

Electronic Supplementary Information

Ep7GT, a glycosyltransferase with sugar donor flexibility from *Epimedium pseudowushanense*, catalyzes the 7-*O*-glycosylation of baohuoside

Keping Feng,^{a, b, c} Ridao Chen,^{b, c, d} Kebo Xie,^{b, c, d} Dawei Chen,^{b, c, d} Jimei Liu,^{b, c, d}
Wenyu Du,^{b, c, d} Lin Yang^{a, *} and Jungui Dai^{b, c, d, *}

^a *College of Life and Environmental Sciences, Minzu University of China, 27 Zhong Guan Cun Southern Street, Beijing 100081, China*

^b *State Key Laboratory of Bioactive Substance and Function of Natural Medicines; ^c CAMS Key Laboratory of Enzyme and Biocatalysis of Natural Drugs; ^d NHC Key Laboratory of Biosynthesis of Natural Products, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, 1 Xian Nong Tan Street, Beijing, 100050, China*

* Corresponding authors.

Tel: +86 10 68936927; fax: +86 10 68936927. E-mail: 15116995486@163.com (L. Y.).

Tel: +86 10 63165195; fax: +86 10 63017757. E-mail: jgdai@imm.ac.cn (J. D.).

Table of Contents

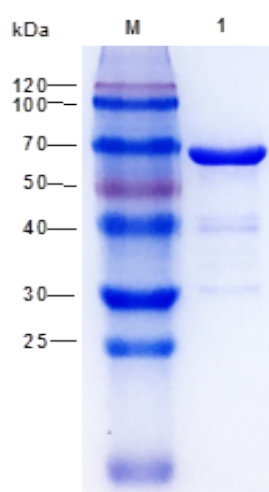
Table S1.	The HPLC method used in this study.....	S3
Table S2.	Details of the plasmids and strains used in this study.....	S3
Figure S1.	SDS-PAGE of recombinant His ₆ -Ep7GT purified by affinity chromatography.....	S3
Figure S2.	Enzymatic reactions catalyzed by crude extracts of recombinant EpGTs with baohuoside (1) as sugar acceptor and UDP-glucose as sugar donor.....	S4
Figure S3.	Multiple alignment of the amino acid sequences of EpGTs, AtUGT78D1 and EpPF3RT.....	S5
Figure S4.	Ep7GT-catalyzed glycosylation of 8-prenylkaempferol (3) with different sugar donors.....	S6
Figure S5.	Ep7GT-catalyzed glycosylation of anhydroicaritin (4) with different sugar donors.....	S7
Figure S6.	Ep7GT-catalyzed glycosylation of baohuoside II (5) with different sugar donors.....	S8
Figures S7–S8.	¹ H and ¹³ C NMR spectra of 1a	S9
Figures S9–S12.	¹ H, ¹³ C NMR, HSQC and HMBC spectra of 1b	S10-S11
Figures S13–S16.	¹ H, ¹³ C NMR, HSQC and HMBC spectra of 1c	S12-S13
Figures S17–S20.	¹ H, ¹³ C NMR, HSQC and HMBC spectra of 2a	S14-S15
Figures S21–S23.	¹ H, ¹³ C NMR and HMBC spectra of 2b	S16-S17
Figures S24–S27.	¹ H, ¹³ C NMR, HSQC and HMBC spectra of 2a	S17-S19
Figure S28.	The HRESIMS spectrum of 1b	S19
Figure S29.	The HRESIMS spectrum of 1c	S20
Figure S30.	The apparent K_m values of recombinant Ep7GT.....	S20
Table S3.	The apparent K_m , K_{cat} and K_{cat}/K_m values of recombinant Ep7GT for different sugar donors with 1 as acceptor.....	S20
Figure S31.	The linear regression models and the regression equations of the external standard method established for the quantitative analysis of bioconversion.....	S21
Figure S32.	Exploring the catalytic reversibility of Ep7GT when 1a and UDP were used as substrate	S21
Figure S33.	Time course assay of 1a with whole cell catalyst of engineered <i>E. coli</i>	S22

Table S1. The HPLC method used in this study

Time (min)	Solvent A (%) 0.1% formic acid	Solvent B (%) CH ₃ OH	flow rate (mL min ⁻¹)
0.00	90.0	10.0	1.0
10.00	50.0	50.0	1.0
30.00	0.0	100.0	1.0
45.00	0.0	100.0	1.0

Table S2. Details of the plasmids and strains used in this study

plasmids or <i>E. coli</i> strain	relevant properties
Plasmids	
pET-28a	pBR322 ori, Kan ^r
Primers	
pE-7GT forward	CAAATGGGTCGCGGATCCATGGGTTCAGAAACTCAT
pE-7GT reverse	GTGGTGGTGGTGCTCGAGTCAGTGTCCAATAGTTGCAGC
Strains	
S0	Transetta (DE3) harboring empty pET-28a
S1	Transetta (DE3) harboring pE-Ep7GT

**Figure S1.** SDS-PAGE of recombinant His₆-Ep7GT purified by affinity chromatography. Lane M: Protein Marker; Lane 1: His-tagged Ep7GT (predicted M.W., 54.1 kDa) purified on Ni Sepharose column chromatography.

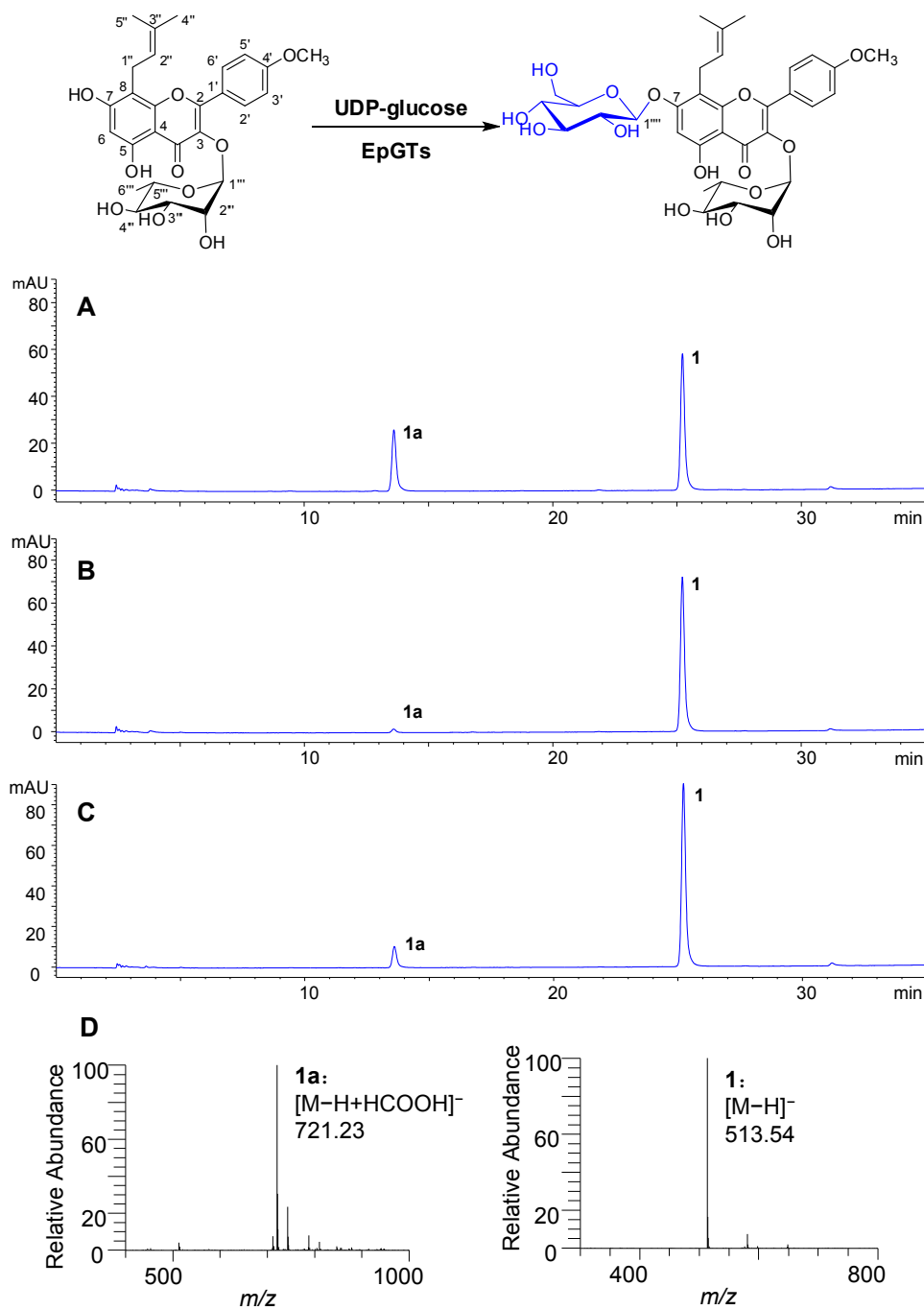


Figure S2. Enzymatic reactions catalyzed by crude extracts of recombinant EpGTs with baohuoside (**1**) as sugar acceptor and UDP-glucose as sugar donor. A) EpGT8 (Ep7GT); B) EpGT2; C) EpGT4; D) MS spectra of **1a** and **1** at negative mode.

