Electronic Supplementary Material (ESI) for ChemComm.

Supporting Information

One-Pot Four-Enzyme Synthesis of Thymidinediphosphate-L-rhamnose

Siqiang Li,[a] Hong Wang,[a] Juncai Ma,[a] Guofeng Gu,[a,*] Zonggang Chen,[a,*] and Zhongwu Guo[b,*]

[a] National Glycoengineering Research Center and Shandong Provincial Key Laboratory of Carbohydrate Chemistry and Glycobiology Shandong University, 27 Shanda Nan Lu, Jinan 250100, China

[b] Department of Chemistry, University of Florida, 214 Leigh Hall, Gainesville, Florida 32611, United States

* Corresponding author E-mails: zguo@chem.ufl.edu; guofenggu@sdu.edu.cn; chenzg@sdu.edu.cn

Table of Contents:

Experimental Section

Materials. *S. pneumonia* serotype 23F, DNA polymerase, BCA protein quantification and bacterial genomic DNA extraction kits, and inorganic pyrophosphatase (YIPP) were purchased from ATCC (700669), Takara, Cwbiotech and Sigma, respectively. Other reagents were obtained from Sangonbiotech unless stated otherwise.

Enzyme Expression and Purification. DNA sequences of complete *cps23FL*, *cps23FN*, *cps23FM*, and *cps23FO* genes were derived from the Gene Bank (*cps23FL*: AAC38757.1; *cps23FM*: AAC38758.1; *cps23FN*: AAC38759.1; and *cps23FO*: AAC38760.1) and employed for the design of primers (Table 1) for polymerase chain reaction (PCR) amplifications. The primers were synthesized by Sangonbiotech. Chromosomal DNA (1 mg/mL) of *S. pneumonia* 23F was extracted from bacterial strain ATCC 700669 with a bacterial genomic DNA extraction kit and used for PCRs. PCRs were performed with a S1000^{TM} Thermal Cycler using 2.5 unit of LA Taq DNA polymerase, 10 μg/ml of template DNA, 0.4 mM of each deoxynucleotide triphosphate (dNTP), and 0.4 μM of corresponding synthetic nucleotide primer. The amplified fragments were digested with the appropriate restriction endonucleases, cloned in plasmid pET-28a, and transferred in *E. coli* BL21 (DE3) competent cell (TIANGEN). The transformants were selected on Luria-Bertani/kanamycin (LB/Kan) plates. Plasmid DNA sequencing was performed by Sangonbiotech Shanghai Co. Ltd.

A desired expression strain was incubated overnight in LB medium containing 100 μg/mL of kanamycin. After growth overnight (37 °C, 200 revolutions/min), the culture was 1:100 diluted with fresh LB medium containing 100 μg/mL of kanamycin typically to a total volume of 1 L. The large scale culture was allowed to grow at 37 ^{o}C (200 revolutions/min) to the midlog phase (OD600 ≈ 0.6), at which point isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. Growth was allowed to continue for another 10 h (200 revolutions/min, 25 °C). The cells were collected through centrifugation (10 min, 8000g/min) and suspended in 100 mL of Tris-HCl buffer (50 mM, pH 7.5) containing 300 mM of NaCl on ice. The cells were lysed via sonication (VCX800, SONICS) on ice using a numerical ultrasonicator at 70% power, 5 s each time with 10 s pause interval, for 15 min in total.

The His-tagged enzymes were purified with a Ni^{+} affinity column, eluted with Tris-HCl buffer (50 mM, pH7.5) containing 0.3 M of NaCl and 0.5 M of imidazole. Fractions containing the enzymes were concentrated with an Amicon Ultra 10-kDa centrifugal filter.

The enzymes were washed twice with Tris-HCl buffer (50 mM, pH 7.5, 10 mL) and then suspended in 2 mL of Tris-HCl buffer (50 mM, pH 7.5) containing 30% glycerol (v/v). Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and Western blot (anti-His-tag rabbit polyclonal antibody) were performed to confirm the purified enzymes. Enzyme concentrations were evaluated with a BCA protein quantitation kit, and the enzymes were stored at -80 °C.

Preliminary enzyme assays. The enzymatic activities of Cps23FL, Cps23FN, Cps23FM, and Cps23FO were assessed by previously reported methods (D. B. James and J. Yother, *J Bacteriol*, 2012, **194**, 6479-6489; D. B. James, K. Gupta, J. R. Hauser and J. Yother, *J Bacteriol*, 2013, **195**, 5469-5478) with minor modifications. For the assay of Cps23FL and Cps23FN, a 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)-buffered system (30 mM, pH 7.5, 500 μL) containing dTTP (3 mM), Glc-1-P (3 mM), $MgCl_2$ (10 mM), YIPP (2U), NAD^+ (0.1 mM), Cps23FL (0.5 μM) and/or Cps23FN (0.5 μM) was incubated at 37 °C for 30 min. Then, a 25 μL aliquot was removed from each reaction mixture and placed in a 100 mM NaOH solution (775 μL). The solution was incubated at rt for 15 min before measuring the absorbance at 320 nm wavelength. For the assay of Cps23FM and Cps23FO, a HEPES-buffered system (500 μL) containing the filtrate of above reactions, 6 mM of NADH, 0.5 μM of Cps23FM, and 0.5 μM of Cps23FO was was incubated at 37 °C, while the reaction progress was analyzed in 2 min intervals by measuring the absorbance at 340 nm wavelength. Furthermore, the reaction products were isolated by HPLC using a DionexCarboPac™ PA-100 (4 × 250 mm) column with 0 to 1 M ammonium acetate solution as eluent and analyzed by ESI-HRMS and NMR studies (data presented below).

Evaluation of the reaction of Cps23FL and its enzyme kinetics. The enzymatic reaction of Cps23FL was studied by a previously reported colorimetric assay (J. Bae, K. H. Kim, D. Kim, Y. Choi, J. S. Kim, S. Koh, S. I. Hong and D. S. Lee, *Chembiochem*, 2005, **6**, 1963-1966). The typical procedure was to add to each well of a 96-well microtiter plate 50 μL of a certain buffer containing glycerol (10%, v/v), dithiothreitol (DTT, 1 mM), and YIPP (2 U/mL), as well as a certain concentration of an ion, dTTP and D-Glc-1-P. Then, Cps23FL was added at 37 °C to initiate the reaction, which after some time of incubation was quenched by adding a 0.7 N HCl solution (50 μL) containing malachite green reagent (0.03%, w/v), ammonium molybdate (0.2%, w/v), and Triton X-100 (0.05%, v/v). The mixture was incubated at 37 °C for 5 min, and its absorbance at 630 nm wavelength was measured with a

microplate spectrophotometer. Negative controls were conducted under the same conditions except for additing heat-deactived Cps23FL in replacement of the native enzyme.

