total sequences	4507423	73045
Total bases (bp)	216176485	68852200
Min length (bp)	25	201
Max length (bp)	16571	14051
Average length (bp)	47.96	942.60
N50	45	1890
(A+T)%	49.28	50.72
(G+C)%	50.71	49.29

Supplementary Table 4. The proportion and number of genes of different lengths in the predicted Unigene sequence.

Predicted gene	Number	Percent (%)
Total gene number	73045	100
Length < 200 bp	0	0.00
Length [200-500) bp	40534	55.49
Length [500-1000) bp	11400	15.61
Length [1000-1500) bp	6363	8.71
Length [1500-2000) bp	4808	6.58
Length > = 2000 bp	9939	13.61

Supplementary Table 5. The unigene was analyzed and annotated according to five databases.

Database	Number of Unigenes
Annotated in Nr	52085
Annotated in COG	29644
Annotated in Swissprot	37744
Annotated in KEGG	13597
Total annotated Unigenes	58026
Total Unigenes	73045

Supplementary Table 6. Transcriptome analysis of glycoside hydrolases of the GH1 and GH3 family in JE0609.

Name*	Gene_ID	Gene_Length	Sample_expected_count	Sample_FPKM	Gene Symbol	Geanbank	Domains	Definition
BG05	DN487_c0_g1_i1	2744	3850	47.86	AtWU_00271	GFN10476	BglX super family, Glyco_hydro_3_C, Fn4-like	Beta-glucosidase
-	-	-	-	-	AtWU_02072	GFN12275	BglX super family, Glyco_hydro_3_C, Fn3-like	Beta-glucosidase
-	-	-	-	-	AtWU_06875	GFN17073	BglX super family, Glyco_hydro_3_C, Fn3-like	Beta-glucosidase
BG17	DN6391_c0_g1_i1	1696	242	5.17	AtWU_02829	GFN13031	Glyco_hydro super family	Beta-glucosidase
BG19	DN6410_c0_g1_i1	565	14	1.34	AtWU_02287	GFN12490	PRK15098 super family-BglX, Glyco_hydro_3_C	Beta-glucosidase
BG23	DN22453_c0_g1_i1	3516	454	4.31	AtWU_00223	GFN10428	BglC-GH1	Exo-1,3-beta-glucanase
-	-	-	-	-	AtWU_08086	GFN18283	BglC-GH1, Glyco_hydro super family	Cellulase
-	-	-	-	-	AtWU_06990	GFN17188	BglC-GH1	Exo-1,3-beta-glucanase

*Analysis of glycoside hydrolases annotated as GH1 and GH3 family in the JE0609 transcriptome.

Supplementary Table 7. Analysis of the transcriptome and structural domains of the three hydrolases detected by PMF.

Protein	Name	Gene_ID	Gene_ Length	Sample_e xpected_ count	Sample_ FPKM	Gene Symbol	Genbank	Domains	Definition
706-130 kDa	BG05	DN487_c0_g 1_i1	2744	3850	47.86	AtWU_0 0271	GFN10476	BglX super family, Glyco_hydr o_3_C, Fn3-like	β-glucosidase
803-110 kDa	BG07	DN40820_c0 _g1_i1	2424	1331	18.97	AtWU_1 0364	GFN13031	Glyco_hydr o_15, CBM20_glu coamylase	Glucoamylase
803-65 kDa	BG15	DN4493_c0_ g2_i1	2545	163	2.2	AtWU_0 7393	GFN17591	Glyco_hydr o_7	1,4-β-D-glucan cellobiohydro lase

144 **Supplementary Table 8.** Physicochemical properties of the proteins predicted by

145 Expasy (https://web.expasy.org/protparam/).

Name	Number of amino acids (aa)	Molecular weight (kDa)	Theoretical *pI*	Instability index	Aliphatic index	Grand average of hydropathicity
BG05	860	93.36	4.70	31.31	75.56	-0.356
BG07	639	68.40	4.20	35.37	72.10	-0.193
BG15	452	48.30	4.16	29.71	57.01	-0.408
BG17	599	66.32	4.51	43.40	77.91	-0.185
BG19	819	89.53	5.06	32.60	74.10	-0.310
BG23	416	45.53	4.97	32.74	71.54	-0.344

146

Supplementary Table 9. Enzyme activity assay of 250 mL shake flask fermentation supernatant of recombinant proteins.