EpGT8	MGSETHQLHALFFPFMAHGMI PM DI ARI FSVRG.....LKSTIIITTPQHATNI TSK..I CRYQNSGLDI KI I T	68
AtUGT73C6	..NAFEKNNPFPLHFVLPFMAQGMI PMVDI ARLLAQRG.....VLI TI VTPPHNAARFKNV..LNRAI ESGLPJ NLVQ	72
EpPF3RT	MSCI SVVKNGTSSI HI AMFWLPFGHVNPF I QLSNQLAANG.....YTI SELTTRDNL PKI KPHNHPDQI HI VPLDVKPP	76
EpGT2	NAKQTI VNYPSPTIGHLI SNVELGKVI TKYHPSFSVTI LI TTPPFNGMTI PQH..I YEI PKTHPSI NFHY	68
EpGT4	NAKTLVLVYPSPGMGLI SNVELGKVI KHHPSFSI TI LI AVPPYNTGSTAQY..I ARVKTTPSI TFHY	68
EpGT1	MGSI QAEKHLVCI PYPACGHNPFMOLSKLHAKG.....FHI TFWNTEFNHRRLLR...SKGPDALKGVSDFR	67
EpGT3	MEKKTI VLYPSPGMGLI SNVELGKLI TKRYPSFSI AI LI TTPPYTGSTAPY..I DRVSKTNPSI I FHY	68
EpGT5	NSKSSSTGPHI LVFPFPAQGMPLPLDLTHQLALRN.....LTI TI LVTPKNLSLSP.....LLSLYPSI QPLV	64
EpGT6	..NALFGEAKEEKI NVLLVAFSAQGMPLMLRLAKRLACKG.....LQVTLATTEI VRERMLKS...KNI DPSTKPTI TI AP	72
EpGT7	MEKKTI VLYPSPGMGLI SNVELGKLI TKRYPSFSI AI LI TTPPYTGSTAPY..I DRVSKTNPSI I FHY	68
EpGT9	..MADPNTNQSPHVALLESNGMLTFLRLAATLS TLN...CKI TFI TTHPTVSLAESQL..I SRFLLVTFPQI TPRQ	72
EpGT8	I PFPSEEFGLPKGCESVDSI PSRDMLFNFFDAVAKLQLEFERI LEEI HP..DCI VSDMFLPWTND..VACKHGTPI RI I FHG	145
AtUGT73C6	VKFPYQAEAGLEQGENVDLLTTVEQI TSFFKAVNLLKEPVQNL I EEMS PRPSCLI SDMCLSYTSE..I AKKFKI PKI L FHG	151
EpPF3RT	PFPFANGQPPSSAGAPPADSN PGLALLQLI LATDSLEDEVESQLALI KPDI I I YEF AHYI PA..I ANRLGI KSAFYCV	155
EpGT2	LPTFLPLSSSI PASNG...TGVLELL...LNNPHVHEALLTI STSSM P...AFVI DGFCSCLSE..I ANQLNI PAFYI T	139
EpGT4	LPVI PLISA..I PNINSANKETVAFGLLL...LSNPDVQKALVSI SATSTI P...AFI LDTFCSSSRE..LAEQNI PTYYFT	141
EpGT1	FETI PDGLPPSDKDQATQVPKLSDSVR..KNCLVPFVELVKKLNSSPDVKVTCI I SDGMVSGVQ..AAQKLGPEVQW	145
EpGT3	LPAL PFDS..I PTSNSSYRETI I FELLR...LSNPHVHEALLTI STYATI P...AFI VDFCFSSLE..LANQLNI PGYFLLT	141
EpGT5	LFPFNSSSLPFGVENAKDLSPI YFI PLI RVI GDLYTPLLQWFSHPSP...CAI LSDCTLGWTHH..LATQLGI PRI VTFSP	141
EpGT6	SI QI HLEYFS DGLPLDYDRKTS LDYDFD..CLGKFGPQNLSNLI GERFS...CI VNNPFVPWAD..VAAQRKTPCAMI WI	147
EpGT7	LPAL PFDS..I PTSNSSYRETI I FELLR...LSNPHVHEALLTI STYATI P...AFI VDFCFSSLE..LANQLNI PGYFLLT	141
EpGT9	FHLLLPDPTLSNSTD...FFLQFETI R...QSAHLLSPLLSSSSPPLSA...LI TDI TLASAFI PI TASLHLPNYLLFT	143
EpGT8	T..SFFSLCVANVTREYSVYE...SE.....GETFLIPLGLPD...EI KKKKSMPLSHLG..SKDRFGEMVDRI RDTE	207
AtUGT73C6	MG..CFCLLCVNRKRNREI LDNLKSD...KEYFI VYFPPD...RVEFTRPQVPEVETI..VPAGWKEI LEDVVEAD	217
EpPF3RT	TS..ATAVAYHVPACQPKSVD.....DLTHPPGHPPSKI SLQHF EAQQFMALAFARFGGLTTHERI TRAM	220
EpGT2	SGVSLSLLLHTPTI HNNTTI SFKDV.....NTLI DFPGLP...PMPSSDMI QPLDR..TSKVYEMF LNI CTSFP	205
EpGT4	SGVSLSLFLHII PTI HEKTTESFKDV.....NSLLVFPGLP...PI PSSDMI KPM LDR..TDKAYEWF I YCCQLI	207
EpGT1	ASACGFVGYLVNRELLKRGTLPLKDES YLTNGFLDMPI DWVPVGP...DI RLKDFPSFMRITDPNDI MLDFLGEEAQC	221
EpGT3	SGI SFLSI FLHLPTVDQNTLSFKDV.....NTLI DFPVGP...PI PSSDMFDPV LDR..TDKAYEWFVDCAKDFP	207
EpGT5	SGAFALQI I HSFLQDLPKDDNP.....DHLVSFPLKNS..PVYQWLQI PPLYHTYKPGDSNSEFI RDGFLAN	216
EpGT6	QPCALYAI YSRFENKLDPFHKFSAQLER.....EASLEVPFLP...LLRYDDLPSFI LPSNPFNRLTKLFTDMF TNM	217
EpGT7	SGI SFLSI FLHLPTVDQNTLSFKDV.....NTLI DFPVGP...PI PSSDMFDPV LDR..TDKAYEWFVDCAKDFP	207
EpGT9	SSARMLSVCASFPI NI SSI D.....EI QI PGSS...PVPKS WVPPLLNA..TNLFTTQFI ENGQKLV	202
EpGT8	VTSYGVLVNSFYELP PAYADHYRNVLGRRAW..HI GPVSLSNNNI I DKAQRGKKAGI DEHY..CUNWNSKEKDSVI YVSF	284
AtUGT73C6	KTSYGVY VNSFOELP AYAKDFKEARS GKAW..TI GPVSLCNKVGVDKAERGKNS DI DQDE...CLEWLDKPEKPSVLYVCL	294
EpPF3RT	SGS DLI I NKTSKEMSKYCNVI KEQY...QKPLVLAGLSLPE...PETDDLEDR..WESWLGQFAPESVLYVSF	286
EpGT2	KA..RGLI VNTFESLPRALKAI TDGLSAPNG..PTPPVYCI GPVI SP...SNQASKPA..CLDWDKQPRKSVFLCF	276
EpGT4	KS..QGI I VNSDSELPRAKAVRDGLSVFDA..PTPPVYCI GPLI ATDDRS GTEAGSGQAS...CLEWLDTPQSKSVFLCF	283
EpGT1	LNATAI I I NTFDDELOEVLDAI SSKFPETY...TI GPLSVLSNFPKSDI NSLRSNLWKEDTCLKWLDEKPEPSVI YVNY	299
EpGT3	KS..RGVI VNSFOELPRAKAI TDGLSVNPS..PTPPVYI GPLI ASEDRSGGEAGI GKVP...CLEWLDQPSKSVFLCF	283
EpGT5	SVS VGVVNTFEKLESVYLQHMVKS...CTNVWAVGPLL PAS...GSRSI SADDAQLVSWLDTCPDNSVYVI CF	275
EpGT6	DKLKWLVNSNEIENDVSVGLSPVRAI GPLVPSLLLDQDES D...DVGVDWMPDET CI KWLDEKPPSSVI YVSF	292
EpGT7	KS..RGVI VNSFOELPRAKAI TDGLSVNPS..PTPPVYI GPLI AS...EAGI GKVP...CLEWLDQPSKSVFLCF	277
EpGT9	EA..DGI LVNTFDSFEKESLVSLSEG VVKG...MPVVI TI GLVLPVG...DFERTDGGDS...GLGWLDEQPNSSVYVSF	272
EpGT8	GSVSRFSVCVQLTEI AHGLEASAVSI WVVR.....QLKDEEDNFLEFGFEERI HGRGLVI KDWAPVLI LD	350
AtUGT73C6	GSICNLPLS QLLELGLGLEESQRPI WVI RG.....WEKYKELVWFVESGFEEDRI QDRGLLI KGSPPNMI LS	363
EpPF3RT	GSQDVLSEKEI TELVLGLEESGVPNAVLK.....FPGDAPQEEI LPEGFTERVKGRGLI HSGVVRQQLLS	353
EpGT2	GSNGAFSAQWKEI AMGLEKSGORLEWVVR.....NQDSDLD DI LPPGLELTKERGMNI NSIAPREEMLN	343
EpGT4	GLGLFLSEAKLEI AI GLEKSGORLEWVVRSPPS..DEETKQLLAPPEPLDALLPQGLELTKERGFVVKSVAPVEVMS	362
EpGT1	GSVTVMSEKDI EFAGVGLANTKLPLFI WVR...PDVVMGDS AALPOEFMEFI TGRGRI SN..WCPQDQVMS	365
EpGT3	GSNGLFSGAQKEI AI GLEKSGORLEWVVRSPPTSDESSKGI AATPEPDLDALIPDGFLELTKERGLVVKSVAPVEVMS	363
EpGT5	GSQTVLNNKQNEQLALCI EOSGTREI VCAKEP...TVGQI S GDFGKVPDGF EQRLAGKGVVI KGSASVMI LK	345
EpGT6	GSITVLS TKQENLALGLEKSKHPLWVVKP...NEQPKDGEGLPSGLEI GDQGLVVE..WCPKI KVLM	359
EpGT7	GSNGLFSGAQKEI AI GLEKSGORLEWVVRSPPTSDESSKGI AATPEPDLDALIPDGFLELTKERGLVVKSVAPVEVMS	357
EpGT9	GSRTAMSREQI RELGDGLVRS GCKLEWVKD.....KKVDKEDEEGVEEVLGSELMDKVKELVVKSVVDGEKI LN	344
EpGT8	HPAI GGFVTEHCWNSI LE GVSAGVPI TWPLFAECPYNEALI TQVMNI GNVGVERVSDWTEQG...HVLVTKETVKKVYN	428
AtUGT73C6	HPSVGGFLTEHCWNSITLGI TAGLPMLTWPLFADGFCNEKLVQI LKVGVS AEVKEVMKAGEEIKI GVLVDKEGVKKA	443
EpPF3RT	HKS VGGYLSISGFLAEAVSSNCQLVLLPKMGDFQINARLMSRDLKI GVEVPRD...PVDG...KFTREDYCKAVK	424
EpGT2	KESVGGFLTEHCWNSILEAVI AGVPVVGWPLYAEQRNRVLLVEEVEALSVEE...RKDG...FVTAEEVEKKVR	413
EpGT4	NESLGGFVTEHCWNSITL EALLAGI PNVAWPLYAEQRNRVLLVEEI KVALSVEE...GVDG...FVTAQVEKKVR	432
EpGT1	HPSVGGFLTEHCWNSITL ESI SS GVPVVCWPFFAEQQTNCRYACNVGMI QMEI DN...NVKRDEVEKI VR	431
EpGT3	KEAVGGLTEHCWNSILEAVCAGVPPVVAWPLDAEQRNRVVI VEELKLALSVE...GEDR...FVTAEEVKKVK	433
EpGT5	HRAVGAFLTEHCWNSI YEAI VAGVPLTWAMNAEQFI DATI LV DHLGVAIVRCE...G...PRTVPDSGELAQ	412
EpGT6	HQAI TCFI TEHCWNSITL ETVTCTGVPVI AF PQWIDQPINAKLVEDVFRVGVRFQCD...QDG...VI SKEEVERCI S	429
EpGT7	KEAVGGLTEHCWNSILEAVCAGVPPVVAWPLDAEQRNRVVI VEELKLALSVE...GEDR...FVTAEEVKKVK	427
EpGT9	HPAVGGFLTEHCWNSITL EAAWHGVPVWVQGGDKN...ADVVDTSGLGKWARS...VDWGGQG...VVAEAI GVKVK	416
PSPG motif		
EpGT8	QLNAT..EEGEEI RNRARMLKGLARKAVKEDGSSFTDLRLI EELHHAATI G	478
AtUGT73C6	ELNGESDDAKERRRRRAKELGES AHKAVEEGGSSHSNI TFLLDQI MQLAQS	494
EpPF3RT	SLNVEVDG...EVGKEI RGNHAKLRDMLDKETQSGYLEQVLEELKLAGV	473
EpGT2	GLNES..EQGSVI RERVVKLSGEAKGAI TEAGSSVKALMNLVESWKL...	458
EpGT4	GLNES..EEGSVI RDRI VKLSREAKAAI SDGGSVRAVADLEELWER...	477
EpGT1	EVNEG..EEGKVRTKALDWKEKVF NACKEGGSSYKNFDRFVKDLVKLSN...	479
EpGT3	GLNES..EEGNLI KERVVNLSGEAKAAVVDKDGSSSEKSLSDLADLWKTGNRDS	483
EpGT5	AFVKSVDKSLERVRI LDLCKAAAGAI AN...SSIGDFVRALCEL...	454
EpGT6	VVDVG..PNSRELKRRALELKEAARKAVADGSSDRNI QI FI DEI SVDDS...	477
EpGT7	GLNES..EEGNLI KERVVNLSGEAKAAVVDKDGSSSEKSLSDLADLWKTGNRDS	477
EpGT9	ENMGD...EKLKAEARVREEARKAVGSGSSNGGLMKCLQVWKKGHHD	463

Figure S3. Multiple alignment of the amino acid sequences of EpGT8 (renamed Ep7GT), AtUGT78D1 (GenBank accession number NM_102790, from *Arabidopsis thaliana*) and EpPF3RT (GenBank accession number MG264429, from *E. pseudowushanense*) and other EpGTs.

