Electronic Supplementary Material (ESI) for Catalysis Science & Technology.
This journal is © The Royal Society of Chemistry 2022

Supporting Information

Heparosan Oligosaccharide Synthesis Using Engineered Single-function Glycosyltransferases

Jian-Qun Deng[a], Zhen Lu[b], Juan Liu[a], Yan Zhao[b], Xu-Ben Hou[a], Xue-Ping Guo[b], Wen-Jie Jiang[a,c], Feng-Shan Wang[a,c], Ju-Zheng Sheng[a,c,*]

[a] Key Laboratory of Chemical Biology of Natural Products (Ministry of Education), School of Pharmaceutical Sciences, Cheeloo College of Medicine, Shandong University, Jinan 250012, China

[b] Bloomage BioTechnology Corp., Ltd., Jinan 250010, China

[c] National Glycoengineering Research Center, Shandong University, Jinan 250012, China

Table S1. Summary of heparosan oligosaccharides used for activity assays.

Compound	Structure of oligosaccharide	Catalyst[1]
H1	GlcA-*p*NP	Purchased
H2	GlcNAc-GlcA-*p*NP	EcKfiA or NaKfiA[2]
H3	GlcA-GlcNAc-GlcA-*p*NP	PmHS2
H4	GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	EcKfiA
H5	GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	PmHS2
H6	GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	EcKfiA
H7	GlcA-GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	PmHS2(D479N/D481N)
H8	GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	EcKfiA

Notes:

1. Enzyme involved in the heparosan oligosaccharide biosynthesis.
2. H2(GlcNAc-GlcA-*p*NP) synthesized by NaKfiA was used for structural confirmation in Figure S6 and S7.

Table S2. Kinetic parameters of EcKfiA and NaKfiA.

Enzyme	V_{max} (μM/s)	k_{cat} (/s)	Acceptor: *p*NP-GlcA		Donor: UDP-GlcNAc	
			K_m^A (μM)	k_{cat}/K_m^A (1/μM s)	K_m^D (μM)	k_{cat}/K_m^D (1/μM s)
EcKfiA	0.26	730	510±20	1.4±0.1	250±40	2.9±0.4
NaKfiA	0.17	470	720±60	0.7±0.1	570±50	0.8±0.1
NaKfiA C16L/S165K	0.37	1020	660±120	1.6±0.3	430±60	2.4±0.4

Table S3. Summary of heparosan oligosaccharides produced by engineered single-function glycosyltransferases.

Compound	Structures of oligosaccharide	Synthetic scale	Overall yield[1]	Enzyme[2]	Product purity[3]
H6a	GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	4.9 mg	98%	NaKfiA C16L/S165K	>99%
H7a	GlcA-GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	5.5mg	96%	EcKfiC	>95%
H8a	GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	6.1 mg	94%	NaKfiA C16L/S165K	>95%
H9	GlcA-GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-GlcNAc-GlcA-*p*NP	6.6 mg	91%	EcKfiC	>95%

Notes:

1. Overall yield was calculated based on the amount of GlcA-*p*NP.
2. Enzyme involved in the heparosan oligosaccharide biosynthesis.
3. The purity was assessed by PAMN-HPLC.

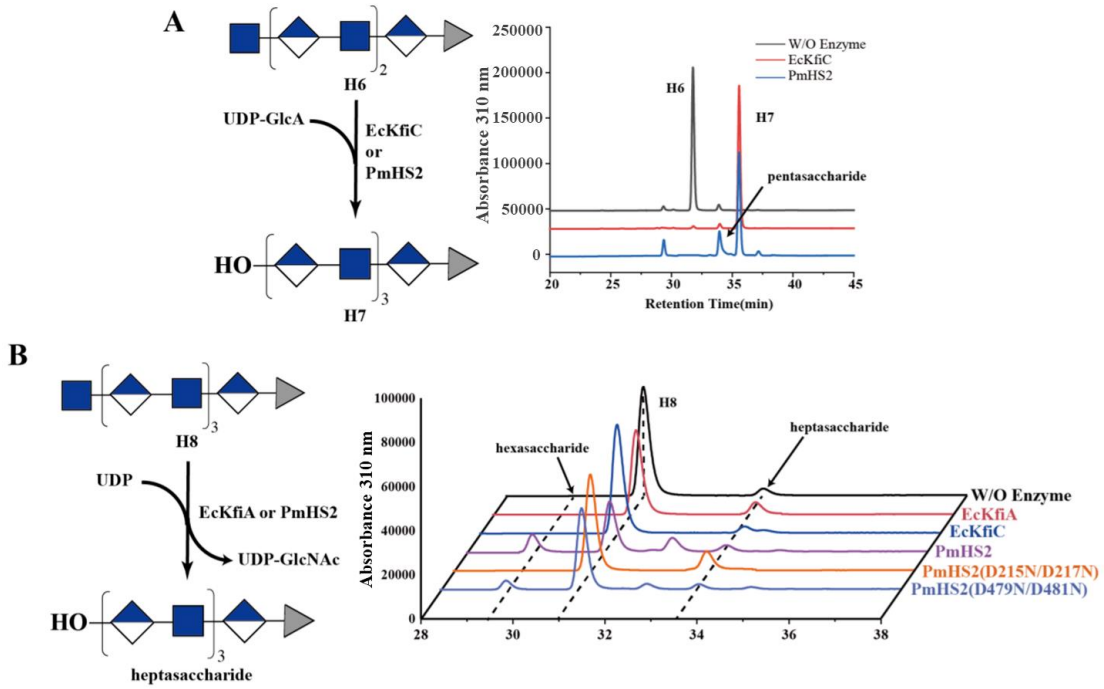

Figure S1. Determination of the reverse-glycosylation activity of single- and dual-function glycosyltransferases. **(A)** Reactions involved in the elongation of substrate H6 by Escherichia coli KfiC (EcKfiC) and Pasteurella multocida heparosan synthase 2 (PmHS2). Left panel: Schematic illustrating the products of the enzyme-catalyzed reactions. Right panel: Polyamine-based anion exchange–high-performance liquid chromatography (PAMN-HPLC) of H6 incubated with EcKfiC or PmHS2. Elongation reactions were carried out in the presence of UDP-N-acetylglucosamine (UDP-GlcNAc). **(B)** Analysis of degradation of H8 by EcKfiA, EcKfiC, PmHS2, or engineered PmHS2 single-function glycosyltransferase mutants. Left panel: Schematic illustrating the products of the reverse glycosylation reactions. Right panel: PAMN-HPLC of H8 incubated with different enzymes. Degradation reactions were carried out in similar conditions in the presence of UDP (2.0 mM), but UDP-monosaccharides were absent.