Protein	GS115-pPIC9K (pNPG)*	GS115-pPIC9K (DNS)†	GS115-pPIC9K-*BG05* (pNPG)*	GS115-pPIC9K-*BG07* (DNS)†	GS115-pPIC9K-*BG15* (DNS)†	GS115-pPIC9K-*BG17* (pNPG)*	GS115-pPIC9K-*BG19* (pNPG)*	GS115-pPIC9K-*BG23* (pNPG)*
Enzyme activity (U/mL)	0.23 ± 0.02	0.27 ± 0.05	0.28 ± 0.02	3.00 ± 0.21	0.26 ± 0.04	0.22 ± 0.09	59.41 ± 8.19	235.73 ± 30.07
Protein concentration (mg/mL)	1.04 ± 0.18	1.04 ± 0.18	1.13 ± 0.04	1.44 ± 0.12	1.32 ± 0.12	1.33 ± 0.09	1.31 ± 0.05	1.15 ± 0.16
Specific activity (U/mg)‡	0.23 ± 0.03	0.26 ± 0.05	0.25 ± 0.02	2.16 ± 0.05	0.20 ± 0.05	0.16 ± 0.05	45.35 ± 6.62	2354.71 ± 382.87

*The enzymatic activities of the transformants GS115-pPIC9K, GS115-pPIC9K-*BG05*, GS115-pPIC9K-*BG17*, GS115-pPIC9K-*BG19*, and GS115-pPIC9K-*BG23* were determined according to the pNPG method. †The recombinants GS115-pPIC9K, GS115-pPIC9K-*BG07* and GS115-pPIC9K-*BG15* were determined by the DNS. ‡Ratio of enzyme activity to protein concentration.

155 **Supplementary Table 10.** Summary of papers on PPD-type ginsenoside hydrolases.

Protein	GH family	Strain	pH	Tm (°C)	Time (h)	Enzyme concentration	Substrates	Yield	Transformation	Reference
β-glycosidase	3	*Microbacterium esteraromaticum*	7.0	40	12	0.1 mg/mL	0.74 mg/mL Rb_2	0.27 mg/mL CY, and 0.1 mg/mL CK	Rb_2 → CY → CK	[1]
β-glucosidase	3	*Microbacterium esteraromaticum*	7.0	40	1	0.1 mg/mL	1 mg/mL Rb_1	0.46 mg/mL CK, 77%	Rb_1 → Rd → CK	[2, 3]
β-glycosidase	2	*Microbacterium esteraromaticum* GS514	7.0	40	2	0.1 mg/mL	1 mg/mL Rb_2	0.47 mg/mL Rg_3	Rb_2 → Rd → Rg_3	[4]
β-glucosidase	3	*Microbacterium esteraromaticum*	7.0	37	1	0.75 U/mL	10 mg/mL Rb_1	6.9 mg/mL Rg3, 97.8%	Rb_1 → Rd → Rg_3	[5]
β-glycosidase		*Sulfolobus solfataricus*	5.5	85	12	40 U/mL	1.9 mg/mL Rb_1, 0.52 mg/mL Rb_2, 0.92 mg/mL Rc, and 0.23 mg/mL Rd	1.63 mg/mL	Rb_1 or Rb_2 → Rd → F_2 → CK, and Rc → CMc → CK	[6]
β-glycosidase		*Sulfolobus*	5.5	95	1.5	0.01 mg/mL	2 mg/mL Rb_1	1.12 mg/mL CK,	Rb_1 → Rd → F_2 → CK	[7]