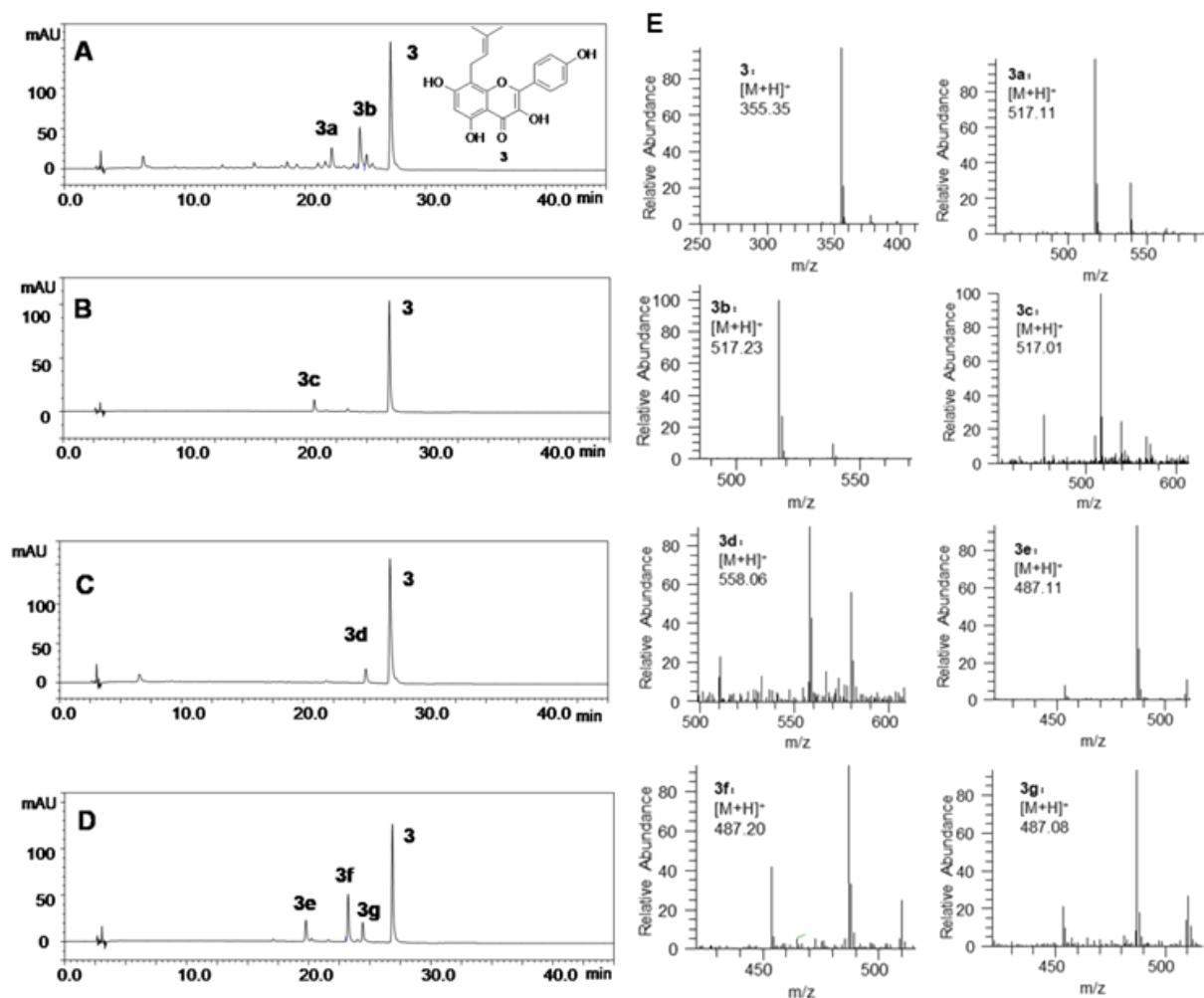


Figure S4. Ep7GT-catalyzed glycosylation of 8-prenylkaempferol (**3**) with different sugar donors. A) UDP-glucose; B) UDP-galactose; C) UDP-*N*-acetylglucosamine; D) UDP-xylose; E) MS spectra of **3** and the main products at positive mode.

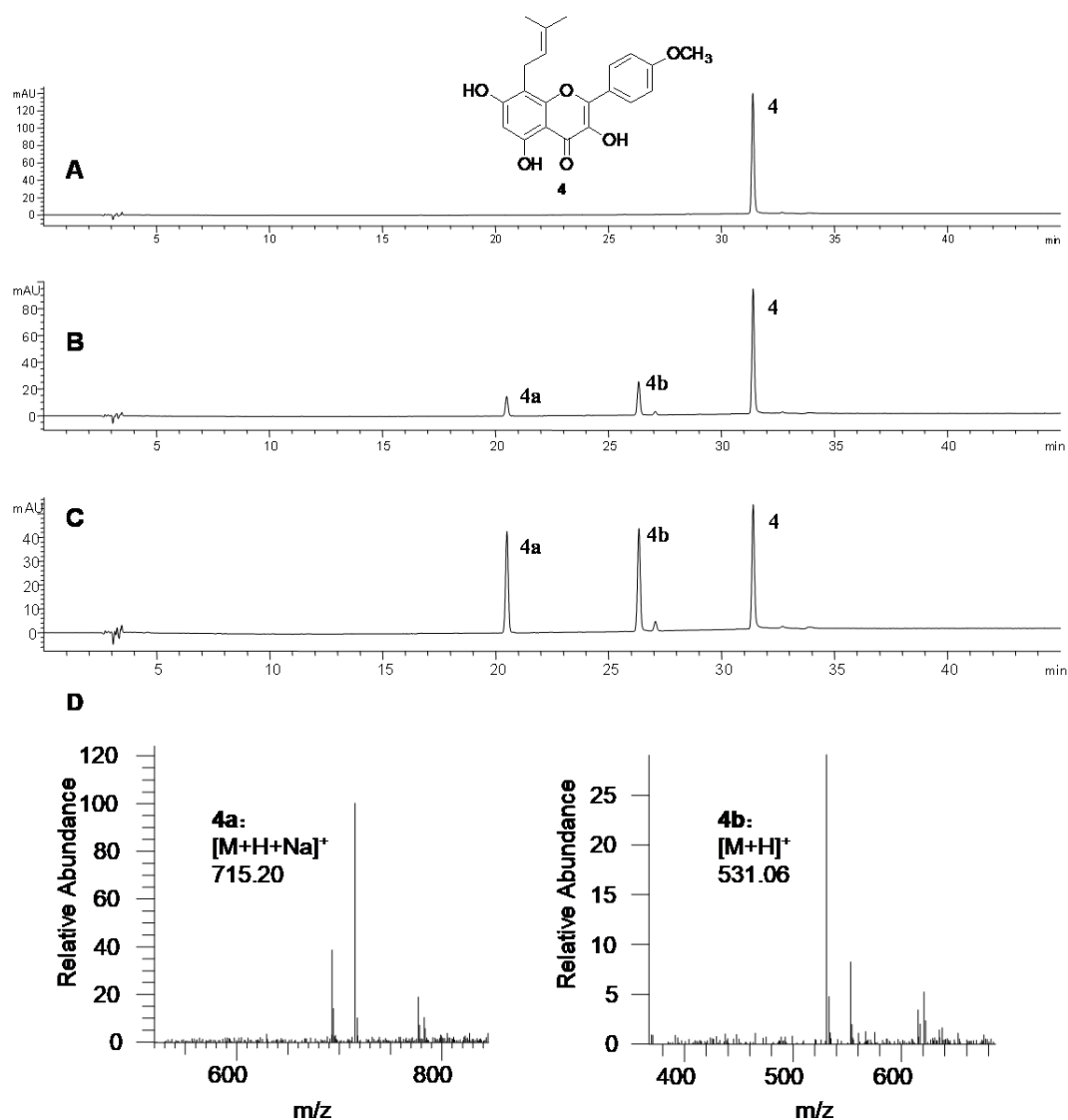


Figure S5. Ep7GT-catalyzed glycosylation of anhydroicaritin (**4**) with different sugar donors. A) Control group; B) UDP-glucose, C) TDP-glucose; D) MS spectra of the main products at positive mode.

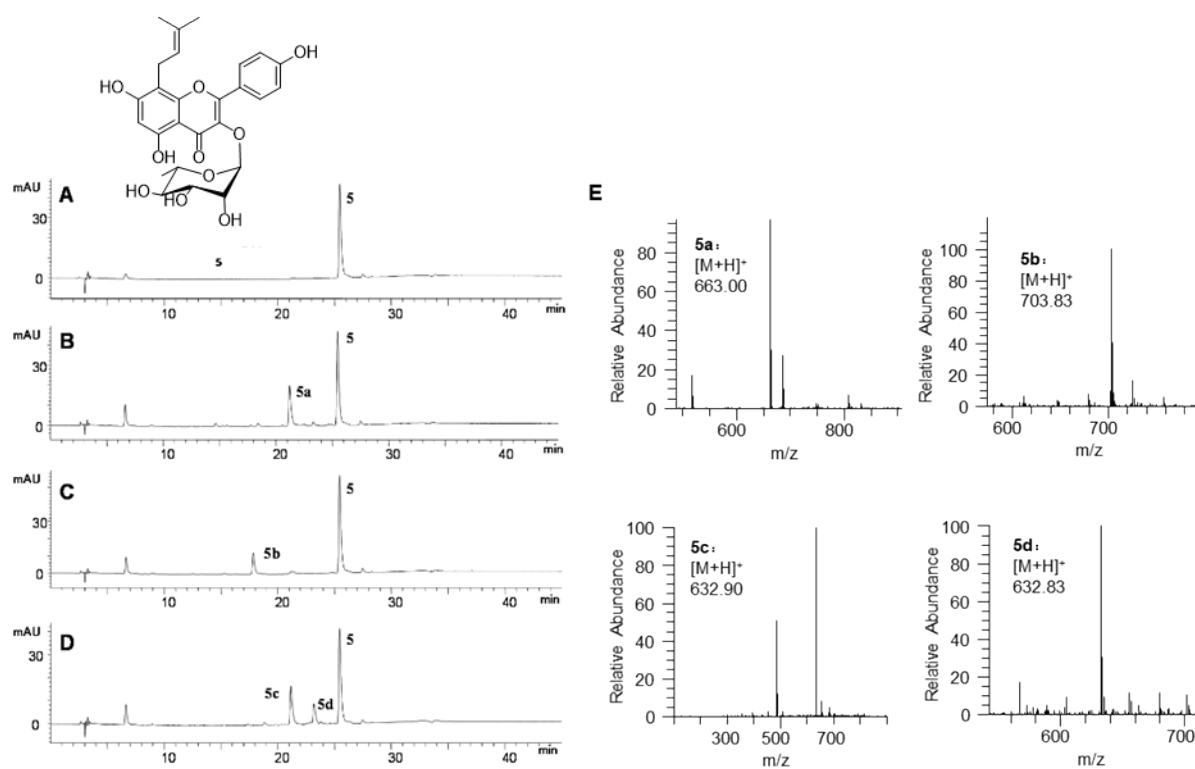


Figure S6. Ep7GT-catalyzed glycosylation of baohuoside II (**5**) with different sugar donors. A) Control group; B) UDP-glucose, C) UDP-*N*-acetylglucosamine; D) UDP-xylose; E) MS spectra of the main products at positive mode.