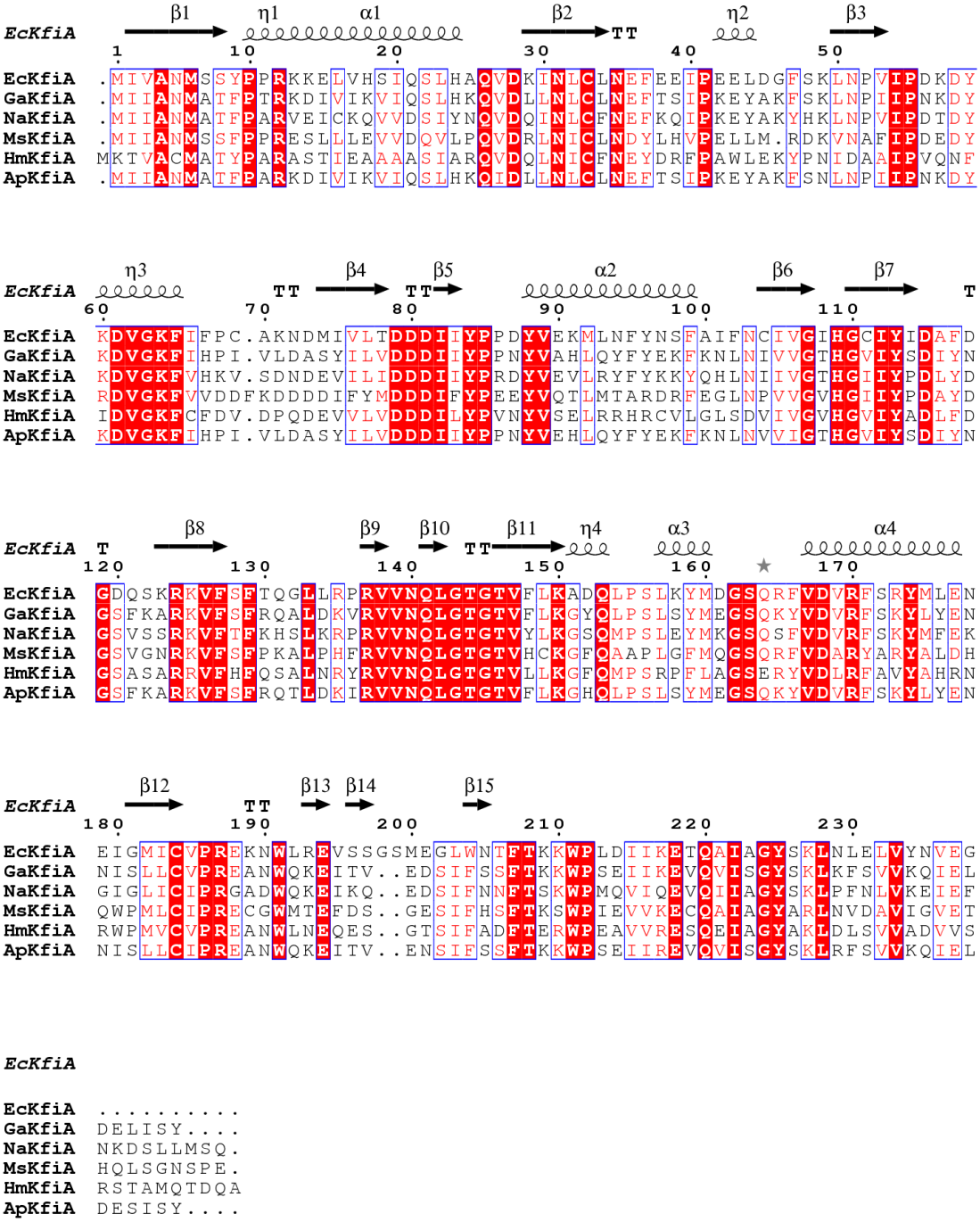

Figure S2. Sequence alignment of EcKfiA (from *Escherichia coli* O10:K5:H4), GaKfiA (from *Gallibacterium anatis*), NaKfiA (from *Neisseria animaloris*), MsKfiA (from *Marinimicrobium* sp. LS-A18), HmKfiA (from *Hyphomicrobium methylovorum*), and ApKfiA (from *Avibacterium paragallinarum*). Strictly conserved residues are highlighted by a red background, and conservatively-substituted residues are boxed. The secondary structure of EcKfiA is shown above the aligned sequences.

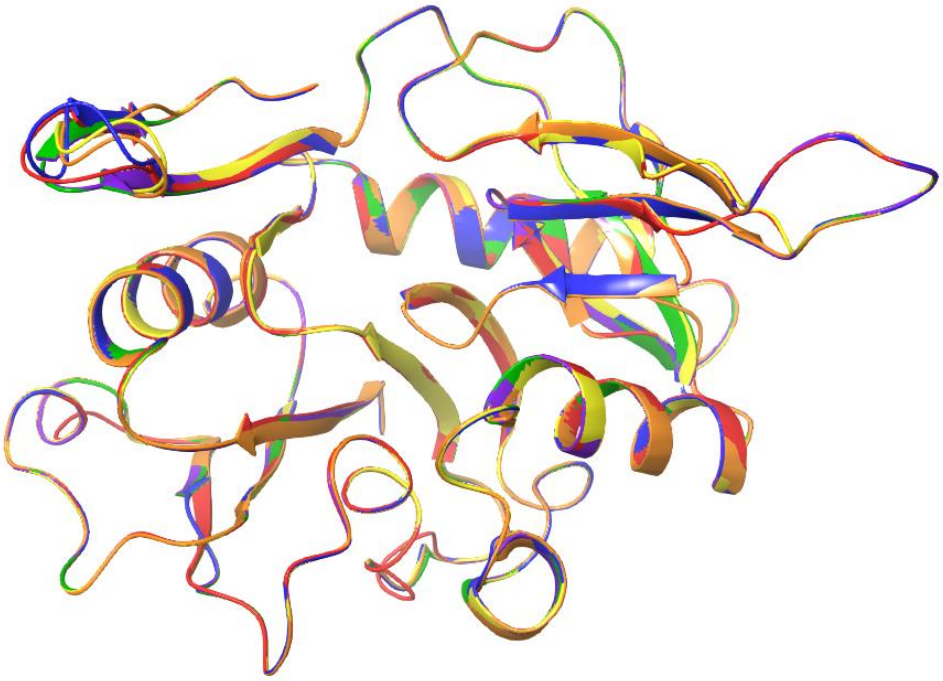

Figure S3. Structural alignment of EcKfiA (blue), GaKfiA (green), NaKfiA (yellow), MsKfiA (red), ApKfiA (purple), and HmKfiA (orange). Three dimensional protein structure model of KfiAs was built based on the structure of EcKfiA (https://www.rcsb.org/structure/5Z8B) and was visualized by PyMOL.