		solfataricus						100%，746.6 mg/L/h		
β-glucosidase		*Paenibacillus mucilaginosus*	7.0	30	24	0.003 U/mL	3.4 mM Rb_1, and 2.4 mM Rd	70% CK	$Rb_1 \rightarrow Rd \rightarrow Rg_3 \rightarrow CK$	[8]
β-glucosidase		*Flavobacterium johnsoniae*	7.0	30	24	0.001 U/mL	3.4 mM Rb_1, and 2.4 mM Rd	61% Rg_3	$Rb_1 \rightarrow Rd \rightarrow Rg_3$	[8]
β-glucosidase	3	*Flavobacterium chilense*	7.0	37	48	0.4 mg/mL	Rb_1, Rb_2, Rb_3, Rc, Rd, F_2, Rg_3, Re, Rg_1, Gyp XVII, and Gyp LXXV		$Rb_1 \rightarrow Rd \rightarrow Rg_3 \rightarrow Rh_2$, Gyp XVII → Gyp LXXV → CK, Gyp LXXV → CK, $F_2 \rightarrow$ CK, $Rb_2 \rightarrow$ CO → CY, $Rb_3 \rightarrow$ CMx1 → CMx, Rc → CMc1→ CMc	[9]
β-glucosidase	3	*Bifidobacterium breve* ATCC 15700	5.0	35	12	6 U/mL	5 mM Rd	4.35 mM CK, 100%	$Rd \rightarrow F_2 \rightarrow$ CK	[10]
β-glucosidase	3	*Lactobacillus brevis*	7.0	30	24	0.1 mg/mL	1 mg/mL Rb_1, 1 mg/mL F_2	0.59 mg/mL Rd, 69% 0.72 mg/mL CK, 91% 30 mg/L/h	$Rb_1 \rightarrow Rd$, $F_2 \rightarrow$ CK	[11]
β-glycosidase	1	*Thermotoga petrophilia*	5.0	90	3	1.2 U/mL	10 mg/mL ginsenoside extract	3.93 mg/mL Rg_3, 98.19%	Rb_1 or $Rb_2 \rightarrow Rd \rightarrow Rg_3$	[12]

β-glucosidase		*Thermotoga petrophila* RKU-1	5.0	95	1.5	10 U/mL	10 mg/mL Rb_1	6.93 mg/mL Rg_3, 97.9%	Rb_1 → Rd → Rg_3	[13]
a-L-arabinofuranosidase	51	*Thermotoga thermarum* DSM5069	5.0	85	1	12 U/mL	25 mg/mL Rc	21·8 mg/mL Rd, 99.4%	Rc → Rd	[14]
β-glucosidase	3	*Corynebacterium glutamicum* ATCC13032	7.0	37	24	75 mg/mL	20 mg/mL PPDGM	7.86 mg/mL CK, 79.2%, 327.5 mg/L/h	Rb_1 → Gyp XVII → Gyp LXXV, F_2 → CK → PPD, Rd → F_2 → CK → PPD	[15]
β-glycosidase	3	*Pyrococcus furiosus*	5.0	95	1.2	60 U/mL	3.9 mM Rb_1, Rb_2, Rc, and Rd	3.1 mM CK, 79.5%	Rb_1 or Rb_2 → Rd → F_2 → CK, and Rc → CMc → CK	[16]
β-glucosidase	3	*Actinosynnema mirum* KACC 20028T	8.0	30		1 mg/mL	0.1% (w/v) Rb_1, Rb_2, Rc, Rd, Rg_3		Rb_1→ Gyp XVII → Gyp LXXV, Rh_2 → PPD, Rb_2 → CO → CY, Rc → CMc 1→ C-Mc, Rd → F_2 → Rh_2 → PPD, Rg_3 → Rh_2 → PPD	[17]
β-glucosidase		*Caldicellulosiruptor bescii*	5.5	80	2	10.2 U/mL	1 mM Rb_1, Rb_2, Rc	100% CK	Rb_1, Rb_2, Rc → Rd → F_2 → CK	[18]
α-L-arabinofuranosidase		*Leuconostoc* sp. 22-3	6.0	30	0.35	10 μg/mL	0.1% Rc	100% Rd	Rc → Rd	[19]
β-glucosidase	3	*Shingopyxis*	5.5	40	1	0.5 mg/mL	8 mg/mL Rb_1	6.8 mg/mL Gyp XVII	Rb_1 → Gyp XVII	[20]