In the study of pH influence on the enzymatic reaction, the buffers used were sodium acetate and acetic acid at pH 4～6 and Tris-HCl at pH 7～9, and the concentrations of ion, dTTP, D-Glc-1-P, and Cps23FL were 5 mM of $MgCl_2$, 0.2 mM, 1 mM, and 2.04 μg/mL, respectively. The reaction was carried out at 37 °C for 3 min before quenching. In the study of temperature influence on the enzymatic reaction, the conditions were Tris-HCl buffer (pH 7.5) with 5 mM of $MgCl_2$, 0.2 mM of dTTP, 1 mM of D-Glc-1-P, and 2.04 μg/mL of Cps23FL, and the mixture was incubated at 16, 25, 30, 37, 42, and 60 °C, respectively, for 3 min before quenching. To evaluate the influence of metal ions on the activity of Cps23FL, the reaction was carried out in Tris-HCl buffer (pH 7.5) containing 0.2 mM of dTTP, 1 mM of D-Glc-1-P, 2.04 μg/mL of Cps23FL, and 10 mM of EDTA or various cations, including Na^+, K^+, Zn^{2+}, Mn^{2+}, Ca^{2+}, Mg^{2+}, Cu^{2+}, Fe^{2+} and Ni^{2+}. The reaction mixture was incubated at 37 °C for 3 min before quenching. To optimize the Mg^{2+} concentration for the enzyme, the reaction was conducted under the same conditions using 0～160 mM of Mg^{2+}.

To analyze the enzyme kinetics of Cps23FL, the reaction was performed under above-optimized conditions, namely, at 37 °C in Tris-HCl buffer (50 mM, pH 7.5) containing 10% of glycerol (v/v), 1 mM of DTT, 2 U/mL of YIPP, 5 mM of $MgCl_2$, and varied D-Glc-1-P and dTTP concentrations. First, different concentrations (0.51, 1.02, 2.04, 4.08, and 8.16 μg/mL) of Cps23FL were tried (incubation for 1～6 min) to establish the appropriate enzyme concentration. Eventually, 2.04 μg/mL of Cps23FL was selected for the study. The concentration of dTTP varied in 0.01～0.08 mM range in the presence of saturated Glc-1-P (1 mM), and the concentration of Glc-1-P varied in 0.05～0.40 mM range with saturated dTTP (1 mM). The reaction was quenched at 3, 6, and 9 min, respectively. The data obtained were used to calculate the initial reaction velocities and then determine the Michaelis constant (Km) and maximal velocity (Vmax) values by GraphPad Prism 6.04. Each experiment was repeated three times, and the results are averaged.

Evaluation of the reaction of Cps23FN and its enzyme kinetics. These studies were carried out along the same line as described above. Accordingly, the reactions were carried out in 300 μl of HEPES buffer (30 mM) at a certain pH, containing 10% of glycerol, 0.01 mM of NAD^+, and a certain concentration of dTDP-Glc and Cps23FN. The reaction mixture was incubated at a temperature for a

fixed period and then terminated by adding 30 μL of NaOH (1 M) at 30 °C. About 20 min later, the OD320 value of the reaction mixture was measured by a microplate spectrophotometer. Negative controls were performed in parallel using heat-deactived Cps23FN under the same conditions.

The influence of pH on the reaction was studied in HEPES buffer with the pH ranging from 6.0 to 9.0 that contained 0.8 mM of dTDP-Glc and 0.01 mg/mL of Cps23FN at 37 °C for 15 min. The influence of temperature, at 16, 25, 30, 37, 41 and 45 °C respectively, was studied at pH 7.5 for 10 min of reaction.

To study the enzyme kinetics of Cps23FN, the reaction was carried out under above optimized conditions in HEPES buffer at pH 7.5 and 37 °C. First, the suitable enzyme concentration for this study was established by analysis of the reaction at various concentrations of Cps23FN (2.5, 5.0, 10, 20 and 40 μg/mL) for 2.5, 5, 10, 15, and 20 min. Then, the initial reaction velocities were derived from reactions carried out with 10 μg/mL of Cps23FN and 0.1～1.0 mM of dTDP-Glc at 37 °C for 5, 10, and 15 min, respectively, which were employed to determine the Km and Vmax values of Cps23FN with GraphPad Prism 6.04. Each parameter was measured in triplicate.

Enzymatic synthesis of dTDP-Glc. To a HEPES buffered solution (30 mM, pH 7.5, 5 mL) of Glc-1-P (3 mM), dTTP (5 mM), $MgCl_2$ (5 mM), and YIPP (2 U/mL) was added Cps23FL (2 μM). After the reaction mixture was incubated at 37 °C for 30 min, the enzyme was removed by filtration through an Amicon Ultra 10-kDa centrifugal filter. The product was purified by Q anion exchange chromatography (GE Q Sepharose Fast Flow, 0-0.6 M ammonium bicarbonate) and lyophilized to provide dTDP-Glc (7.26mg, 86%) as a white solid. ^{1}H-NMR (600 MHz, D_2O): δ 7.55 (d, 1 H, *J* = 1.2 Hz, H-6''), 6.14 (dd, 1 H, *J* = 7.2, 6.6 Hz, H-1′), 5.42 (dd, 1 H, *J* = 7.2, 3.6 Hz, H-1), 4.41 (dt, 1 H, *J* = 6.0, 3.0 Hz, H-3′), 4.02-3.97 (m, 1 H, H-4′, H-5a,b′), 3.71 (ddd, 1 H, J = 9.6, 4.8, 2.4 Hz, H-5), 3.66 (d, 1 H, *J* = 12.6, 2.4 Hz, H-6a), 3.59 (dd, 1 H, J = 12.6, 4.8 Hz, H-6b), 3.58 (t, 1 H, *J* = 9.6 Hz, H-3), 3.34 (dt, 1 H, *J* = 9.6, 3.6 Hz, H-2), 3.27 (t, 1 H, *J* = 9.6 Hz, H-4), 2.24-2.13 (m, 2 H, H-2′), 1.75 (s, 3 H, -C*H*$_3$); ^{13}C-NMR (150 MHz, D_2O): *δ* 166.5 (C-4''), 151.6 (C-2''), 137.2 (C-6''), 111.6 (C-5''), 95.4 (d, *J* = 6.6 Hz, C-1), 85.2 (d, *J* = 9.0 Hz, C-4′), 84.9 (C-1′), 72.73 (C-3), 72.67 (C-5), 71.5 (d, *J* = 6.6 Hz, C-2), 70.9 (C-3′), 69.0 (C-4), 65.4 (d, *J* = 5.4 Hz, C-5′), 60.2 (C-6), 38.5 (C-2′), 11.5 (-*C*H_3); ^{31}P-NMR (242 MHz,

D_2O): δ -11.4 (d, *J* = 20.6 Hz), -13.0 (d, *J* = 20.6 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{25}N_2O_{16}P_2$ $[M-H]^{-1}$ 563.0679, found 563.0706.