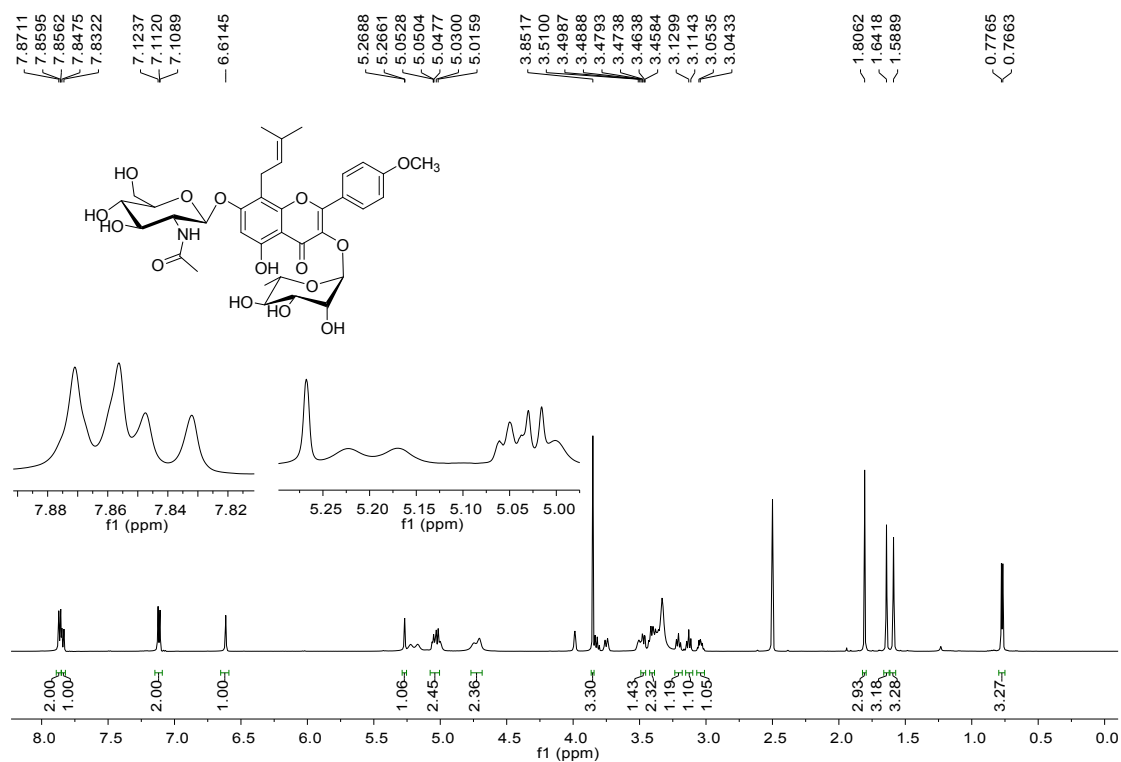


Figure S9. ¹H NMR spectrum of **1b** in DMSO-*d*₆

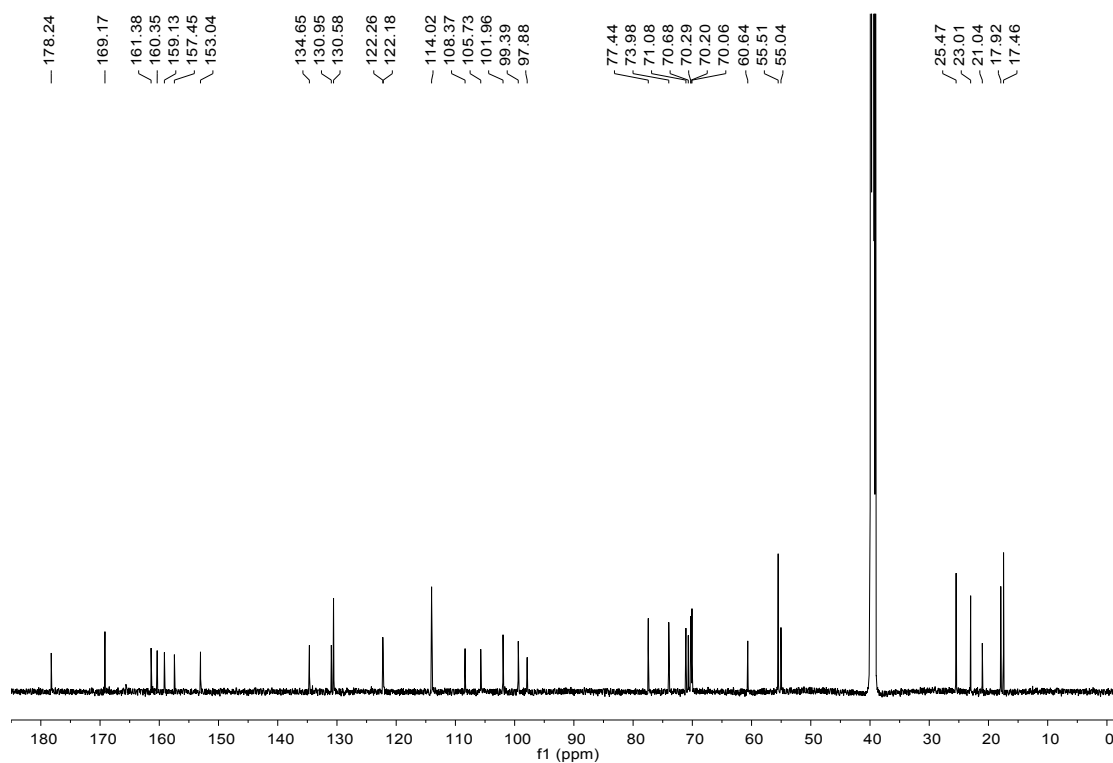


Figure S10. ¹³C NMR spectrum of **1b** in DMSO-*d*₆

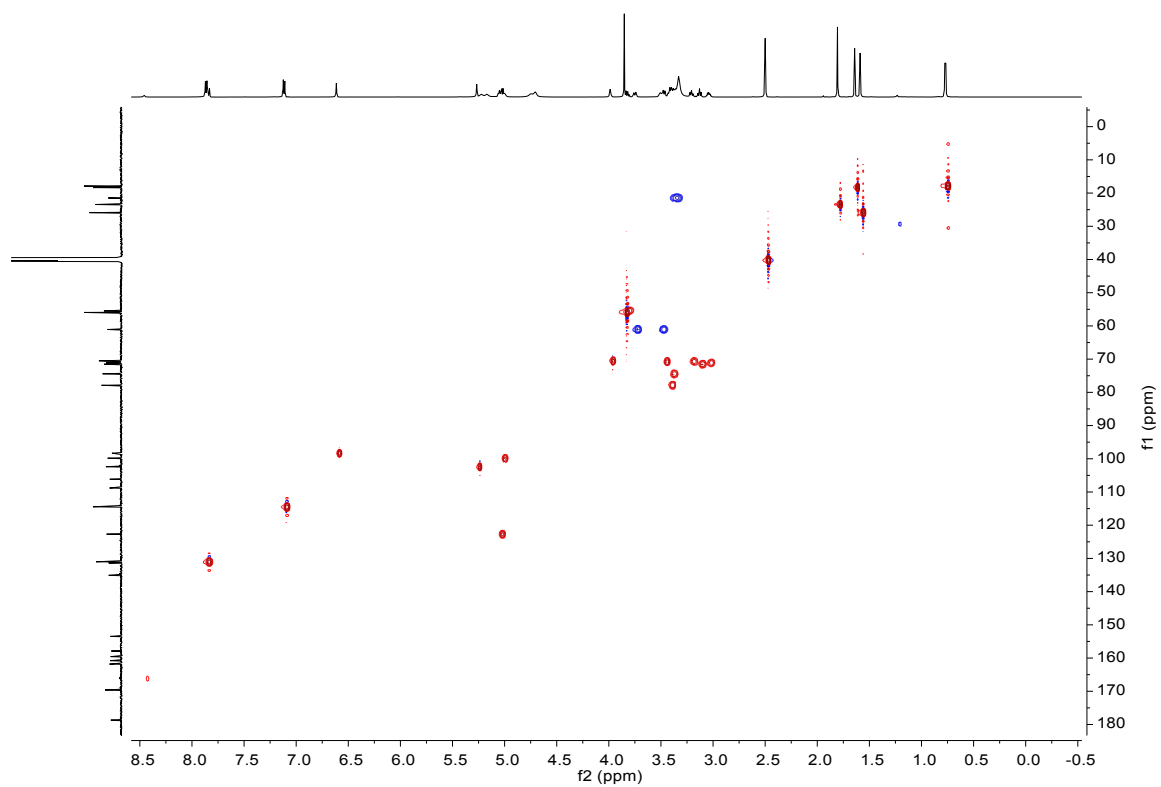


Figure S11. HSQC spectrum of **1b** in DMSO- d_6

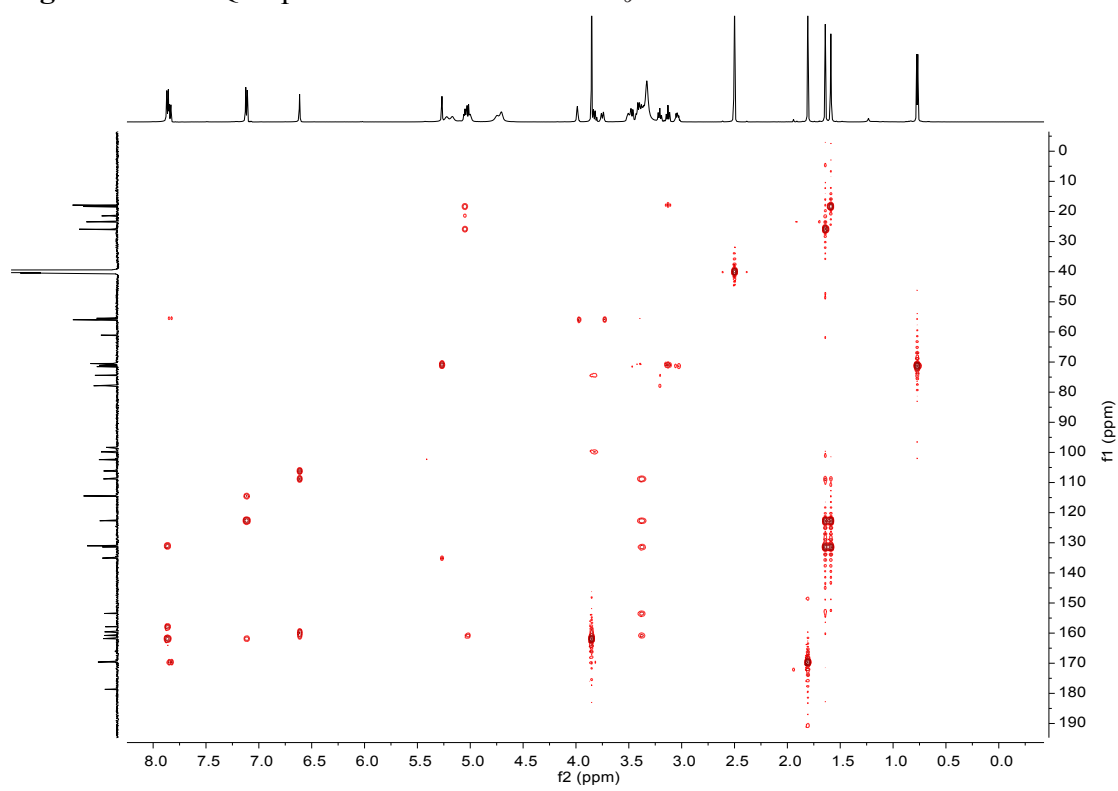


Figure S12. HMBC spectrum of **1b** in DMSO- d_6

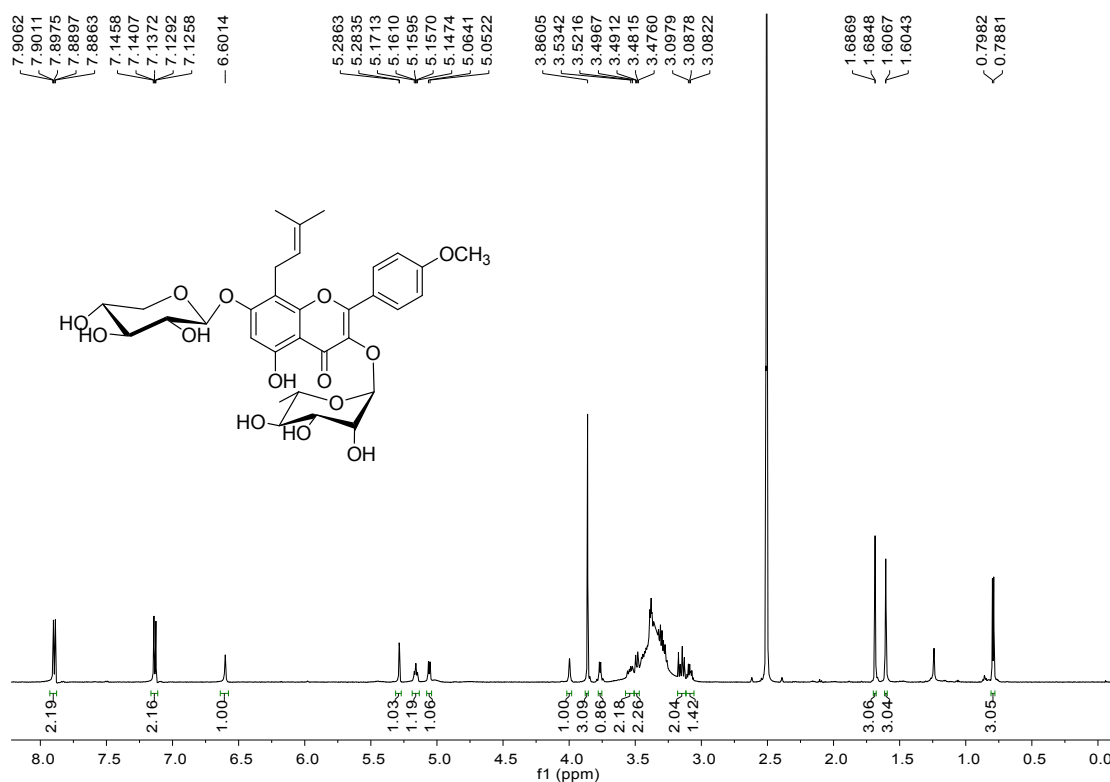


Figure S13. ¹H NMR spectrum of **1c** in DMSO-*d*₆

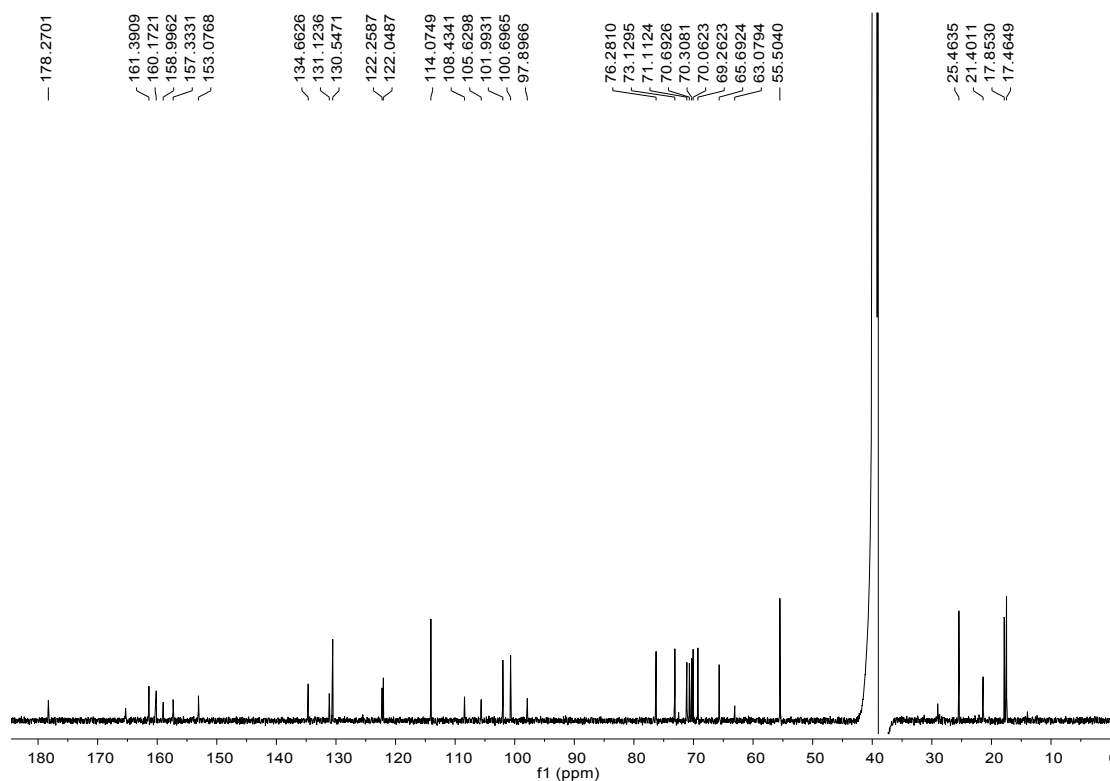


Figure S14. ¹³C NMR spectrum of **1c** in DMSO-*d*₆

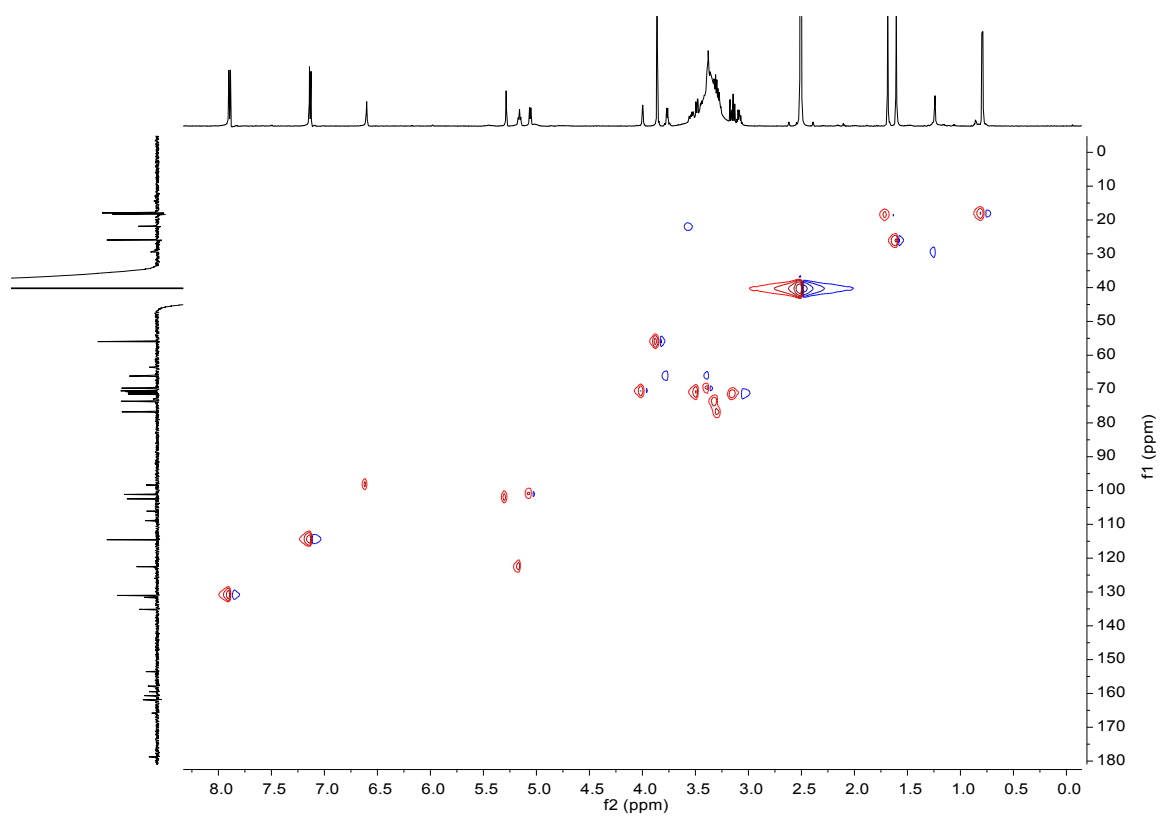


Figure S15. HSQC spectrum of **1c** in DMSO- d_6