		alaskensis								
β-glucosidase	1	*Cellulosimicrobium cellulans* sp.	5.5	35	2	1 mg/mL	1g PPDGM	292 mg Gyp XVII, 134 mg CO, 184 mg Mb, and 62 mg F_2	$Rb_1 \rightarrow$ Gyp XVII, $Rb_2 \rightarrow$ CO, Rc $\rightarrow$ Mb, Rd $\rightarrow F_2$	[21]
β-glucosidase	1	*Bacillus* sp. 3 KP	6.0	37	48	100 µL of the cell lysate	0.1% (w/v) Rb_1		$Rb_1 \rightarrow$ Gyp XVII $\rightarrow F_2$	[22]
cellulase		*Trichoderma viride*	4.5	37	72	0.8 g	0.4 g of PPD-type saponin		Rb_1, Rb_2, Rb_3 $\rightarrow$ Rd	[23]
lactase	35	*Aspergillus oryzae*	4.5	37	72	2 g	0.4 g of PPD-type saponin		Rb_1, Rb_2, Rb_3 $\rightarrow$ Rd $\rightarrow$ Rg_3, Rb_1, Rb_2, Rb_3 $\rightarrow$ Rd $\rightarrow F_2 \rightarrow$ CK	[23]
β-galactosidase	35	*Aspergillus oryzae*	4.5	37	72	0.8 g	0.4 g of PPD-type saponin		Rb_1, Rc $\rightarrow$ Rd $\rightarrow F_2 \rightarrow$ CK, Rb_2, Rc $\rightarrow$ CV $\rightarrow$ CVII $\rightarrow$ CK	[23]

156

Supplementary Table 11. The lowest binding energy of molecular docking of BG19 and BG23 with ginsenosides.

Protein	Binding energy (kcal/mol)				
	Rb_1	Rc	Rb_2	Rb_3	F_2
BG19	-4.39	-	-	-5.31	-
BG23	-3.65	-4.91	-3.76	-5.02	-6.47

Supplementary Table 12. CK yield of strains with β-glucosidase producing ability.

Strains*	JE0609	JE0512	HW0114	HWGS0607	YN191213
CK content (mg/mL)	5.22	0.83	0.63	0.60	0.39

*Only 5 of the 42 strains had CK production capacity.

Supplementary Table 13. List of primers of the β-tubulin genes of different *Aspergillus* sp. in NCBI.

Organism	GenBank	Base Number (bp)	Primer
Aspergillus niger	FJ629297.1	418	F 5’ - TGAGTTCAGATGTTGTCCATTAGGTAC - 3’ R 5’ - GAAGACGAAGTTGTCGGG - 3’
Aspergillus piperis	KJ469441.1	524	F 5’ - TGCTTTCTGGTACGTATTCAC - 3’ R 5’ - CCTTGGCCCAGTTGTTAC - 3’
Aspergillus costaricensis	LC516862.1	507	F 5’ - TGGTACGTATTCACTGCCAC - 3’ R 5’ - CCAGTTGTTACCAGCACC - 3’
Aspergillus tubingensis	GFN19051.1	1334	F 5’ - ATGCGTGAGATCGTTCACC - 3’ R 5’ - TTACTCCTCGGCCTCAAG - 3’
Aspergillus welwitschiae	MT293604.1	461	F 5’ - TGAGTTCAGATGTTGTCCATTAGG - 3’ R 5’ - TGTAGTGACCCTTGGCCCAG - 3’

Supplementary Table 14. List of primers of the screened genes.