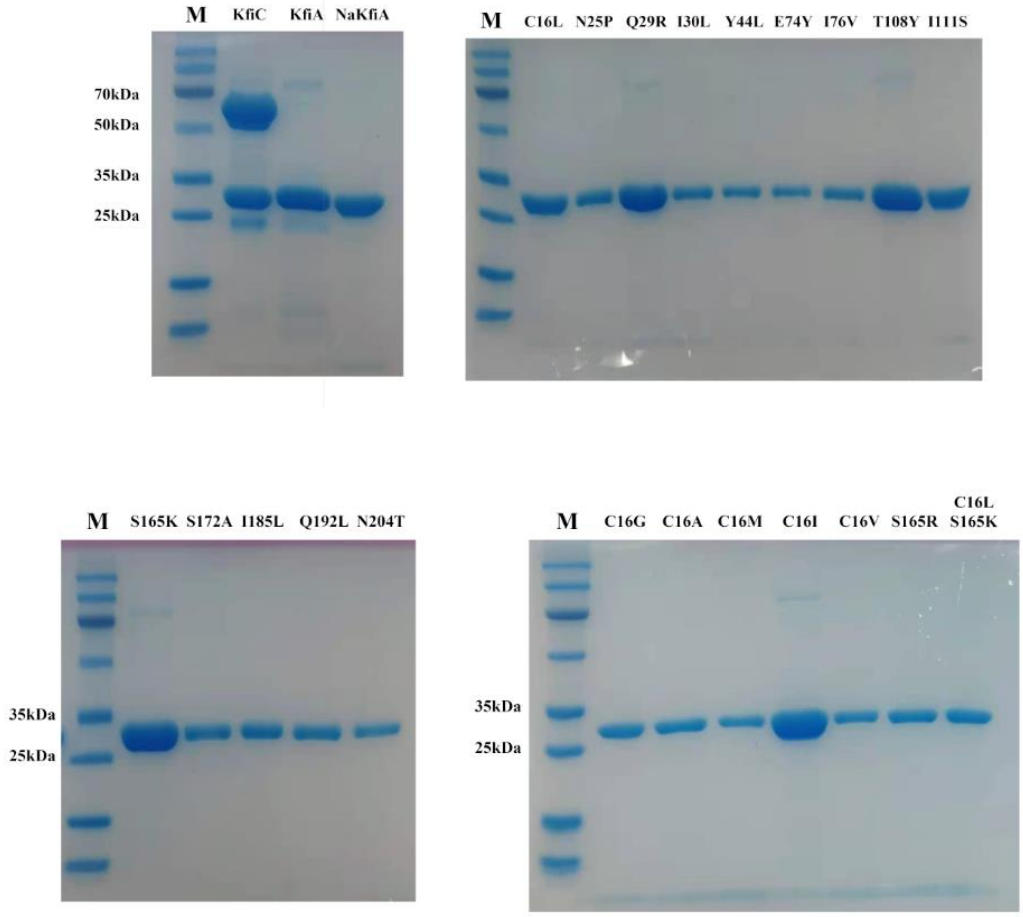

Figure S4. SDS-PAGE analysis of purified proteins. Lane M, protein molecular weight markers (Sangon Biotech TrueColor Pre-stained Protein Marker, 15–130 kDa). In addition, TF-KfiC was treated with thrombin (Solarbio, T8021) at 25 ℃ overnight to cut off TF factors. Each band corresponds to the expected molecular weight: KfiC, 60.81 kDa; KfiA, 28.15 kDa; and NaKfiA and its mutants, 29.37 kDa.