Enzymatic synthesis of dT4k6dG. To a HEPES buffered solution (50 mM, pH 7.5, 30 mL) of dTDP-Glc (0.8 mM) and NAD^+ (0.01 mM) was added Cps23FN (0.24 μM). After the reaction mixture was incubated at 37 ^oC for 30 min, the enzyme was removed by filtration through an Amicon Ultra 10-kDa centrifugal filter. The product was purfied by HPLC (Dionex CarboPacTM PA-100 Semipreparative Column 22 × 250mm, 0～1 M ammonium acetate). Fractions containing the product were pooled and desalted on a G10 column (1.2 × 120 cm) to give dT4k6dG (9.2 mg, 71%) as a white solid after lyophilization. ^{1}H-NMR (600 MHz, D_2O): *δ* 7.57 (s, 1 H, H-6''), 6.17 (t, 1 H, *J* = 6.9 Hz, H-1'), 5.36 (dd, 1 H, *J* = 6.6, 3.6 Hz, H-1), 4.43–4.46 (m, 1 H, H-3'), 4.02–3.98 (m, 3 H, H-4', H-5a,b'), 3.92 (q, 1 H, *J* = 6.5 Hz, H-5), 3.60 (d, 1 H, *J* = 10.2 Hz, H-3), 3.46–3.41 (dt, 1 H, *J* = 10.2, 3.6, 3.6 Hz, H-2), 2.20–2.17 (m, 2 H, H-2'), 1.75 (s, 3 H, -C*H*$_3$), 1.04 (d, 3 H, *J* = 6.5 Hz, H-6); ^{13}C-NMR (150 MHz, D_2O): *δ* 166.5 (C=O), 151.6 (C=O), 137.22 (=*C*H), 111.7 (C=O), 95.2 (d, *J* = 6.8 Hz, C-1), 93.6 (=*C*-Me), 85.2 (d, *J* = 9.1 Hz, C-4'), 84.9 (C-1'), 73.0 (C-3), 70.9 (C-3'), 70.5 (d, *J* = 8.6 Hz, C-2), 69.5 (C-5), 65.3 (d, *J* = 5.3 Hz, C-5'), 38.5 (C-2'), 11.5 (-*C*H$_3$), 11.2 (C-6); ^{31}P-NMR (242 MHz, D_2O): *δ* -11.4, -13.00; ESI-TOF HRMS m/z calcd for $C_{16}H_{22}N_2O_{15}P_2$ $[M-H]^{-1}$ 545.0574, found 545.0546.

One-pot enzymatic synthesis of dTDP-Rha. To a HEPES buffered solution (30 mM, pH 7.5, 120 mL) of Glc-1-P (3 mM), dTTP (5 mM), $MgCl_2$ (5 mM), NAD^+ (0.1 mM), and NADH (6 mM) was added Cps23FL (2 μM). After the reaction mixture was incubated at 37 ^oC for 15 min, Cps23FN (2 μM), Cps23FM (2 μM), and Cps23FO (2 μM) were added within 30 min. The reaction was incubated for another 30 min, and the enzymes were removed by filtration through an Amicon Ultra 10-kDa centrifugal filter. The product was purified by HPLC (Dionex CarboPacTM PA-100 Semipreparative Column 22 × 250 mm, 0-1 M ammonium acetate). Fractions containing the product were pooled and then desalted on a G10 column (1.2 × 120 cm) to produce dTDP-Rha (124.1 mg, 63%) as a white solid after lyophilization. ^{1}H-NMR (600 MHz, D_2O): *δ* 7.57 (d, 1 H, *J* = 1.2 Hz, H-6''), 6.18 (dd, 1 H, *J* = 7.8, 6.6 Hz, H-1'), 5.05 (br d, 1 H, *J* = 9.0 Hz, H-1), 4.47–4.43 (dt, 1 H, *J* = 6.0 , 3.0 Hz, H-3), 4.03–3.98 (m, 3 H, H-4', H-5a,b'), 3.92 (d, 1 H, *J* = 3.6 Hz, H-2), 3.47 (dd, 1 H, *J* = 9.6, 3.6 Hz, H-3), 3.30–3.25 (m, 1 H, H-5), 3.20 (t, 1 H, *J* = 9.6 Hz, H-4), 2.26–2.13 (m, 2 H, H-2'),

1.75 (br s, 3 H, -CH_3), 1.14 (d, 3 H, J = 6.0 Hz, H-6); ^{13}C-NMR (150 MHz, D_2O): δ 166.5 (C=O), 151.7 (C=O), 137.3 (2 C, C=O, =CH), 111.7 (=*C*-Me), 95.45 (d, J = 4.0 Hz, C-1), 85.2 (d, J = 9.0 Hz, C-4'), 84.8 (C-1'), 72.6 (C-5), 72.0 (C-3), 70.9 (C-4), 70.7 (d, J = 5.8 Hz, C-2), 65.4 (d, J = 5.8 Hz, C-5'), 38.5 (C-2'), 16.6 (C-6), 11.6 (-*C*H_3); ^{31}P-NMR (242 MHz, D_2O): δ -11.7 (d, J = 20.6 Hz), -13.8 (d, J = 20.6 Hz); ESI-TOF HRMS m/z calcd for $C_{16}H_{25}N_2O_{15}P_2$ $[M-H]^{-1}$ 547.0736, found 547.0711. Quantitive analyses of dTDP-Rha and dTDP-Glc were performed by UV/Vis spectroscopic measurement of thymidine absorbances at 260 nm wavelength with an extinction coefficient of 7.4×10^3 M^{-1} cm^{-1} using Beer's law as previously reported (D. B. James and J. Yother, *J Bacteriol*, 2012, **194**, 6479-6489; D. B. James, K. Gupta, J. R. Hauser and J. Yother, *J Bacteriol*, 2013, **195**, 5469-5478).