Gean	Base Number (bp)	Primer*
BG05	2433	F 5' - GCTGAAGCTTACGTAGAATTCGTCTCGCAAATGACATTGGATGAG - 3' † R 5' - AAGGCGAATTAATTCGCGGCCGCTTAATGATGATGATGATGATGGTGAACAG TAGGCAGAGACGCC - 3' †
BG07	1863	F 5' - GCTGAAGCTTACGTAGAATTCGTTATTTCCAAGCGCGCGA - 3' † R 5' - AAGGCGAATTAATTCGCGGCCGCGTATGGATTGTCTACCGCC - 3' †
BG15	1285	F 5' - GCTGAAGCTTACGTAGAATTCGGAAGTCCATCCTTCCTTG - 3' † R 5' - AAGGCGAATTAATTCGCGGCCGCGCTTCAGCTTATGCGGAAG - 3' †
BG17	1734	F 5' - AGAGAGGCTGAAGCTTACGTATCGCCTGTCACGCAGAAAG - 3' † R 5' - AAGGCGAATTAATTCGCGGCCGCACACCTACCTGCTAACATGCTCATG - 3' †
BG19	2394	F 5' - GCTGAAGCTTACGTAGAATTCAGCACTTGCCAGGAACCTATCAAC - 3' † R 5' - AAGGCGAATTAATTCGCGGCCGCTCAATGATGATGATGATGATGCTTCCTAA GCAAGGTAGCAGTAG - 3' †
BG23	1167	F 5' - GCTGAAGCTTACGTAGAATTCCAAAGCAACGCAAGCTCTTTCG - 3' † R 5' - AAGGCGAATTAATTCGCGGCCGCTTGACCGTTCAAGAACACTGGTGG - 3' †

*The sequences of the designed primers need to be stripped of the signal peptide sequences. †The underlined sequences represent homologous arm sequences required for homologous recombination.

References

1. Quan L-H, Jin Y, Wang C, Min J-W, Kim Y-J, Yang D-C. Enzymatic transformation of the major ginsenoside Rb_2 to minor compound Y and compound K by a ginsenoside-hydrolyzing β-glycosidase from *Microbacterium esteraromaticum*. *J Ind Microbiol Biotechnol* **39**, 1557-1562 (2012).
2. Quan L-H, Min J-W, Jin Y, Wang C, Kim Y-J, Yang D-C. Enzymatic Biotransformation of Ginsenoside Rb_1 to Compound K by Recombinant β-Glucosidase from *Microbacterium esteraromaticum*. *J Agric Food Chem* **60**, 3776-3781 (2012).
3. Quan L-H, Min J-W, Yang D-U, Kim Y-J, Yang D-C. Enzymatic biotransformation of ginsenoside Rb_1 to 20(S)-Rg_3 by recombinant *β*-glucosidase from *Microbacterium esteraromaticum*. *Appl Microbiol Biotechnol* **94**, 377-384 (2012).
4. Quan L-H, Wang C, Jin Y, Wang T-R, Kim Y-J, Yang DC. Isolation and characterization of novel ginsenoside-hydrolyzing glycosidase from *Microbacterium esteraromaticum* that transforms ginsenoside Rb_2 to rare ginsenoside 20(*S*)-Rg_3. *Antonie Van Leeuwenhoek* **104**, 129-137 (2013).
5. Pei J, *et al.* Enzymatic transformation of ginsenoside Rb1 to ginsenoside 20(S)-Rg3 by GH3 *β*-glucosidase from *Thermotoga thermarum* DSM 5069T. *J Mol Catal B: Enzym* **113**, 104-109 (2015).
6. Noh K-H, Son J-W, Kim H-J, Oh D-K. Ginsenoside Compound K Production from Ginseng Root Extract by a Thermostable β-Glycosidase from *Sulfolobus solfataricus*. *Biosci Biotechnol Biochem* **73**, 316-321 (2009).
7. Shin K-C, Choi H-Y, Seo M-J, Oh D-K. Improved conversion of ginsenoside Rb_1 to compound K by semi-rational design of *Sulfolobus solfataricus* β-glycosidase. *AMB Express* **7**, 186 (2017).
8. Li L, Lee SJ, Yuan QP, Im WT, Kim SC, Han NS. Production of bioactive ginsenoside Rg_3(*S*) and compound K using recombinant *Lactococcus lactis*. *J Ginseng Res* **42**, 412-418 (2018).
9. Siddiqi MZ, Medjebouri S, Liu Q, Park HY, Kim G-R, Im W-T. Efficient Production of Various Minor Ginsenosides from PPD- and PPT-type Major Ginsenosides Using a Single Recombinant BglFc Isolated from *Flavobacterium chilense*. *Biotechnol Bioprocess Eng* **26**, 232-246 (2021).
10. Zhang R, *et al.* Highly Selective Production of Compound K from Ginsenoside Rd by Hydrolyzing Glucose at C-3 Glycoside Using *β*-Glucosidase of *Bifidobacterium breve* ATCC 15700. *J Microbiol Biotechnol* **29**, 410-418 (2019).
11. Zhong F-L, *et al.* Cloning and Characterization of Ginsenoside-Hydrolyzing *β*-Glucosidase from *Lactobacillus brevis* That Transforms Ginsenosides Rb1 and F2 into Ginsenoside Rd and Compound K. *J Microbiol Biotechnol* **26**, 1661-1667 (2016).