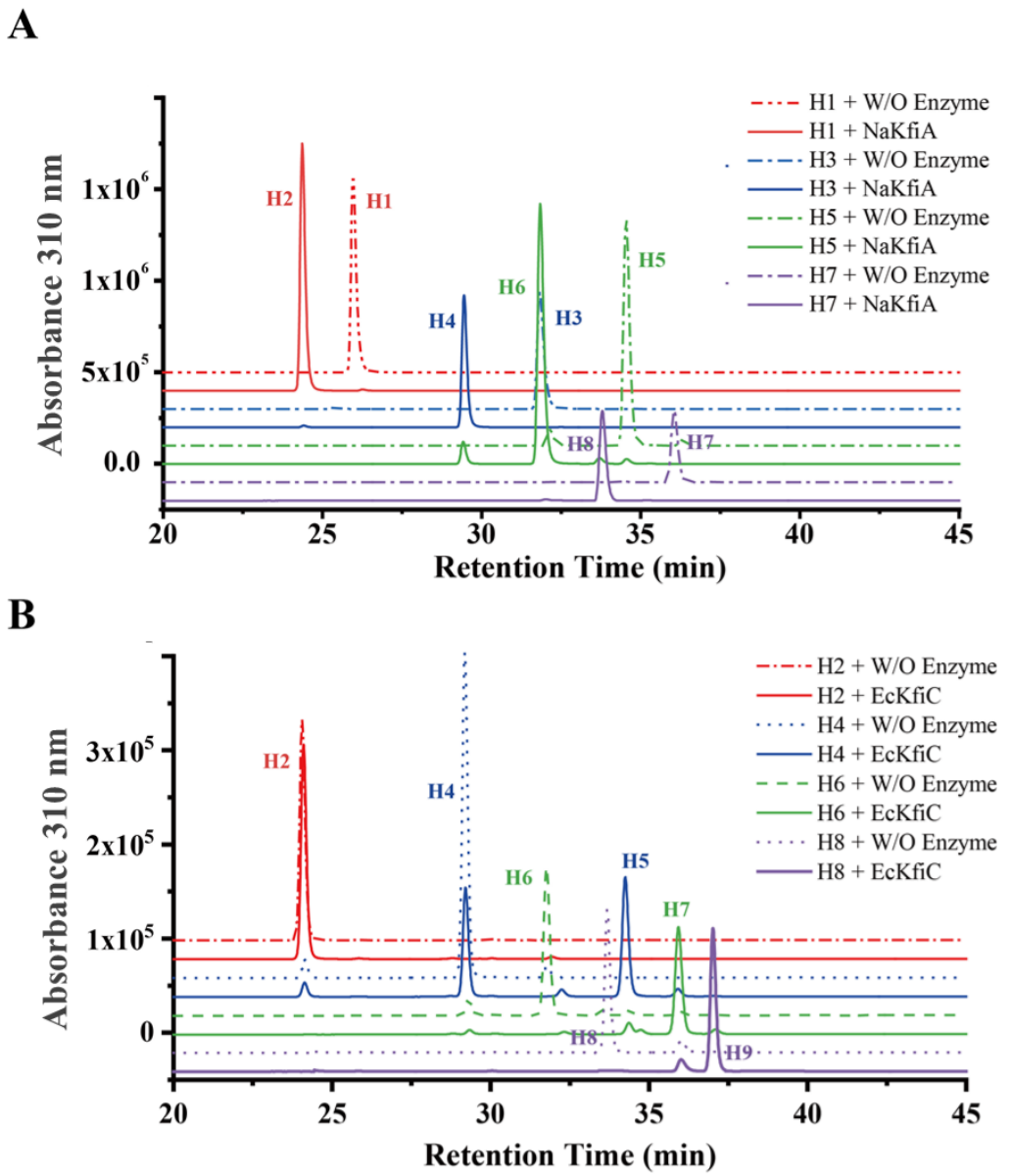

Figure S5. PAMN-HPLC analysis of the GlcNAc-T activity of NaKfiA **(A)** and the GlcA-T activity of KfiC **(B)**. Reactions were carried out in similar conditions in the presence of 25 mM Tris-HCl (pH 7.5), 10 mM $MnCl_2$, 0.2mM acceptor, 0.4 mM UDP-sugar(UDP-GlcNAc for NaKfiA or UDP-GlcA for EcKfiC) and 0.1mg/ml enzyme. All experiments were carried out at 37 ℃ for 18 h, and analyzed by PAMN-HPLC. NaKfiA shows GlcNAc-T activity using **H1**, **H3**, **H5**, and **H7** as aeceptors. EcKfiC shows GlcA-T activity using **H4**, **H6**, and **H8** as aeceptors.

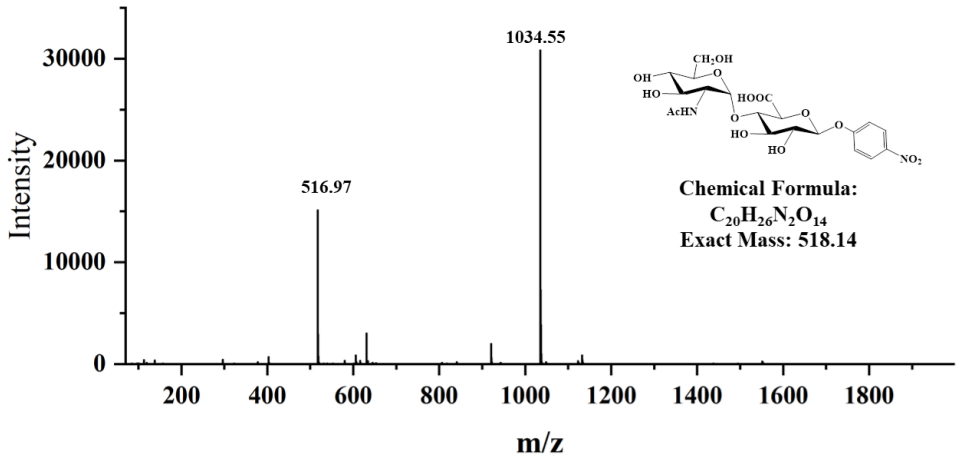

Figure S6. Electrospray ionization mass spectrum of NaKfiA-treated GlcA-*p*NP.

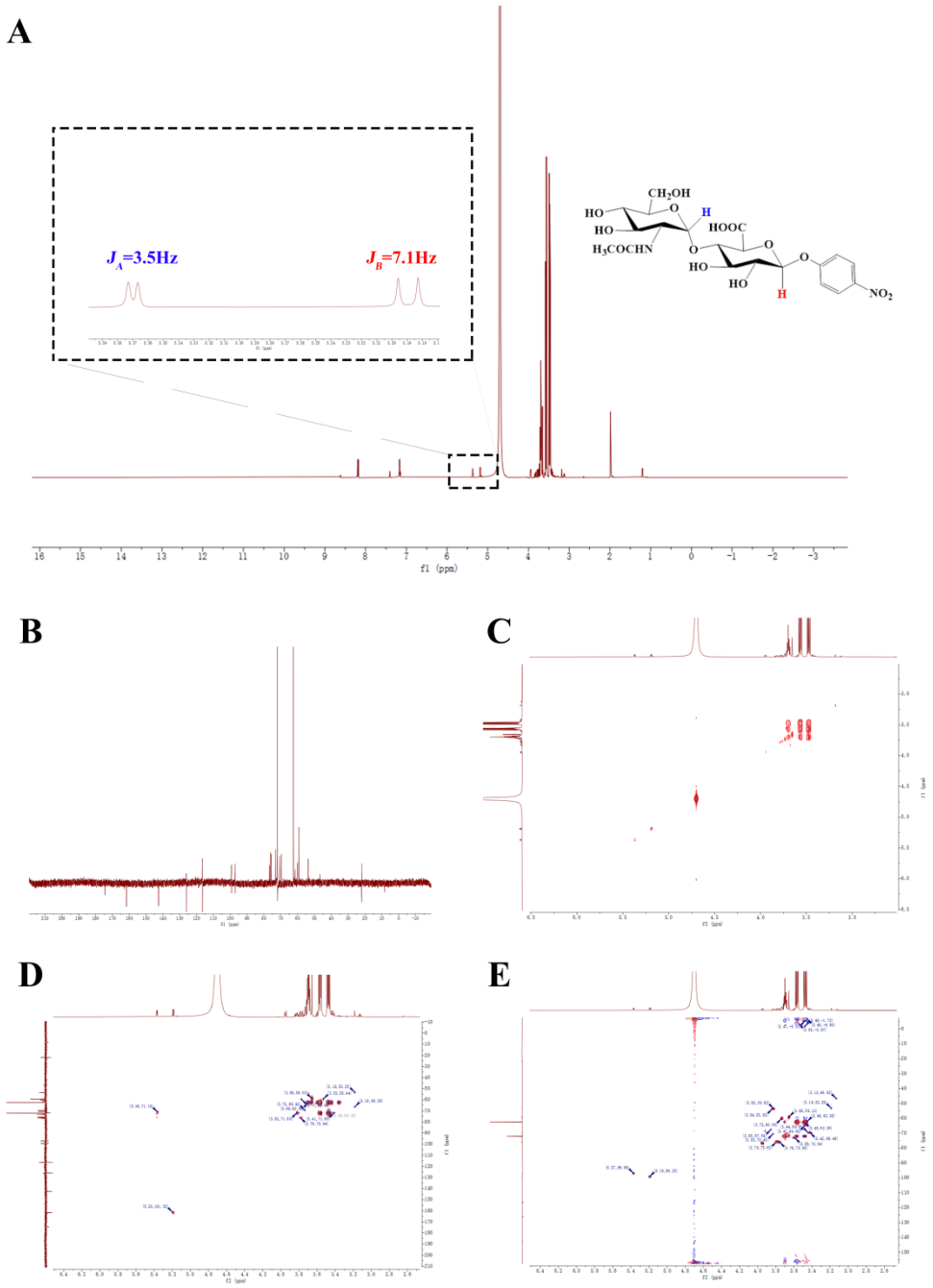

Figure S7. NMR spectra of GlcA-*p*NP treated with NaKfiA. **(A)** ^{1}H NMR (600 MHz, D_2O) δ 5.3 (d, J = 3.45 Hz, 1H). **(B)** ^{13}C NMR (600 MHz, D_2O). **(C)** H–H COSY (600 MHz, D_2O). **(D)** HMBC (600 MHz, D_2O). **(E)** HSQC (600 MHz, D_2O). The coupling constant (~3.5 Hz), indicates that the glycosidic bonds between GlcA and GlcNAc residues were α-(1, 4) linkages; the coupling constant of the anomeric proton was near to 4 Hz.