Table S1. The primers used to clone *Cps23FL*, *Cps23FN*, *Cps23FM*, and *Cps23FO* genes

Enzyme	Direction	Primer sequence (5'→3')
P23FL05	Forward	GGGAATTCCATATGAAAGGTATTATTCTAGCAGGTG
P23FL03	Reverse	CCGCTCGAGCTATGCTTCTCCAATCAAACGGAGCAG
P23FN05	Forward	GGGAATTCCATATGACTGAATACAAAAATATTATCGT
P23FN03	Reverse	CCGCTCGAGTTATACTGTAATAATCTCCTGAGTTTTA
P23FM05	Forward	GGGAATTCCATATGACAGATAATTTTTTCGGTAAGAC
P23FM03	Reverse	ATAAGAATGCGGCCGCTTACAAATCTTCTTTTTTCAAAGGTT
P23FO05	Forward	GGGAATTCCATATGATGATTTTAATTACAGGGGCAAATG
P23FO03	Reverse	CCGCTCGAGTTATCTCACTTCTTGTTTGTAAAAT

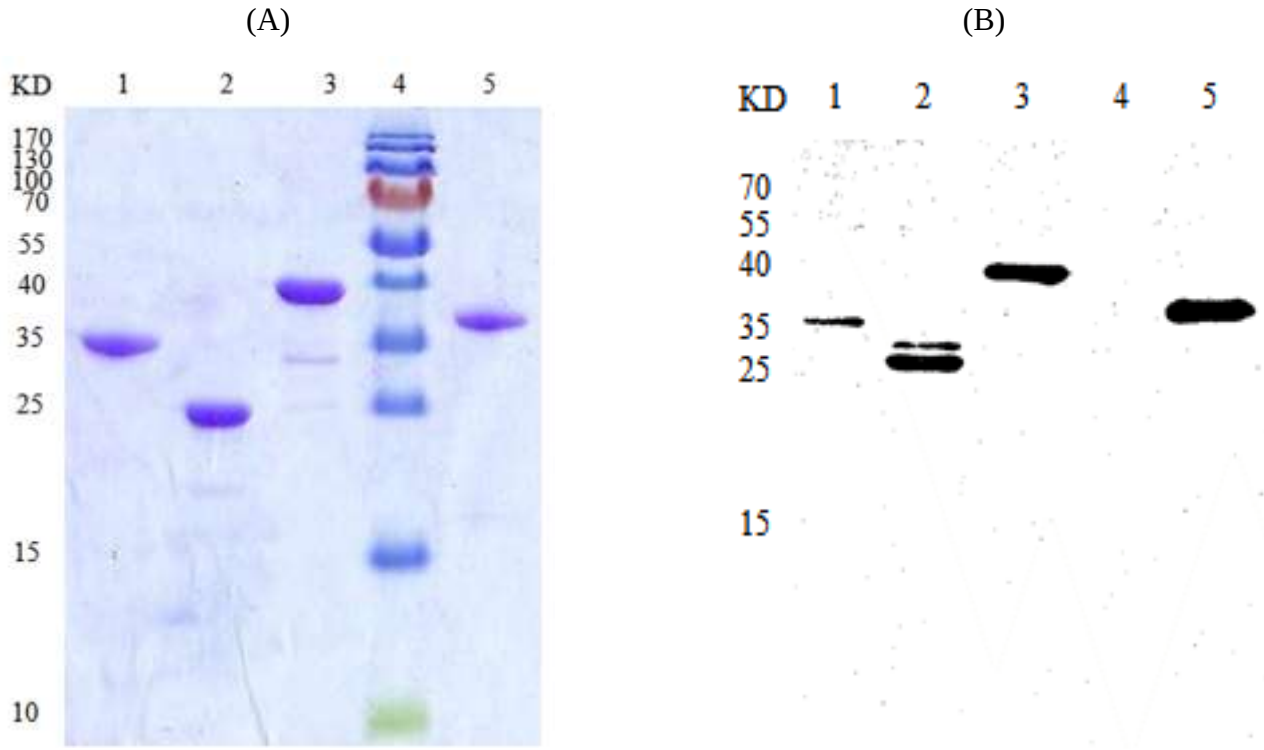


Figure S1. (A) SDS-PAGE results of purified Cps23FL (lane 1), Cps23FM (lane 2), Cps23FN (lane 3) and Cps23FO (lane 5), and molecular marker (lane 4); and (B) Western blot results of purified Cps23FL (lane 1), Cps23FM (lane 2), Cps23FN (lane 3), Cps23FO (lane 5), and molecular marker (lane 4), stained with anti-His-tag rabbit polyclonal antibody.