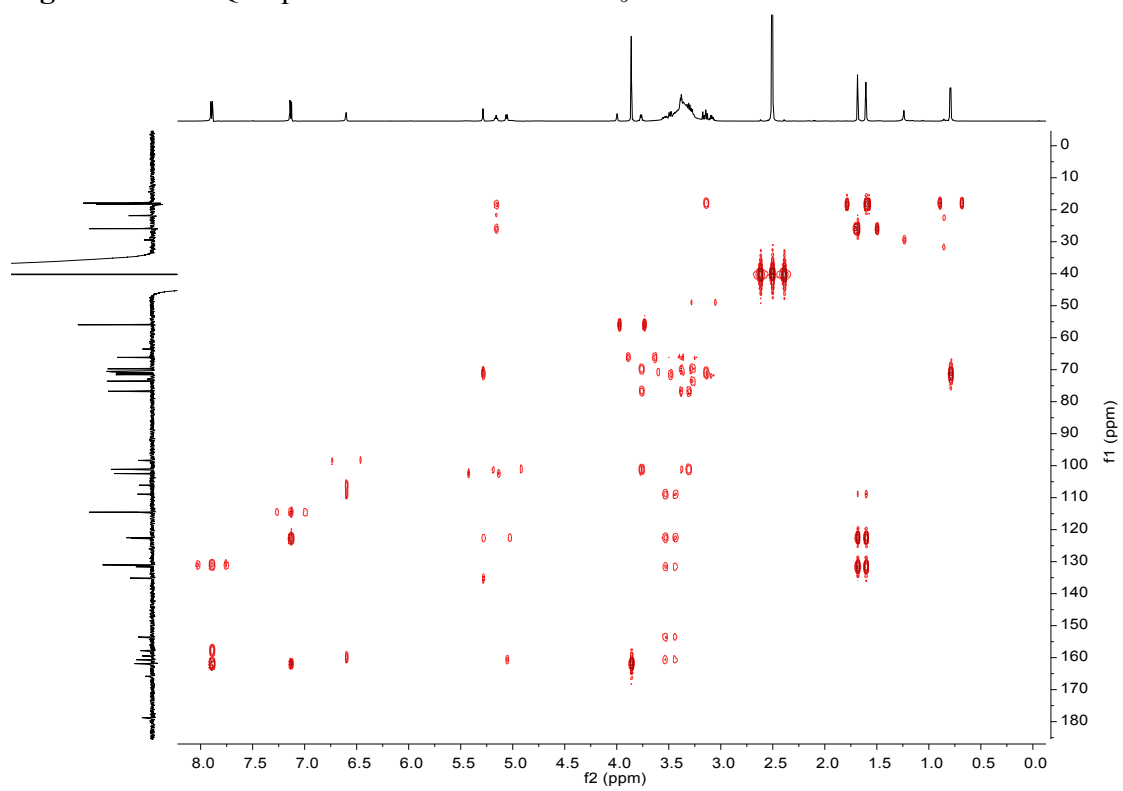


Figure S16. HMBC spectrum of **1c** in DMSO- d_6

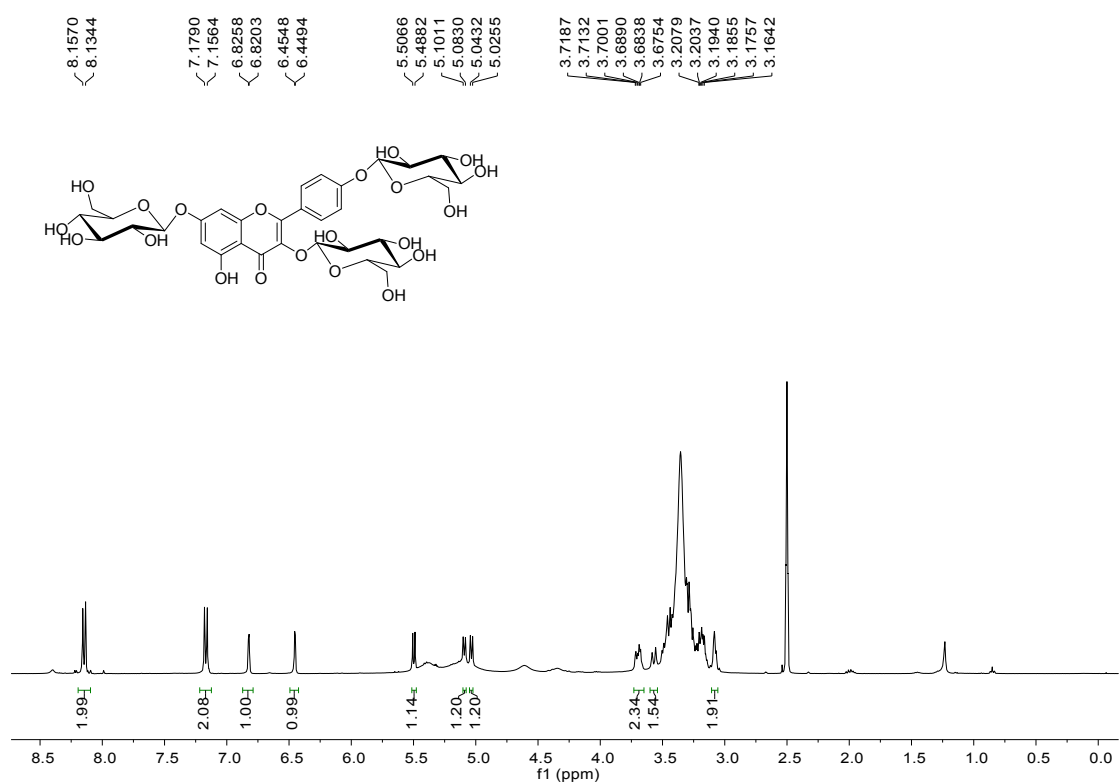


Figure S17. ^1H NMR spectrum of **2a** in $\text{DMSO}-d_6$

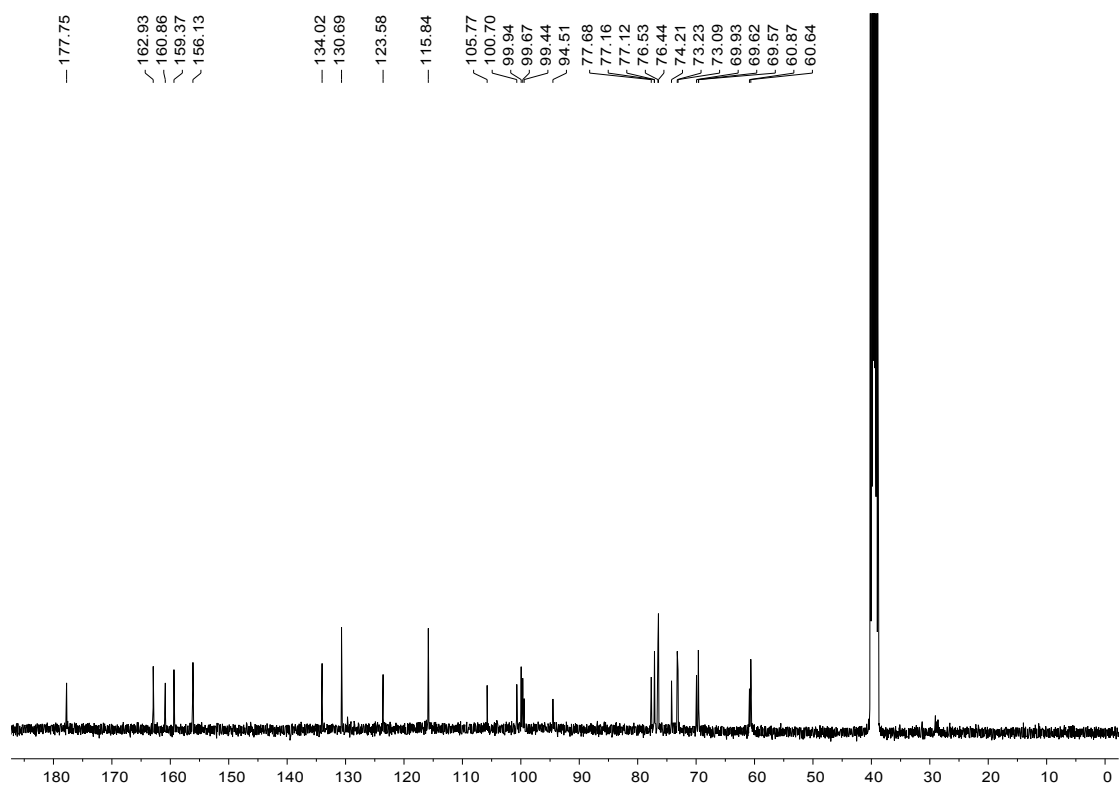


Figure S18. ^{13}C NMR spectrum of **2a** in $\text{DMSO}-d_6$

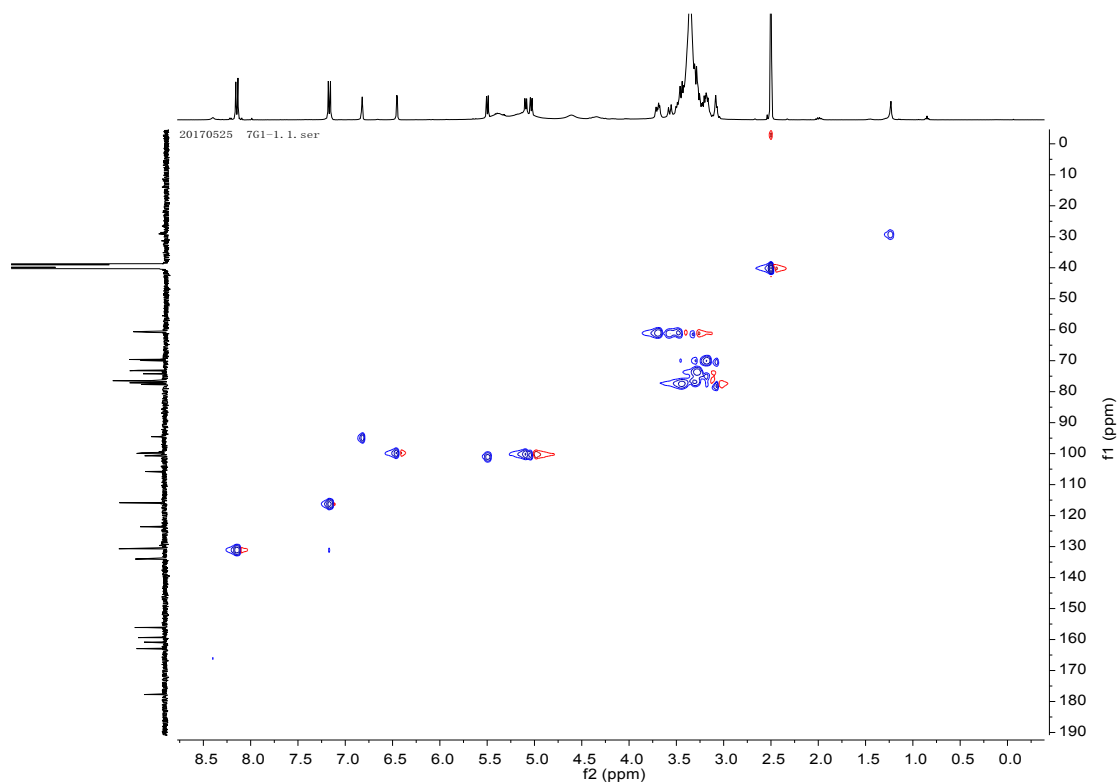


Figure S19. HSQC spectrum of **2a** in DMSO- d_6

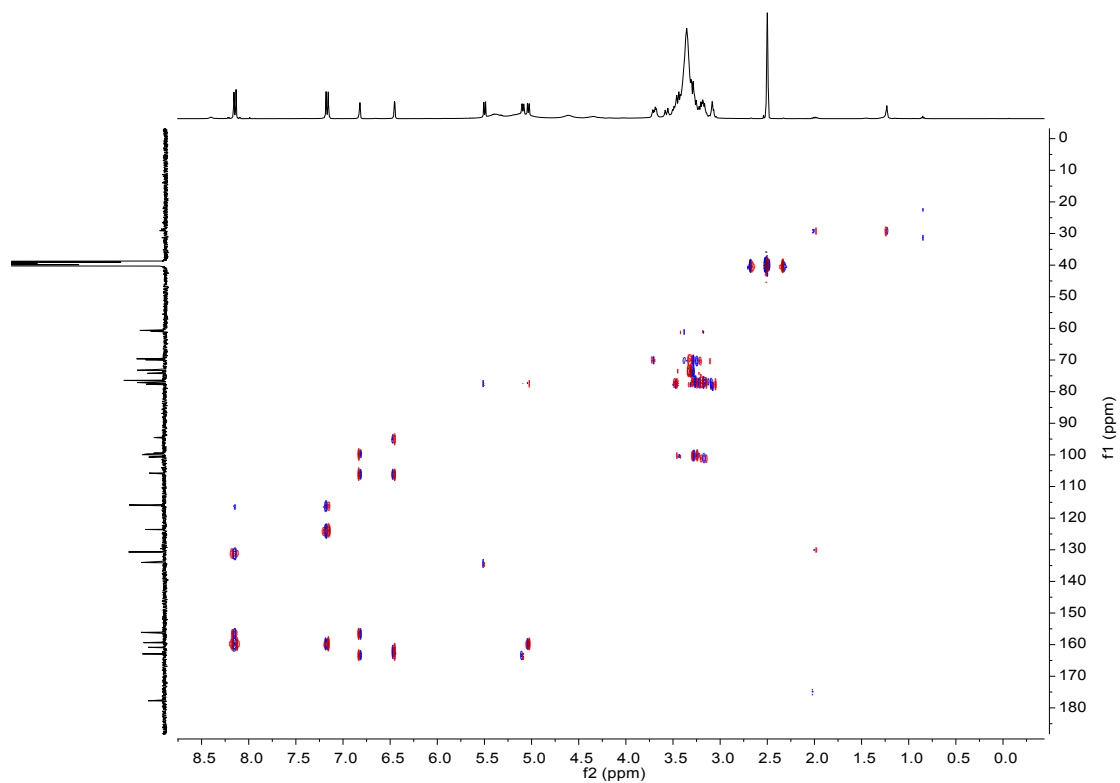


Figure S20. HMBC spectrum of **2a** in DMSO- d_6

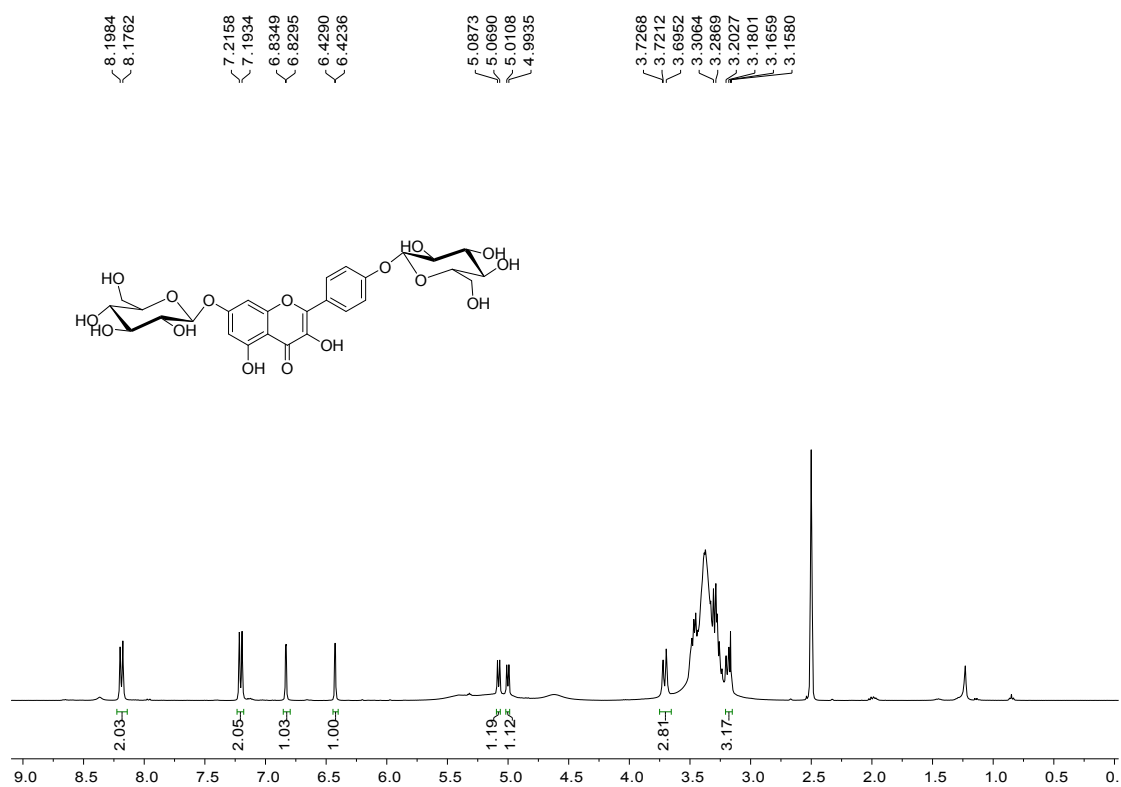


Figure S21. ^1H NMR spectrum of **2b** in $\text{DMSO}-d_6$

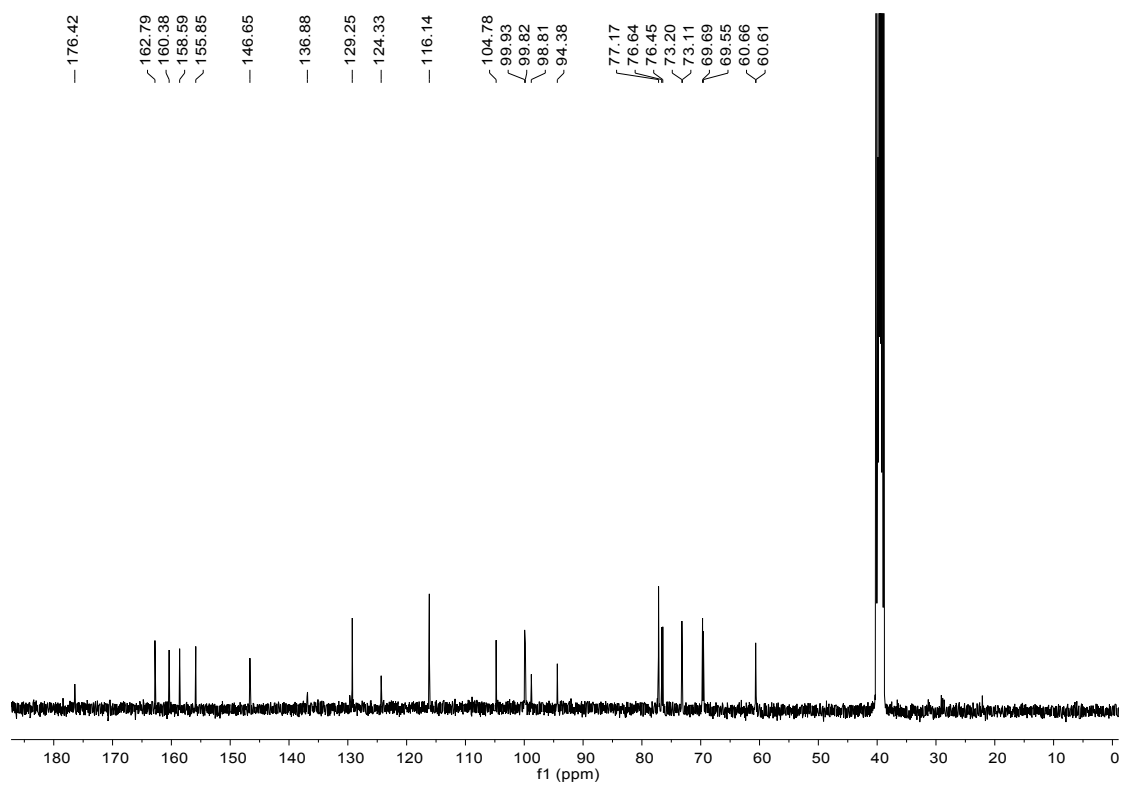