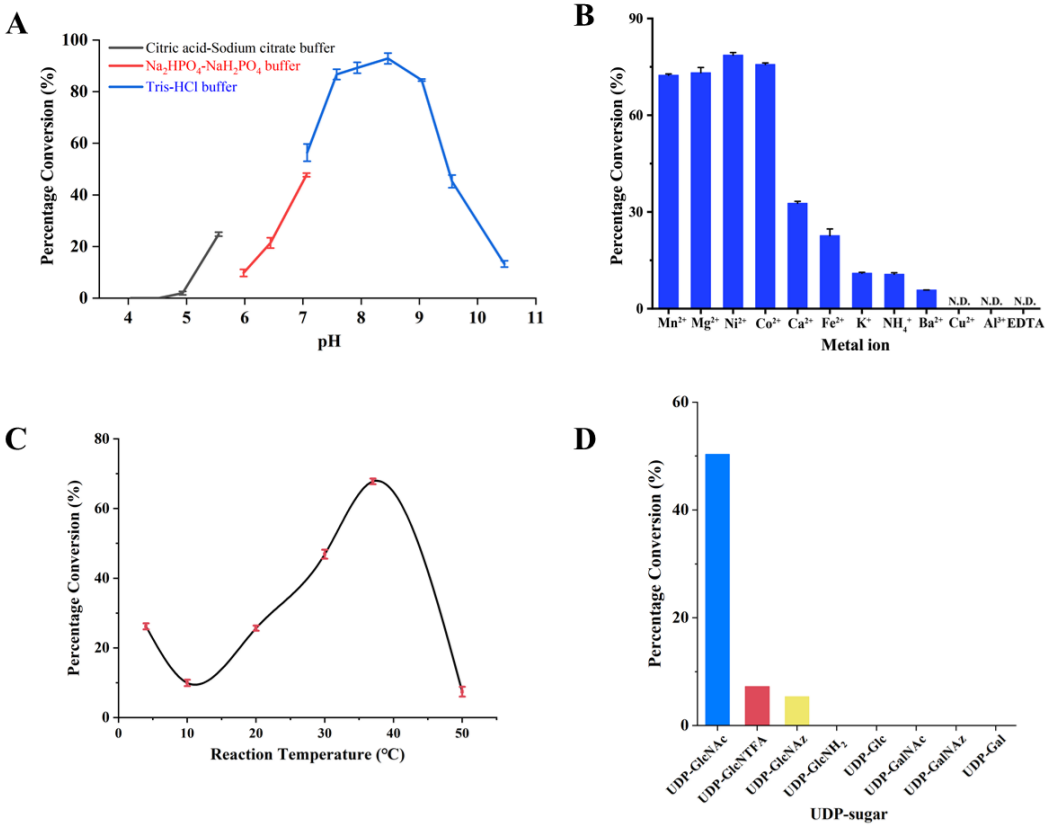

Figure S8. Determination of conditions for optimal catalytic activity of NaKfiA. **(A)** The effect of pH profile on GlcNAc-T activity. Recombinant NaKfiA had its highest activity at pH ~8 in Tris-HCl buffer. **(B)** The optimum mental ion of NaKfiA. NaKfiA exhibited its optimal activity in the presence of Mn^{2+}, Mg^{2+}, Ni^{2+} or Co^{2+}ions (10 mM). **(C)**The temperature preference of NaKfiA. The activity of purified NaKfiA peaked at 37 ℃. **(D)** Substrate tolerance of NaKfiA. Reactions contained GlcA-*p*NP(0.2mM) as the acceptor substrate in a reaction mixture containing Tris-HCl (25 mM, pH 7.5), $MnCl_2$ (10 mM) and UDP-sugar(0.3mM), in one hour reaction time.