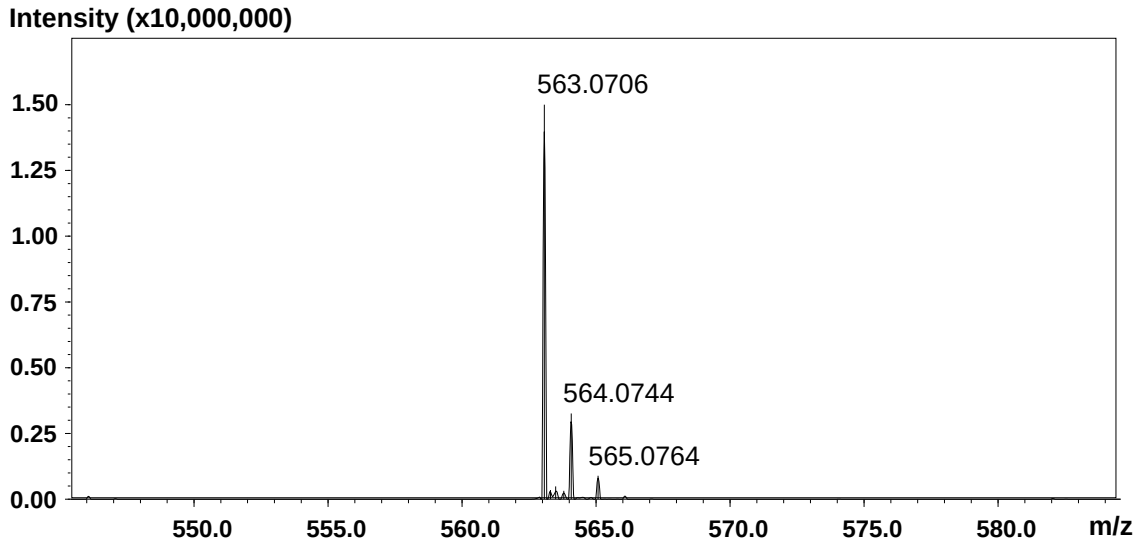
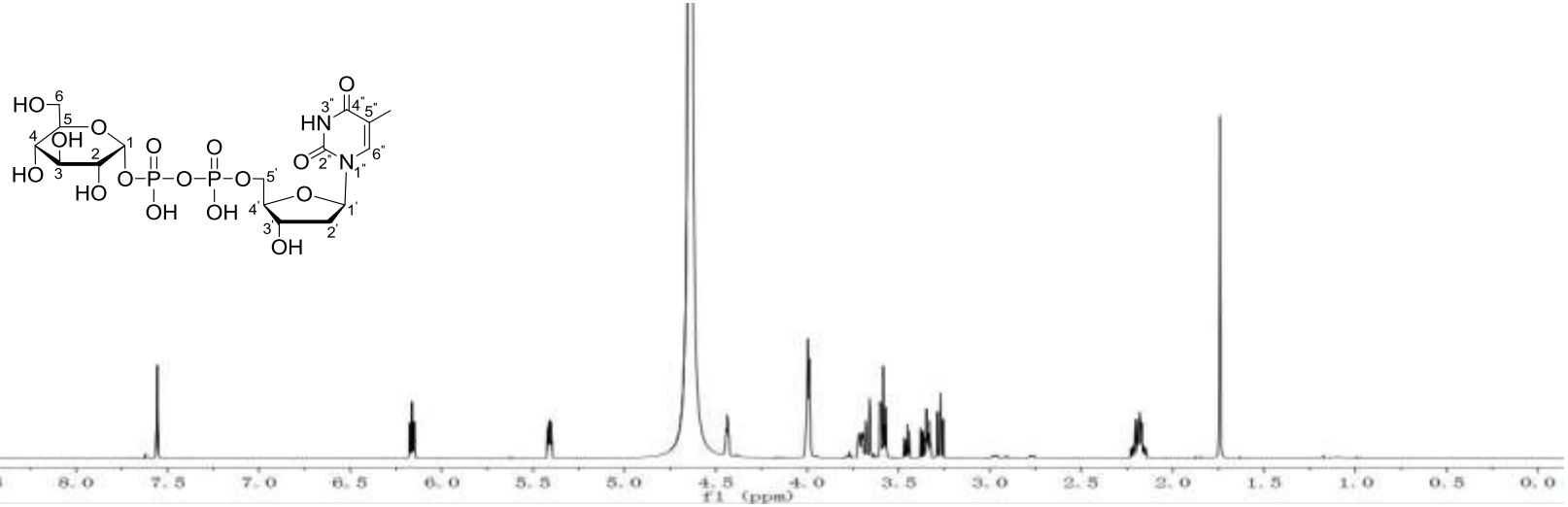
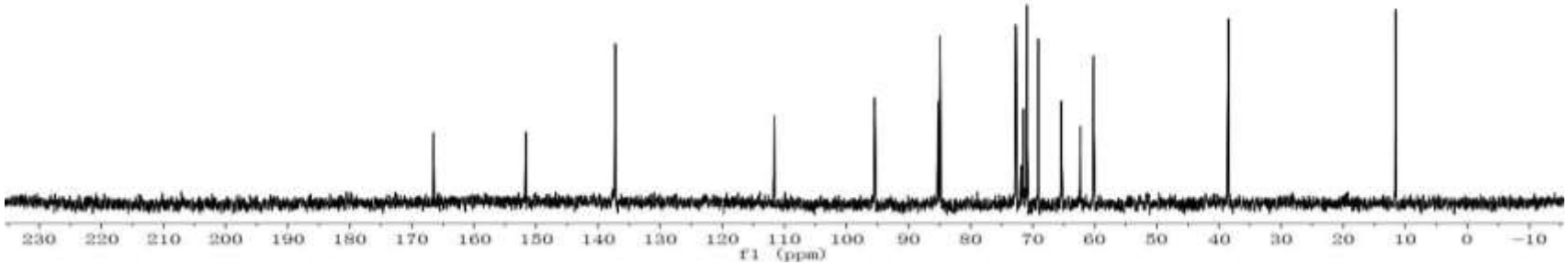


Figure S2. ESI-HRMS spectrum of the reaction product of enzyme Cps23FL

(A)


8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

(B)


230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 −10
f1 (ppm)

(C)

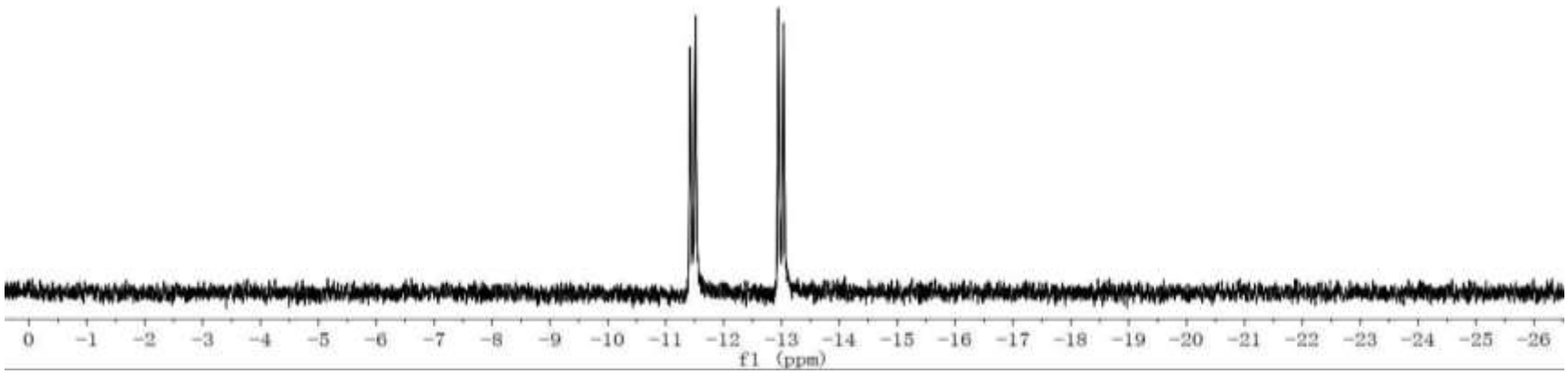


Figure S3. NMR spectra of dTDP-Glc in D_2O: (A) ^{1}H-NMR (600 MHz), (B) ^{13}C-NMR (125 MHz), and (C) ^{31}P-NMR (242 MHz)