Figure S22. ^{13}C NMR spectrum of **2b** in $\text{DMSO}-d_6$

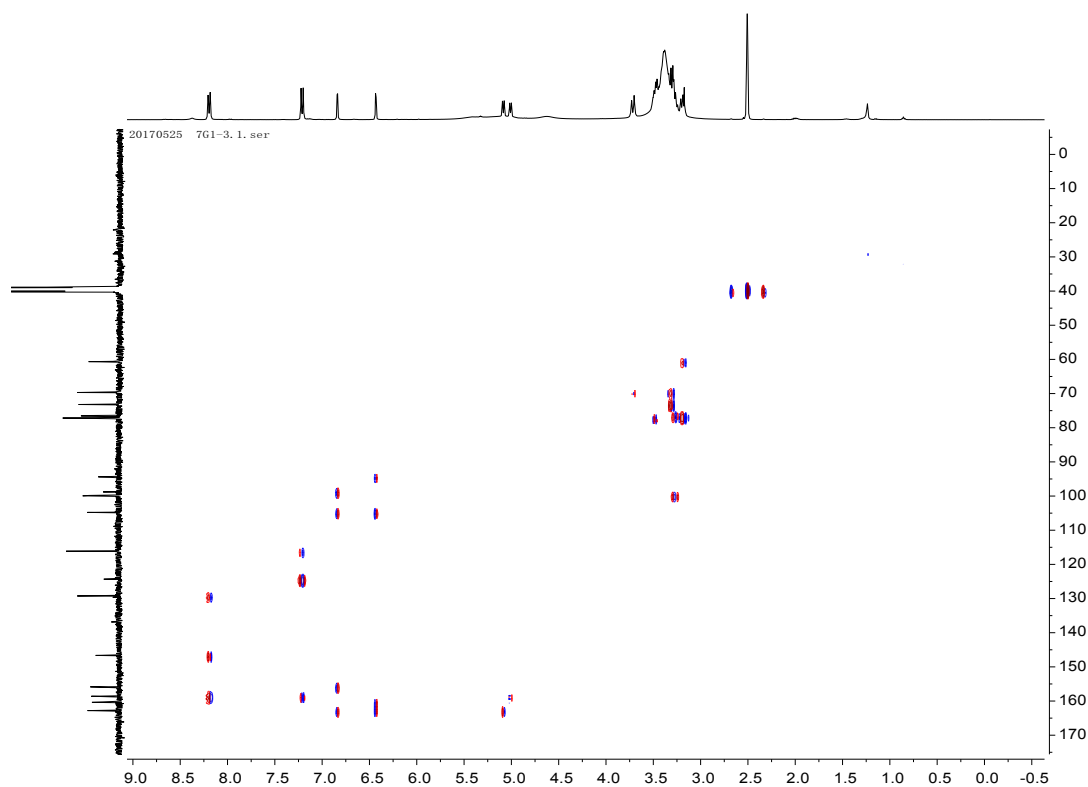


Figure S23. HMBC spectrum of **2b** in DMSO- d_6

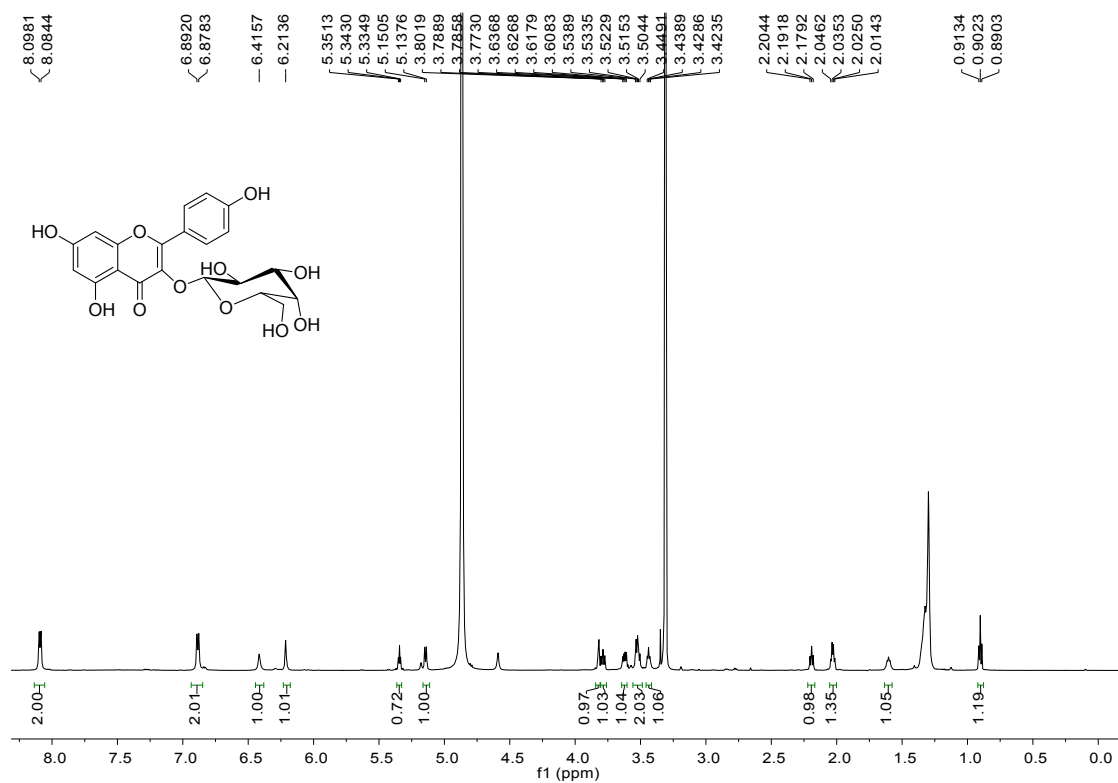


Figure S24. ^1H NMR spectrum of **2c** in Methanol- d_4

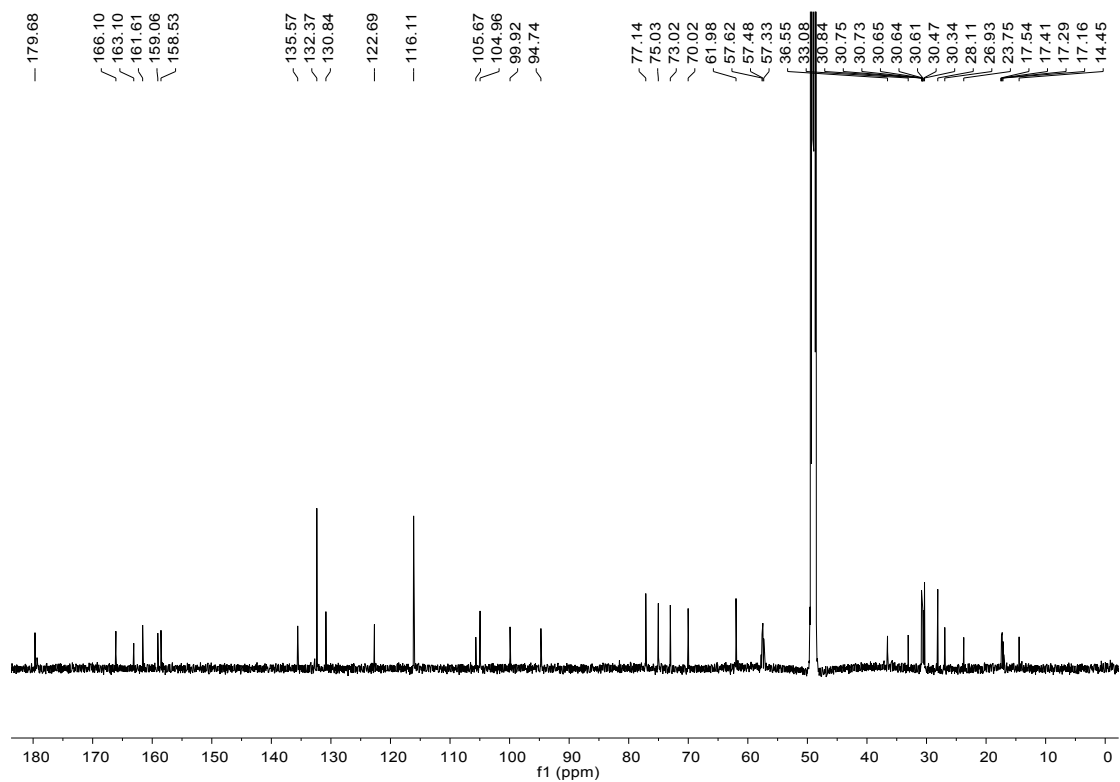


Figure S25. ^{13}C NMR spectrum of **2c** in Methanol- d_4

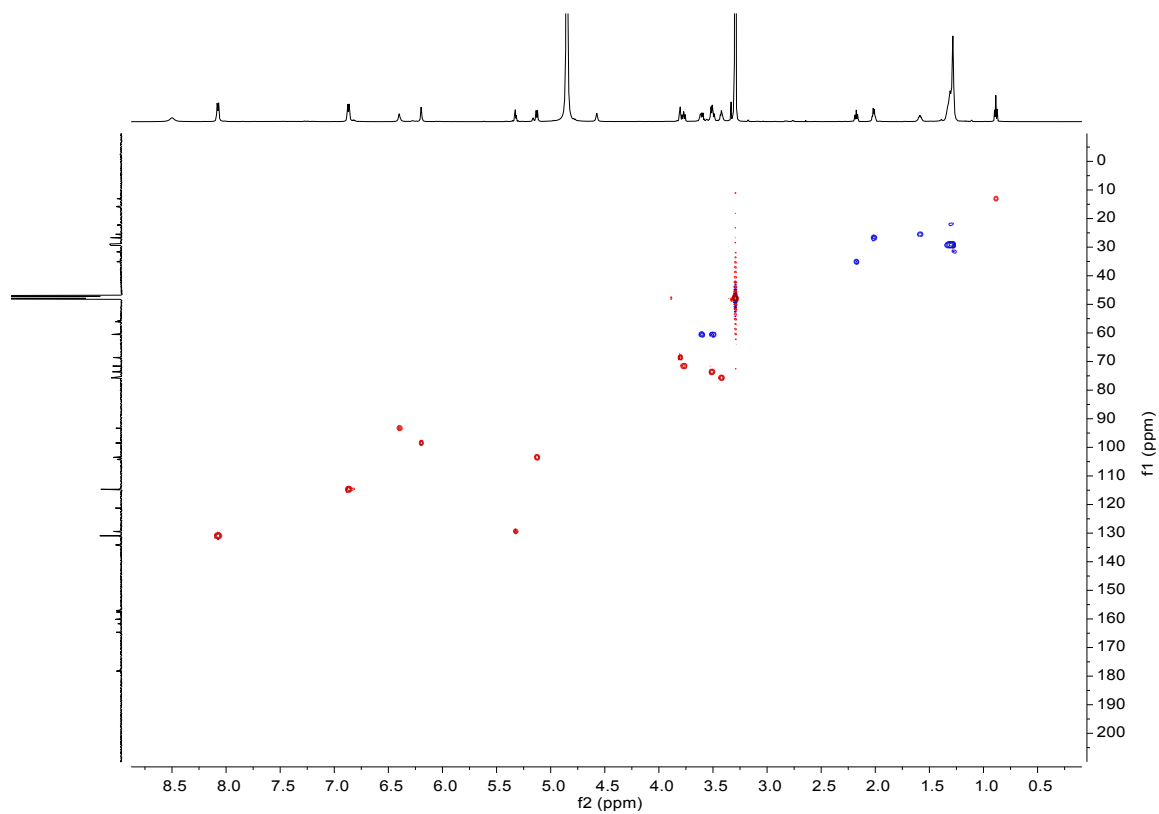


Figure S26. HSQC spectrum of **2c** in Methanol- d_4

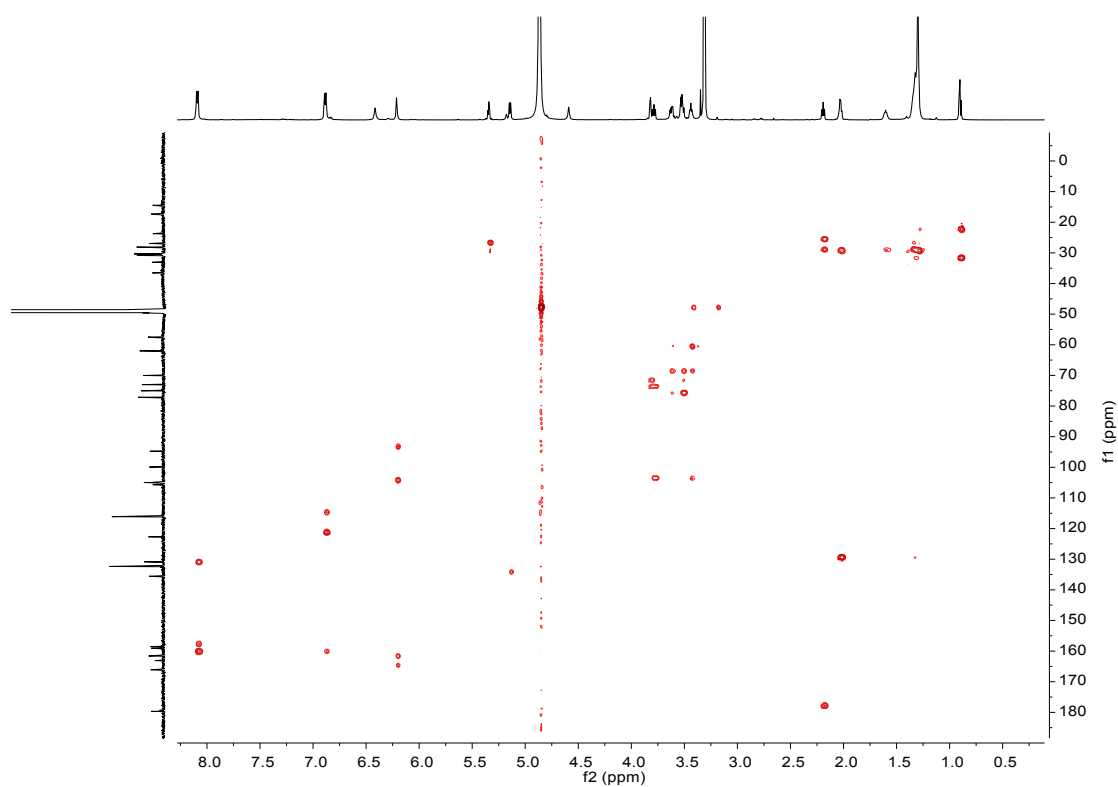


Figure S27. HMBC spectrum of **2c** in Methanol- d_4

BHG-NAC-2 #1241 RT: 3.12 AV: 1 NL: 3.96E4
T: FTMS - c ES+ Full