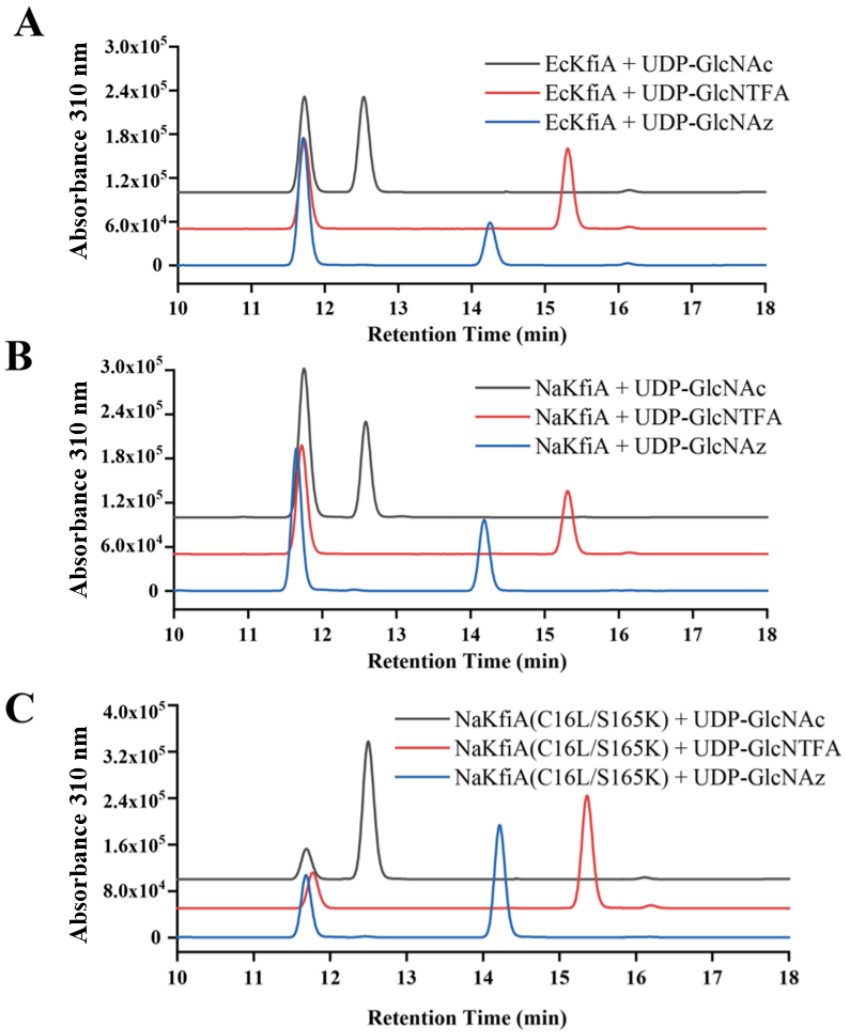

Figure S9. HPLC assays of GlcA-*p*NP treated with EcKfiA(A), NaKfiA(B) and NaKfiA(C16L/S165K)(C) using various donors. Reactions contained GlcA-*p*NP(0.2mM) as the acceptor substrate in a reaction mixture containing Tris-HCl (25 mM, pH 7.5), $MnCl_2$ (10 mM) and UDP-sugar(0.3mM), expect UDP-GlcNAc (black) as natural donor substrate, the other nucleic acid sugar derivatives were also analyzed specificity as donor substrate, including UDP-GlcNTFA (red), UDP-GlcNAz (blue). Reverse phase C18 column (250×4.6 mm, 5 μm, Shimadzu) was used. The analytical method was gradient elution from 20% methanol/ H_2O to 60% methanol/H_2O in 20 min. Retention time: GlcA-*p*NP, eluted at ~11.7 min; GlcNAc-GlcA-*p*NP, eluted at ~12.5 min; GlcNAz-GlcA-*p*NP, eluted at ~14.2 min; GlcNTFA-GlcA-*p*NP, eluted at ~15.4min.

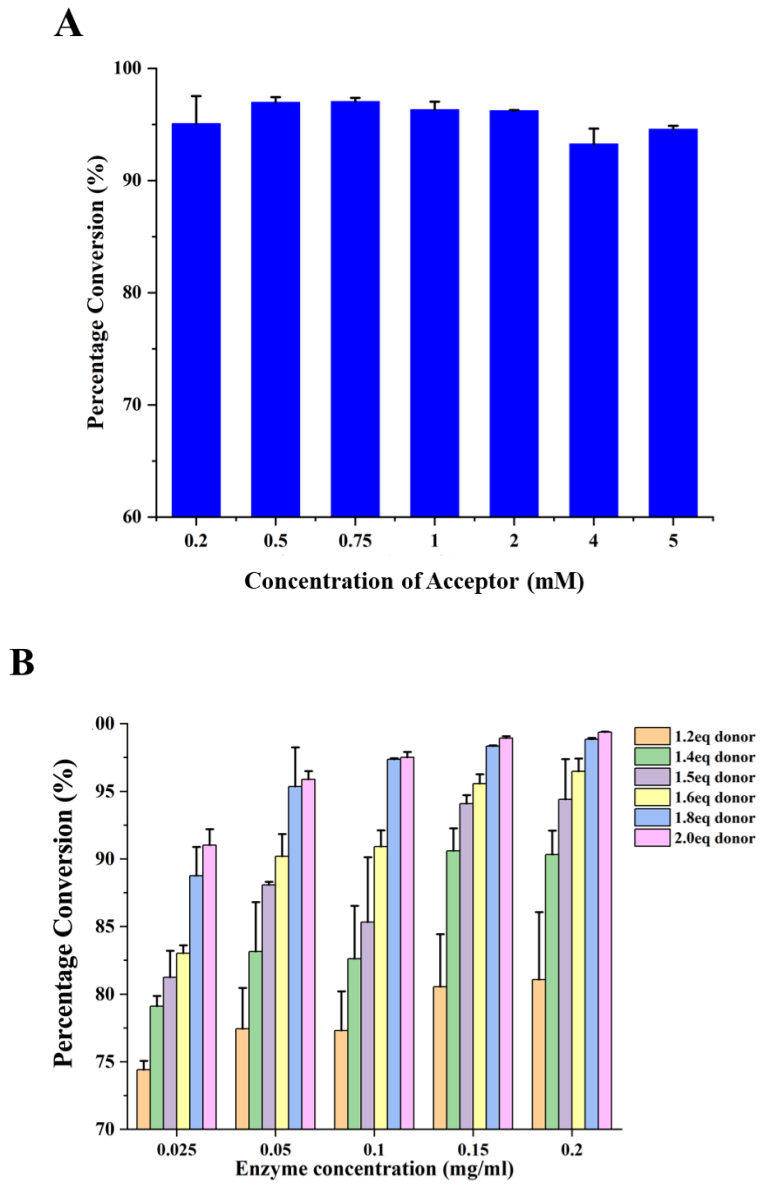

Figure S10. Determination of the optimum reaction conditions for NaKfiA. (A) The substrate tolerance of NaKfiA. Reactions were carried out in similar conditions in the presence of 25 mM Tris-HCl (pH 7.5), 10 mM $MnCl_2$, different concentrations of aeceptor, 1.5 equivalent UDP-GlcNAc and 0.1mg/ml enzyme. All experiments were carried out at 37 ℃ for 18 h, and analyzed by PAMN-HPLC. NaKfiA showed broad tolerance toward pNP-GlcA concentrations (0.2 to 5 mM). **(B)** The effects of the ratio of donor concentration to acceptor concentration, and NaKfiA concentration. Reactions were carried out in similar conditions in the presence of 25 mM Tris-HCl (pH 7.5), 10 mM $MnCl_2$, 5mM *p*NP-GlcA , different equivalent UDP-GlcNAc and different concentrations of enzyme. All experiments were carried out at 37 ℃ for 18 h, and analyzed by PAMN-HPLC. The conversion of acceptor (5 mM) to longer products was almost complete (~99%) when the starting material was incubated with 1.8 equivalents of donor and 0.15 mg/ml purified NaKfiA for 18 h.