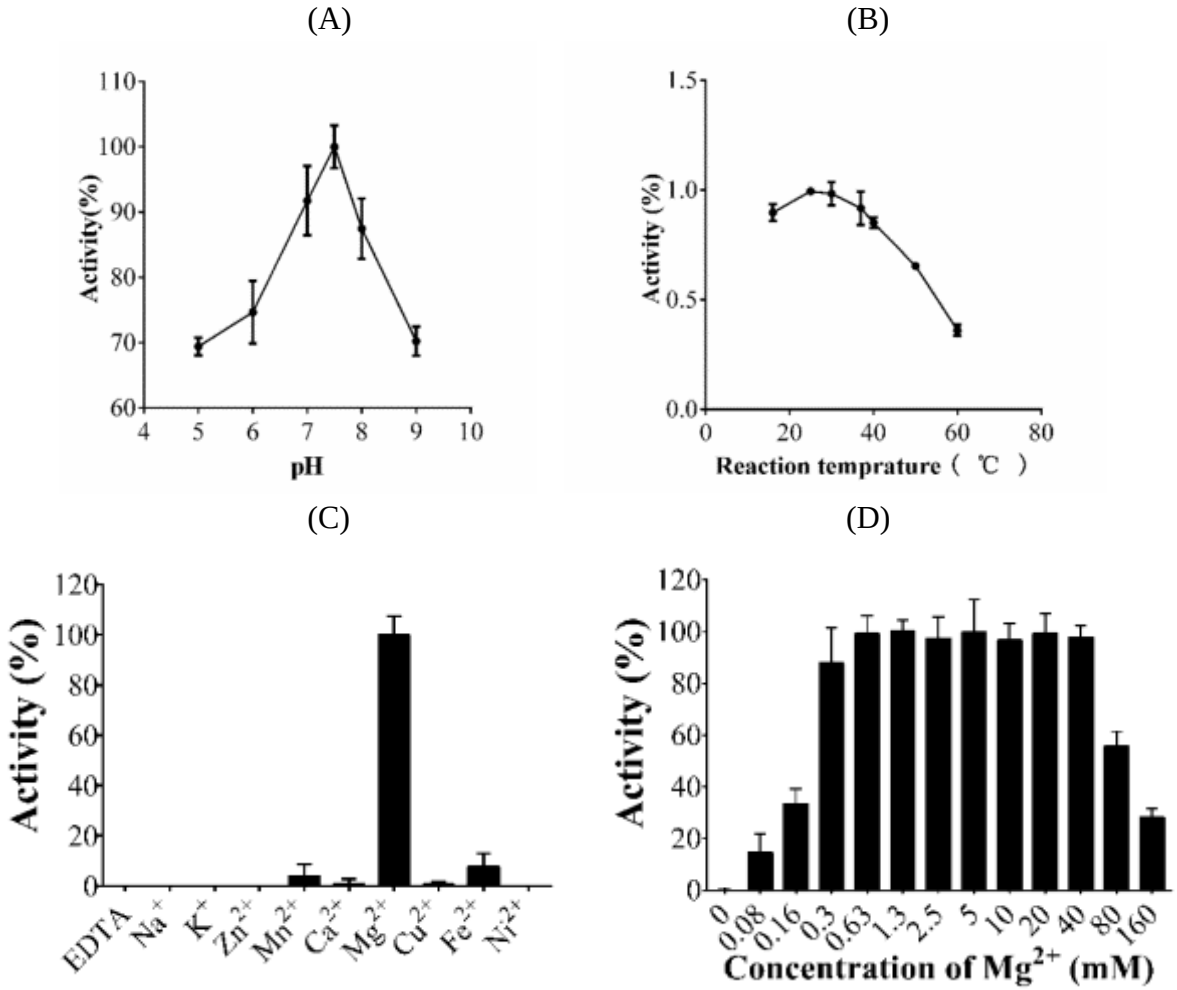


Figure S4. The influence of pH (A), temperature (B), cation (C), and Mg^{2+} concentration (D) on the reaction between Glc-1-P and dTTP catalyzed by Cps23FL. Error bar represents the standard deviation of three experiments

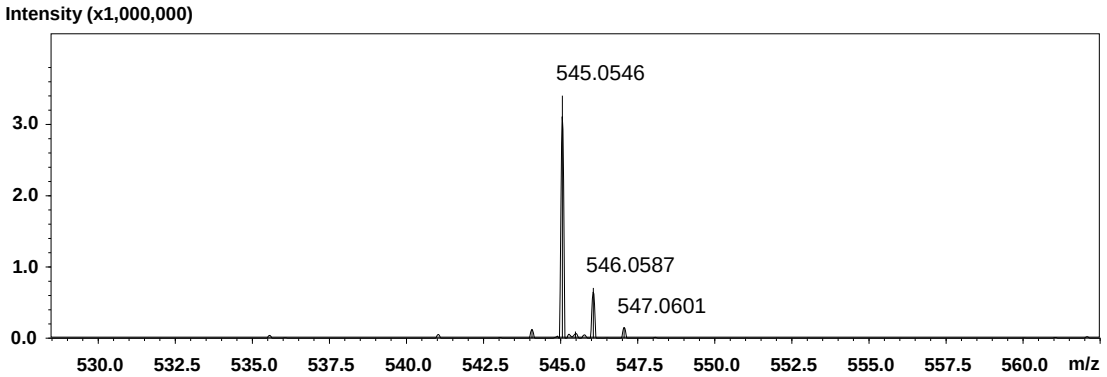
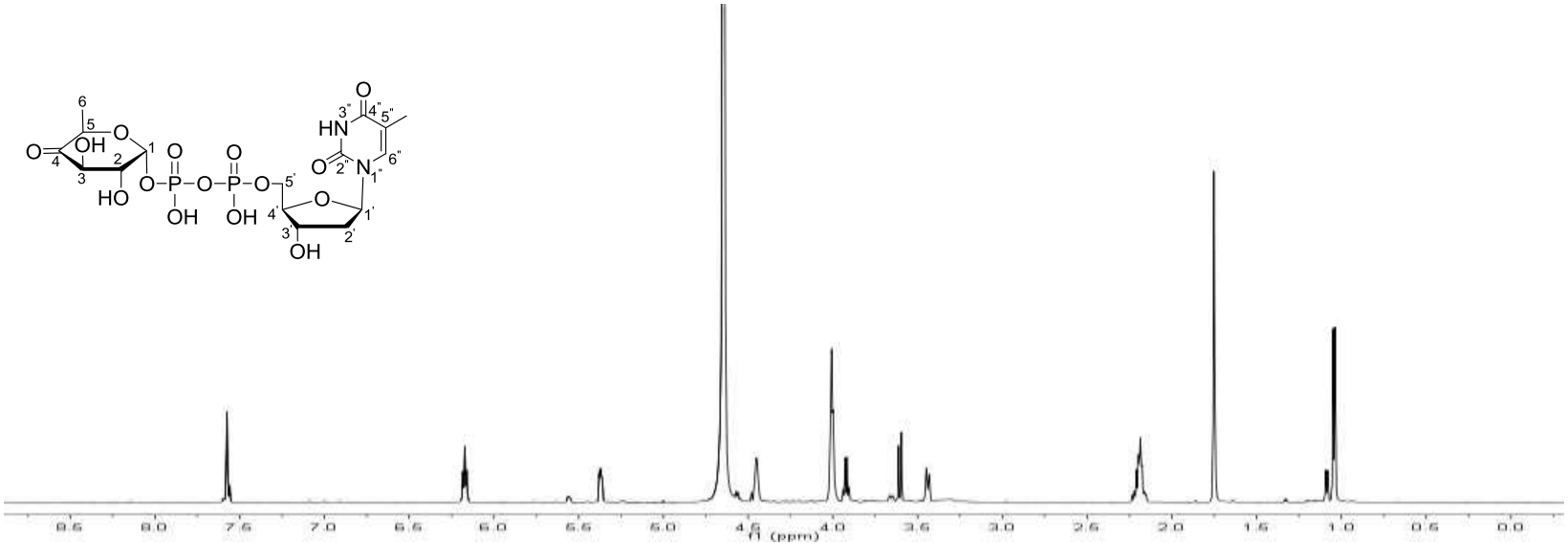
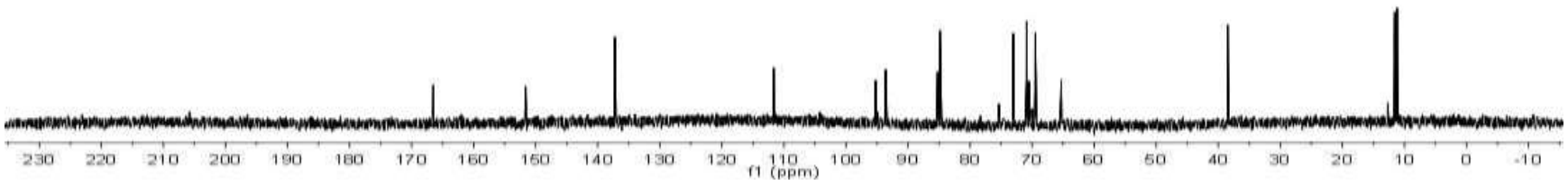