12. Zhang S, Luo J, Xie J, Wang Z, Xiao W, Zhao L. Cooperated biotransformation of ginsenoside extracts into ginsenoside 20(*S*)-Rg_3 by three thermostable glycosidases. *J Appl Microbiol* **128**, 721-734 (2020).
13. Xie J, *et al.* Overexpression and characterization of a Ca^{2+} activated thermostable *β*-glucosidase with high ginsenoside Rb1 to ginsenoside 20(S)-Rg3 bioconversion productivity. *J Ind Microbiol Biotechnol* **42**, 839-850 (2015).
14. Xie J, *et al.* Characterization of a novel arabinose-tolerant *α*-L-arabinofuranosidase with high ginsenoside Rc to ginsenoside Rd bioconversion productivity. *J Appl Microbiol* **120**, 647-660 (2016).
15. Cui C-h, Jeon B-M, Fu Y, Im W-T, Kim S-C. High-density immobilization of a ginsenoside-transforming β-glucosidase for enhanced food-grade production of minor ginsenosides. *Appl Microbiol Biotechnol* **103**, 7003-7015 (2019).
16. Yoo M-H, Yeom S-J, Park C-S, Lee K-W, Oh D-K. Production of aglycon protopanaxadiol via compound K by a thermostable *β*-glycosidase from *Pyrococcus furiosus*. *Appl Microbiol Biotechnol* **89**, 1019-1028 (2011).
17. Cui C-H, Kim S-C, Im W-T. Characterization of the ginsenoside-transforming recombinant *β*-glucosidase from *Actinosynnema mirum* and bioconversion of major ginsenosides into minor ginsenosides. *Appl Microbiol Biotechnol* **97**, 649-659 (2013).
18. Shin K-C, Kim T-H, Choi J-H, Oh D-K. Complete Biotransformation of Protopanaxadiol-Type Ginsenosides to 20-O-*β*-Glucopyranosyl-20(S)-protopanaxadiol Using a Novel and Thermostable *β*-Glucosidase. *J Agric Food Chem* **66**, 2822-2829 (2018).
19. Liu Q-M, *et al.* Bioconversion of ginsenoside Rc into Rd by a novel *α*-L-arabinofuranosidase, Abf22-3 from *Leuconostoc* sp. 22-3: cloning, expression, and enzyme characterization. *Antonie Van Leeuwenhoek* **103**, 747-754 (2013).
20. Shin K-C, Oh D-K. Characterization of a novel recombinant *β*-glucosidase from *Sphingopyxis alaskensis* that specifically hydrolyzes the outer glucose at the C-3 position in protopanaxadiol-type ginsenosides. *J Biotechnol* **172**, 30-37 (2014).
21. Yuan Y, *et al.* Overexpression and characterization of a glycoside hydrolase family 1 enzyme from *Cellulosimicrobium cellulans* sp. 21 and its application for minor ginsenosides production. *J Mol Catal B: Enzym* **120**, 60-67 (2015).
22. Geraldi A, Ni'matuzahroh, Fatimah, Cui C-H, Nguyen TT, Kim SC. Enzymatic biotransformation of ginsenoside Rb_1 by recombinant β-glucosidase of bacterial isolates from Indonesia. *Biocatal Agric Biotechnol* **23**, 101449 (2020).
23. Ko S-R, Suzuki Y, Suzuki K, Choi K-J, Cho B-G. Marked Production of Ginsenosides Rd, F_2, Rg_3, and Compound K by Enzymatic Method. *Chem Pharm Bull* **55**, 1522-1527 (2007).