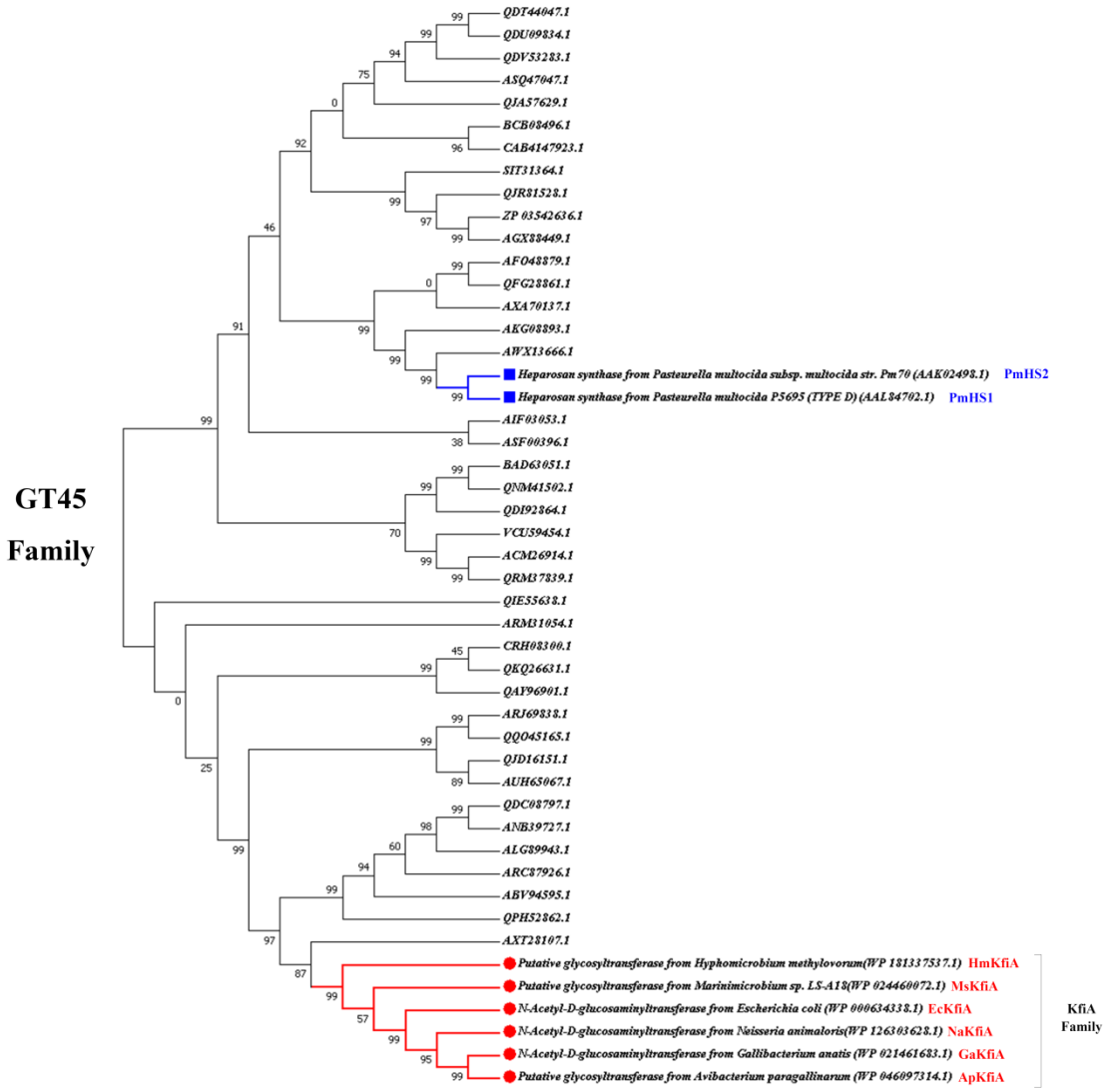

Figure S11. Phylogenetic tree based on GT45 family enzymes. GT45 domain sequences(delete gaps) were derived from CAZy (http://www.cazy.org/), and only one sequence was taken from each species. The ClustalW program was used to analyze amino acid sequences, and the bootstrap consensus tree was constructed using MEGA 7.0 software with the neighbor-joining method and 1000 bootstrap replicates.