Figure S5. ESI-HRMS spectrum of the reaction product of enzyme Cps23FN

(A)



(B)



(C)

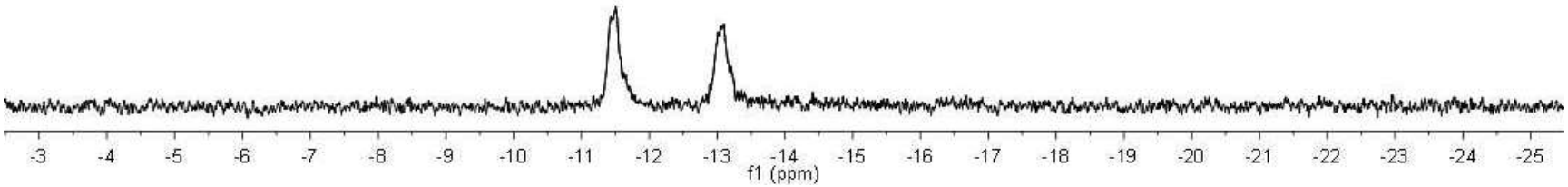


Figure S6. NMR spectra of dT4k6dG in D_2O: (A) ^{1}H-NMR (600 MHz), (B) ^{13}C-NMR (125 MHz), and (C) ^{31}P-NMR (242 MHz)

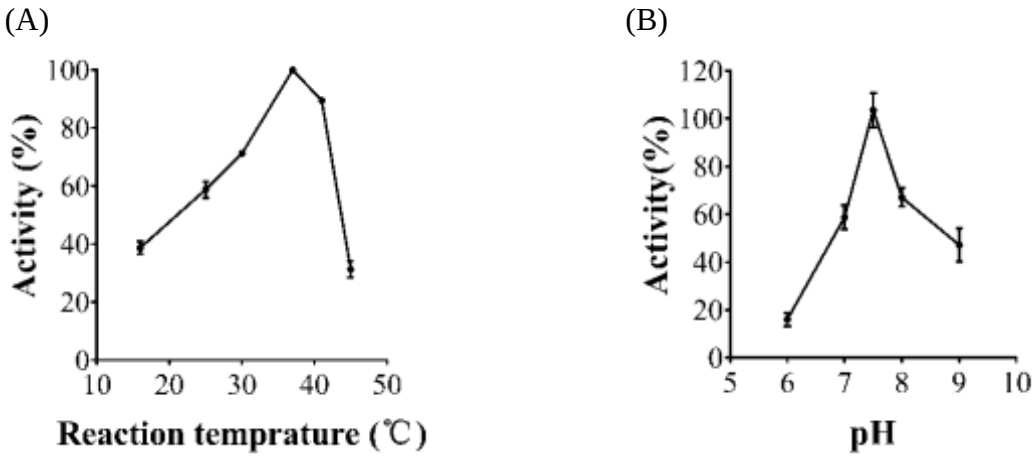


Figure S7. The influence of temperature (A) and pH (B) on the reaction catalyzed by Cps23FN. Error bars represent standard deviations of three parallel experiments.