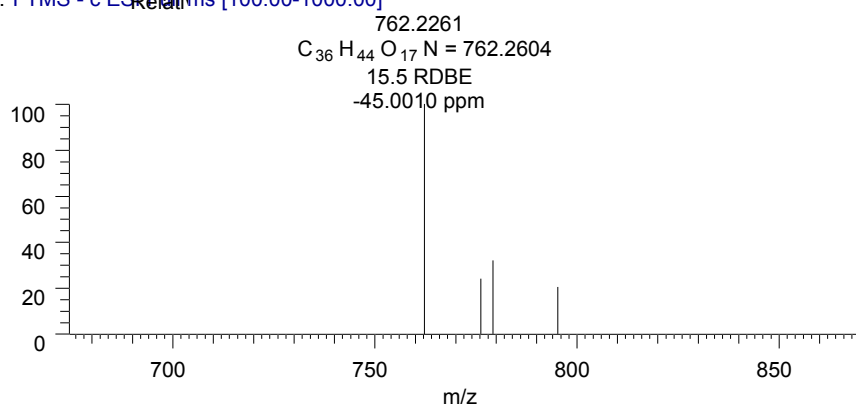


Figure S28. The HRESIMS spectrum of **1b**

BHG-XYL-2 #1387 RT: 3.58 AV: 1 NL: 4.50E7
T: FTMS - c ESI Full ms [100.00-1000.00]

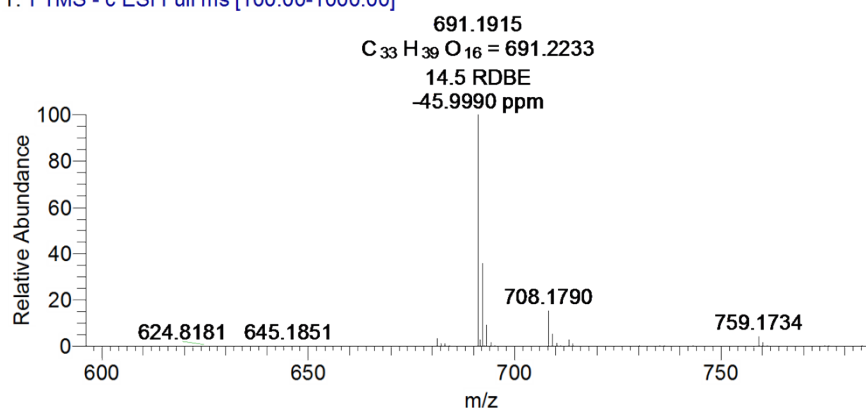


Figure S29. The HRESIMS spectrum of **1c**

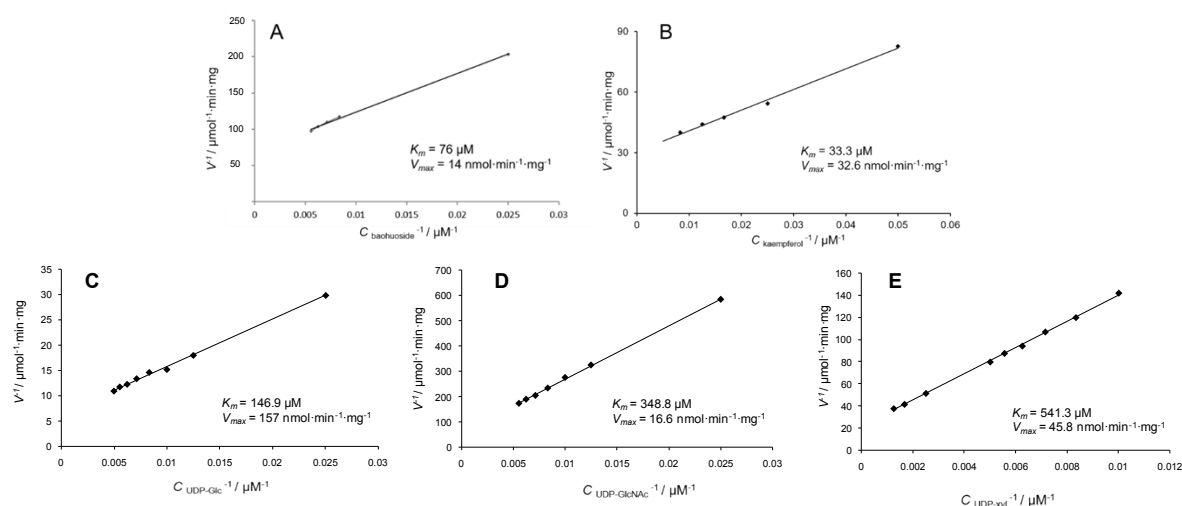


Figure S30. The apparent K_m values of recombinant Ep7GT for baohuoside (A), kaempferol (B), UDP-glucose (C), UDP-*N*-acetylglucosamine (D) and UDP-xylose (E).

Table S3. The apparent K_m , K_{cat} and K_{cat}/K_m values of recombinant Ep7GT for different sugar donors with **1** as the acceptor.

Sugar donors	K_m (μM)	K_{cat} (S^{-1})	K_{cat}/K_m ($\text{S}^{-1} \text{ mM}^{-1}$)