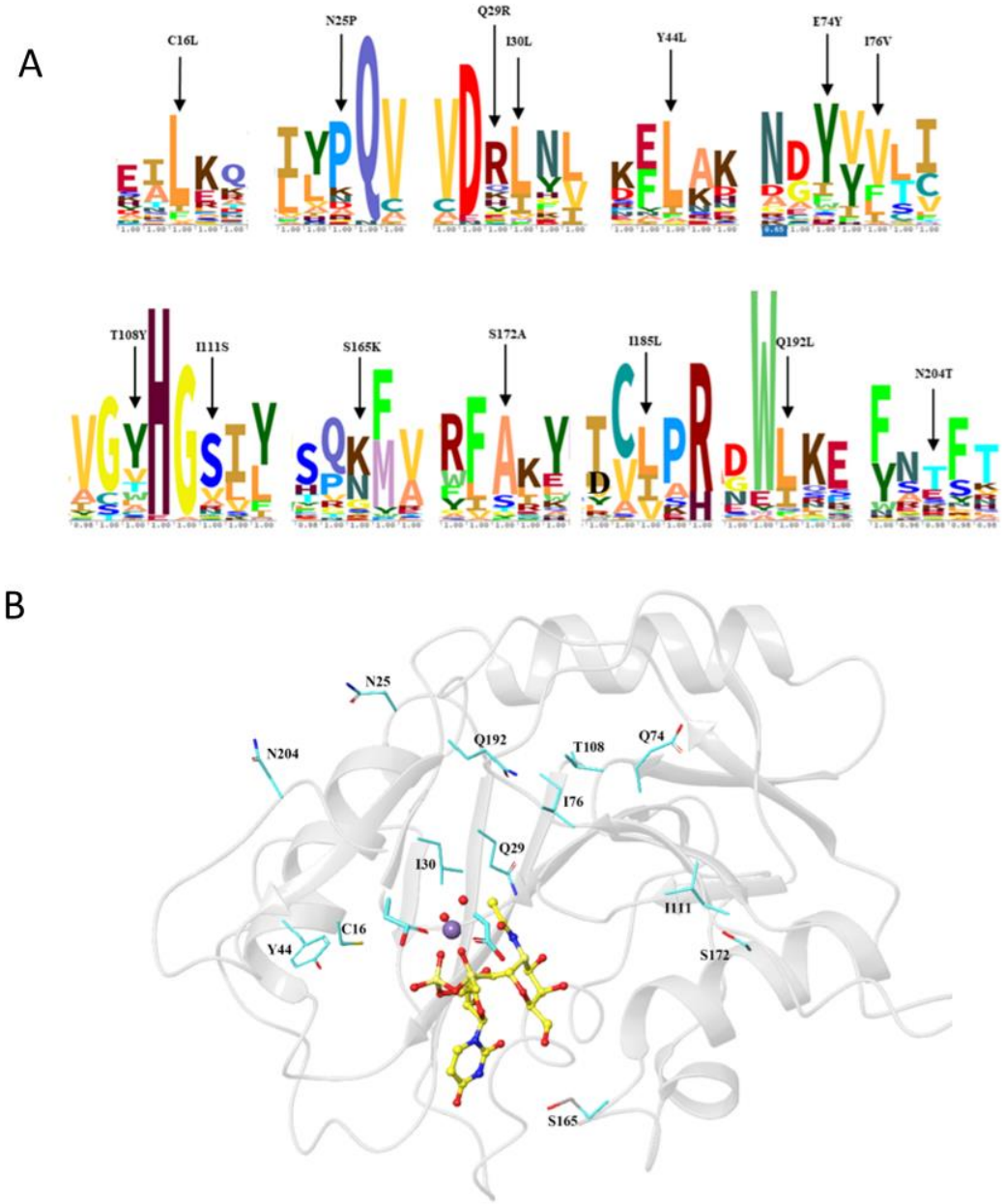

Figure S12. Design of NaKfiA14 point mutations. (A) Weblogo seqlogo of GT45 family enzymes. The amino acid sequence of NaKfiA is given at the bottom. GT45 sequences were derived from CAZy, and only one sequence was taken from each species. The Cluster program was used to compare multiple sequences (https://www.genome.jp/tools-bin/clustalw.), and then visualization used Skylign (http://skylign.org/).The residue preferences at particular sites provide indications for site-directed mutation. (B) The locations of the 14 point mutations made in NaKfiA in this study.

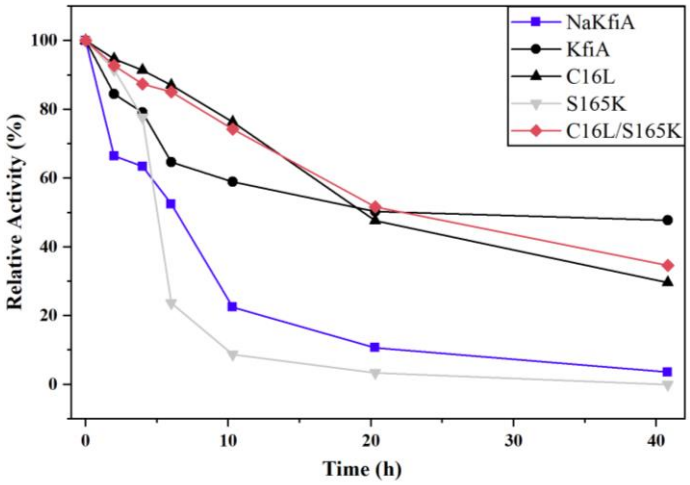

Figure S13. Thermal stability of EcKfiA and NaKfiA and its mutants. Enzymatic stability was determined by incubating enzyme at 37 ℃ for 0, 2, 4, 6, 10, 20, or 40 h. The enzymes were then assayed at 37 ℃ for 1 h and the products analyzed by HPLC.

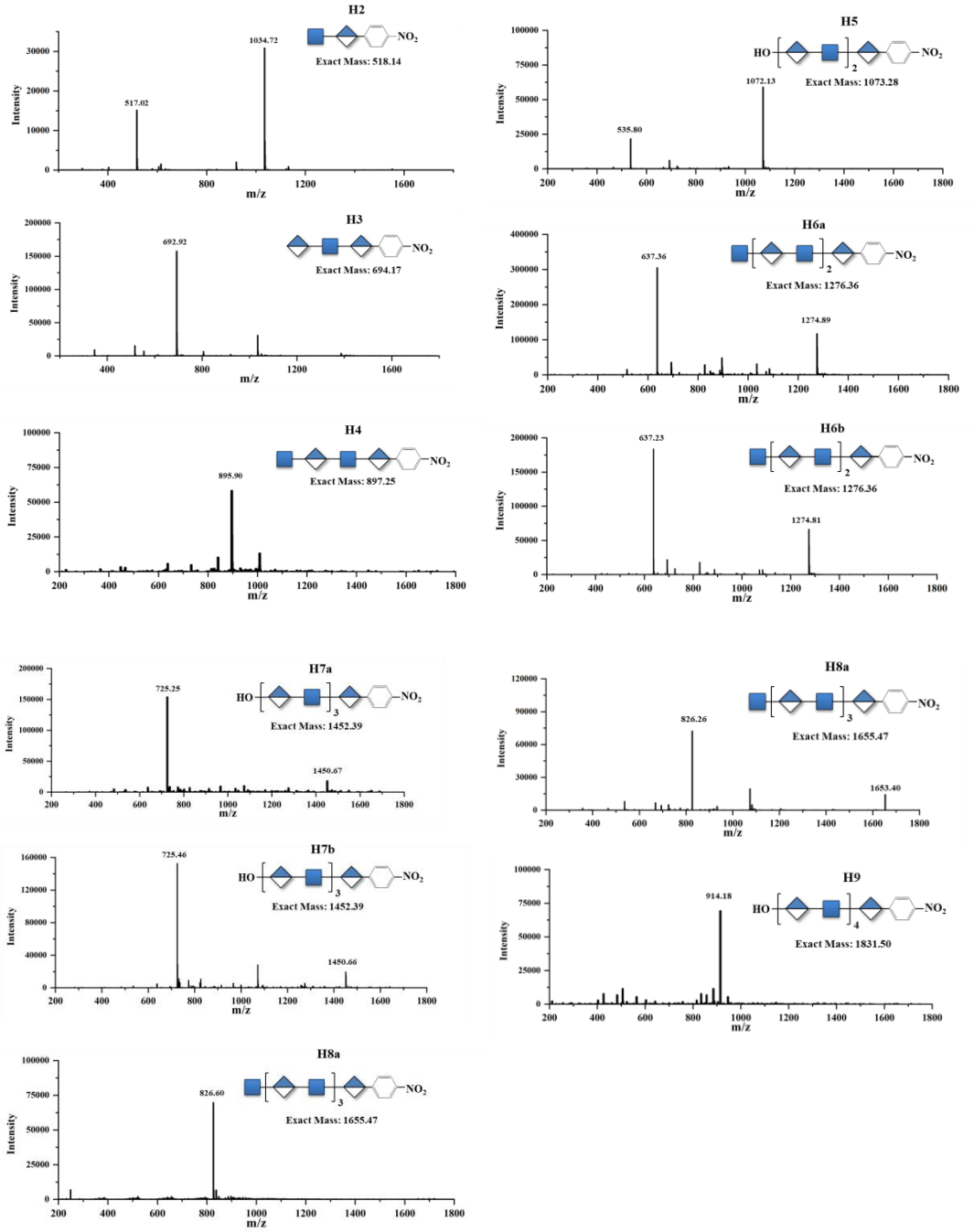

Figure S14. Electrospray ionization mass spectra of H2 to H9. The molecular weights of the products were almost identical to the calculated molecular weights. **H2**, **H4**, **H6**, and **H8** were synthesized using PmHS2 or its mutant. **H3**, **H5**, and **H7** were synthesized using EcKfiA. **H6a** and **H8a** were synthesized using NaKfiA(C16L/S165K). **H7a** and **H9** were synthesized using EcKfiC.