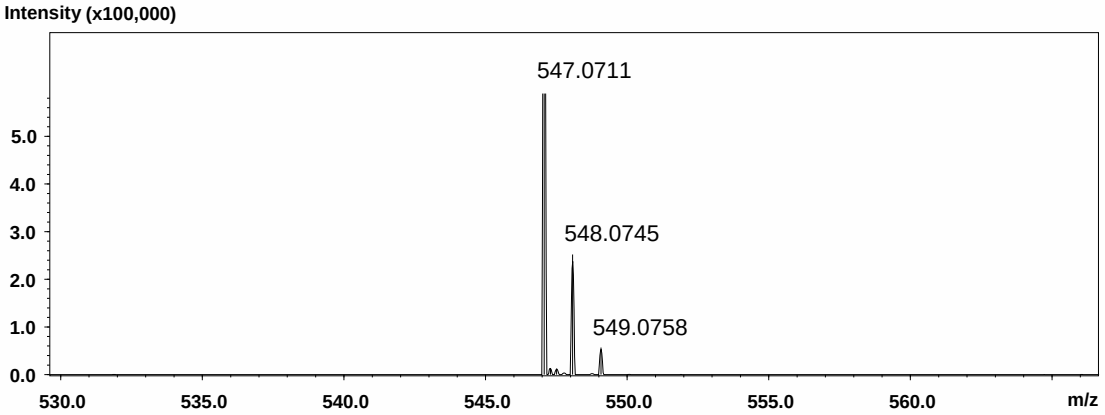
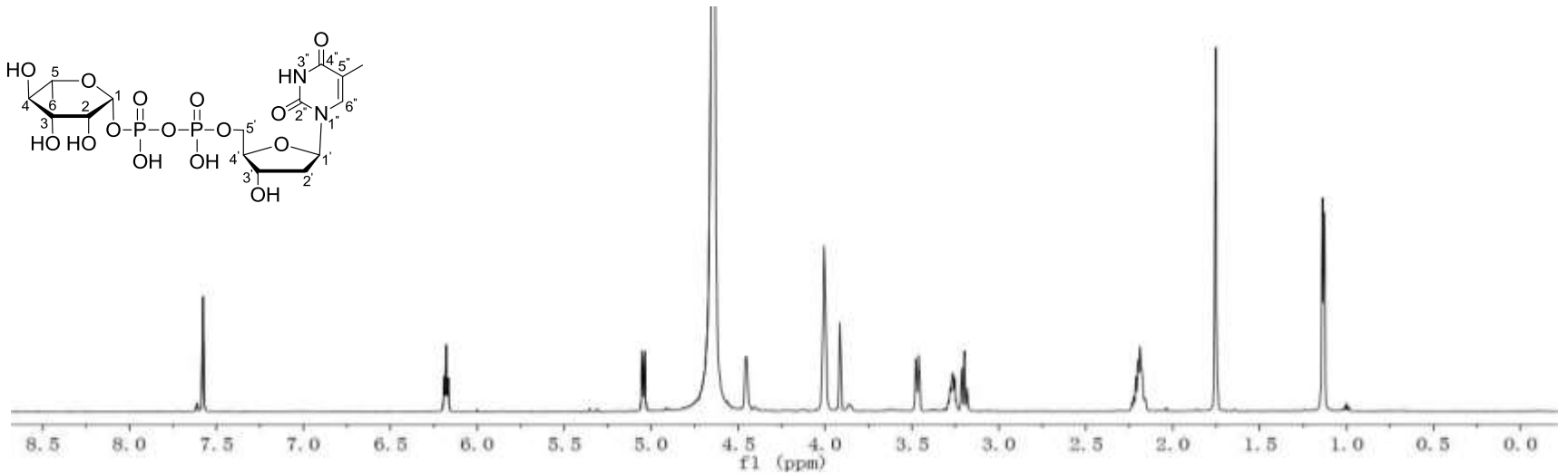
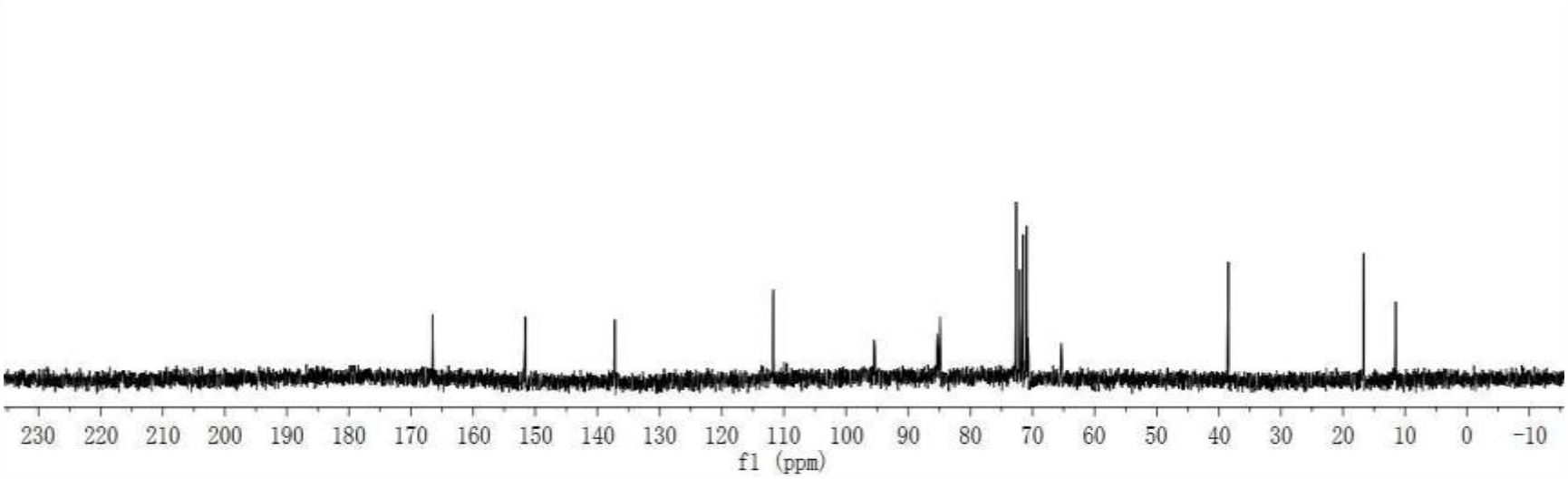


Figure S8. ESI-HRMS spectrum of the reaction product of enzyme Cps23FO

(A)


8.5 8.0 7.5 7.0 6.5 6.0 5.5 5.0 4.5 4.0 3.5 3.0 2.5 2.0 1.5 1.0 0.5 0.0
f1 (ppm)

(B)


230 220 210 200 190 180 170 160 150 140 130 120 110 100 90 80 70 60 50 40 30 20 10 0 -10
f1 (ppm)

(C)

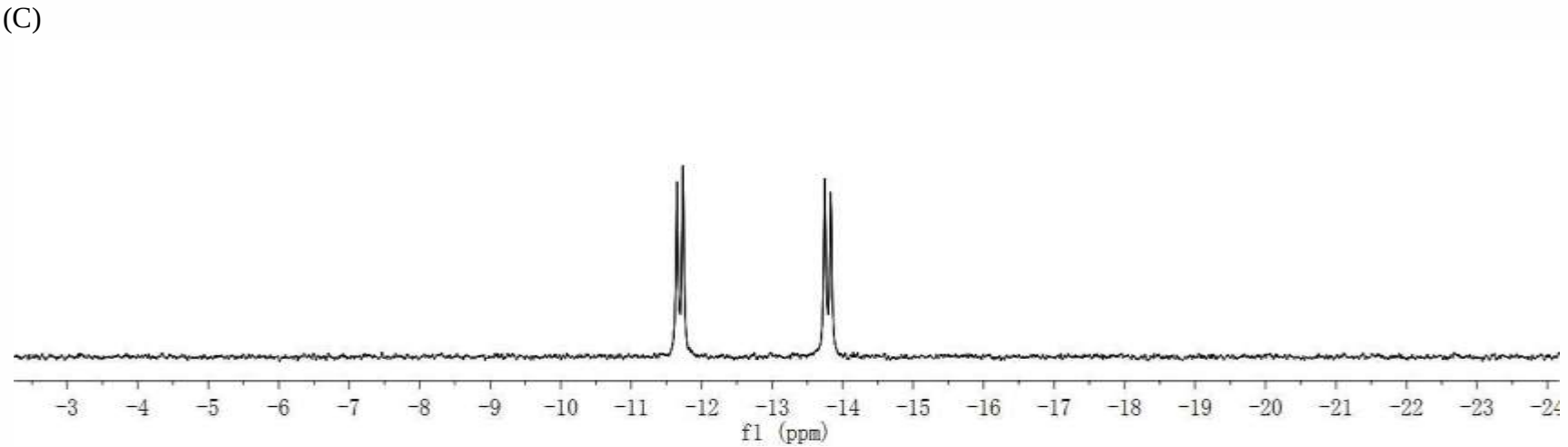


Figure S9. NMR spectra of dTDP-Rha in D_2O: (A) ^{1}H-NMR (600 MHz), (B) ^{13}C-NMR (120 MHz), and (C) ^{31}P-NMR (242 MHz)