UDP-glucose	146.9	0.13	0.88
UDP- <i>N</i> -acetylglucosamine	348.0	0.014	0.04
UDP-xylose	541.3	0.038	0.07

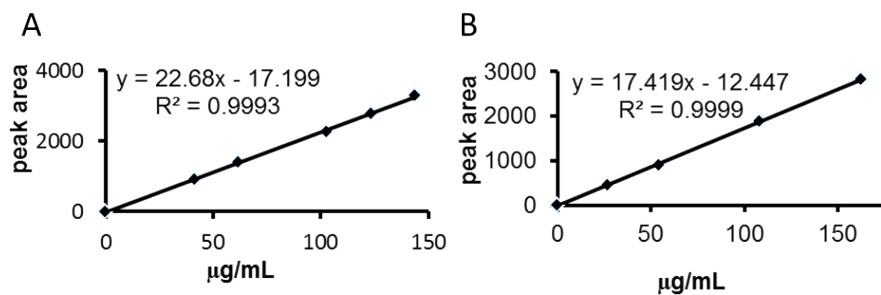


Figure S31. The linear regression models and the regression equations of the external standard method established for the quantitative analysis of baohuoside (A) and icariin (B).

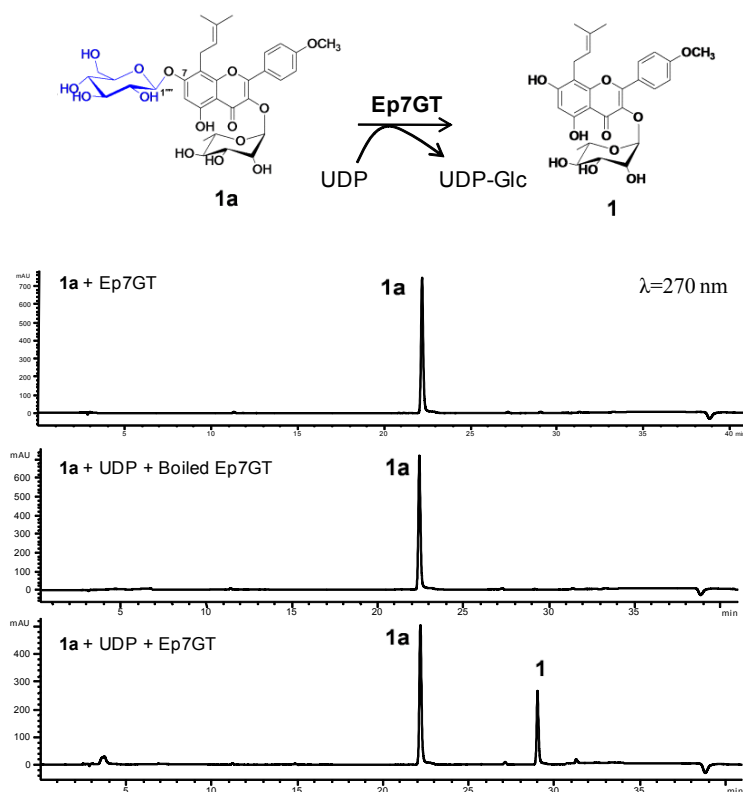


Figure S32. Exploring the catalytic reversibility of Ep7GT when **1a** and UDP were used as substrate.

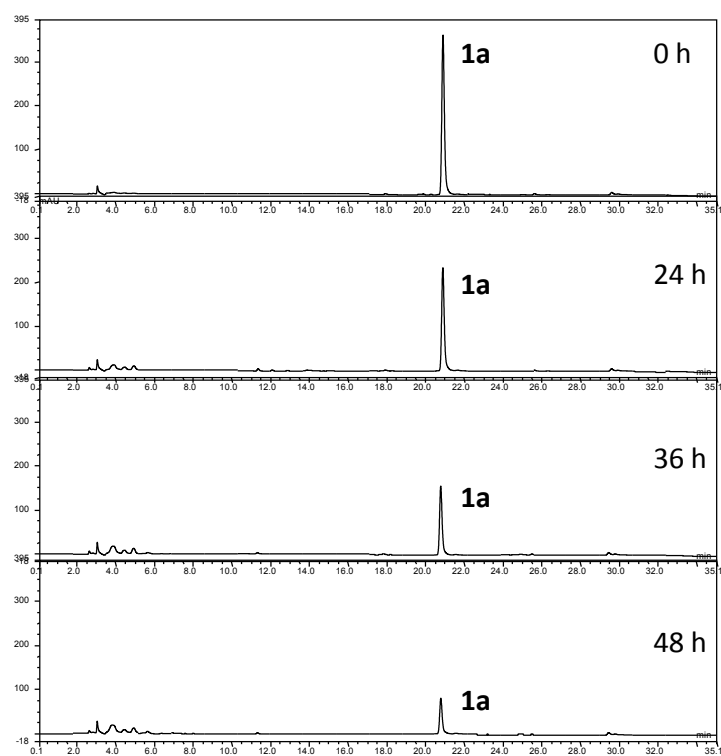


Figure S33. Time course assay of **1a** with whole cell catalyst of engineered *E. coli*.