

Electronic Supplementary Information

**Transformation of formaldehyde into functional sugars via multi-enzyme
stepwise cascade catalysis**

Jiangang Yang^{a§}, Shangshang Sun^{ba§}, Yan Men^a, Yan Zeng^a, Yueming Zhu^a, Yuanxia
Sun^{a*}, Yanhe Ma^a

^a National Engineering Laboratory for Industrial Enzymes, Tianjin Institute of
Industrial Biotechnology, Chinese Academy of Sciences,
Tianjin 300308, China.

E-mail: sun_yx@tib.cas.cn

^b University of Chinese Academy of Sciences, Beijing, China

[§] These authors contributed equally to this work.

* Corresponding author: E-mail: sun_yx@tib.cas.cn; Tel.: +86 22 84861960; fax:
+86 22 84861961.

Experimental Section

Bacterial Strains, Plasmids and Materials

Compounds FALD, GALD, DHA, ampicillin and isopropyl- β -Dthiogalactopyranoside (IPTG) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Strains, plasmids, and primers used in this study were shown in Table S1 and S2.

Vector construction

Full-length gene of FLS as reported in Siegel's study was synthesised by GeneScript (Nanjing, China). The *fsa* gene from *E. coli* was amplified using primer sets FSA-1(*Nde*I)/FSA-2(*Hind*III). The resulting fragment was inserted into plasmid pET-21a (+) at *Nde*I and *Hind*III sites to obtain pET21-FSA. Mutant FSA^{A129S} was constructed with the aid of the QuikChange site-directed mutagenesis kit (Stratagene), the oligonucleotides listed in the Table S2 and with the plasmid pET21-FSA as template. DNA sequencing analysis confirmed the expected mutations in the gene sequence.

Recombinant proteins expression and purification

E. coli BL21(DE3) strains harboring expression plasmids were cultured at 37 °C in 1 L LB medium containing 100 mg/L ampicillin to an optical density OD₆₀₀ of 0.6. 0.5 mM IPTG was added into the culture to induce protein expression and the temperature was adjusted to 16 °C to avoid inclusion body formation. After incubation for an additional 20 h, cells were harvested, washed twice and suspended in 50 mM triethanolamine (TEA) (pH 7.5) buffer. The suspension cells were then

lysed by sonication and centrifuged at 14,000×g and 4 °C for 10 min. Clear supernatant was collected and loaded onto a Ni²⁺-NTA-agarose column pre-equilibrated with binding buffer (50 mM TEA buffer, 300 mM NaCl, 20 mM imidazole, pH 7.5). The retained proteins were recovered with elution buffer (50 mM TEA buffer, 300 mM NaCl, 300 mM imidazole, pH 7.5). The eluted fraction containing purified protein was dialyzed and stored at -20 °C.

Enzyme activity assays.

FLS activity: The activity was assayed in a reaction mixture (200 µl) containing FLS (0.8 mg), 25 mM TEA buffer (pH 7.0), and 25 mM formaldehyde (FALD). After the conversion at 30°C for 30 min, the product was determined via high-performance liquid chromatography (HPLC) as described below. One unit of enzyme activity was defined as the enzyme amount catalyzing the formation of 1 µmol product per min.

FSA activity: The FSA activity was measured in a reaction mixture (200 µl) containing FSA (0.1 mg), 25 mM FALD, 25 mM glycolaldehyde (GALD) and 25 mM TEA buffer (pH 7.0). The reaction mixture was incubated at 30°C for 10 min. The concentration of L-glyceraldehyde (**3a**) produced from FALD and GALD was determined via HPLC. One unit of enzyme activity was defined as the enzyme amount catalyzing the formation of 1 µmol product per min.

FSA^{A129S} activity for L-erythrulose (4a**) synthesis:** The activity of FSA^{A129S} was determined in a reaction mixture (200 µl) containing FSA^{A129S} (0.1 mg), 25 mM FALD, 25 mM dihydroxyacetone (DHA) and 25 mM TEA buffer (pH 7.0). The reaction mixture was incubated at 30°C for 10 min. The concentration of **4a** produced

from FALD and DHA was determined via HPLC. One unit of enzyme activity was defined as the enzyme amount catalyzing the formation of 1 μmol product per min.

FSA^{A129S} activity for L-sorbose (**6a**) synthesis: The activity of FSA^{A129S} was determined in a reaction mixture (200 μl) containing FSA^{A129S} (0.1 mg), 25 mM **3a**, 25 mM (DHA) and 25 mM TEA buffer (pH 7.0). The reaction mixture was incubated at 30°C for 10 min. The concentration of **6a** produced from **3a** and DHA was determined via HPLC. One unit of enzyme activity was defined as the enzyme amount catalyzing the formation of 1 μmol product per min.

In vitro cascade reaction

To synthesise C3 precursor **3a**, one-pot reaction mixture (200 μl) containing 50 mM TEA (pH 7.5), 25 mM FALD, FLS and FSA was carried out at 25 °C and 120 rpm for 12 h. To obtain **3a** in a preparative scale, the reaction medium (50 mL) containing 50 mM FALD, FLS (2 U), FSA (12 U) in 100 mM tris-ethanolamine (TEA) buffer pH 7.5 was carried out at 25°C and 100 rpm for 36 h. Samples were treated with 10 % H₂SO₄ and centrifuged at 22,000 rpm for 20 min. The supernatant was concentrated under reduced pressure and purified using semi-preparative HPLC to afford the desired product **3a**. Isolated yield: 26%. ¹H NMR and ¹³C spectra were identical to those in previous study. [¹] ¹H NMR (400 MHz, D₂O) δ 4.96 (s, 1H), 3.76 (d, J = 1.5 Hz, 1H), 3.64 (s, 1H), 3.61 (d, J = 0.7 Hz, 1H). ¹³C NMR (101 MHz, D₂O) δ 89.77, 74.05, 61.96.

To synthesise **4a**, the reaction mixture (2 mL) containing 50 mM TEA (pH 7.5), 25 mM FALD, and FLS (20 mg, 0.1 U) was initially carried out at 25 °C and 120 rpm

for 12 h. Then, FSA^{A129S} (1 mg, 0.64 U) and FALD (8 mM) were added into the reaction system, and it performed for another 12 h. Samples (100 µl) were captured every one hour, treated with 10 % H₂SO₄, centrifuged (22,000 rpm, 20 min) and analysed by HPLC. For a preparative scale, the reaction mixture (50 mL) containing 50 mM TEA (pH 7.5), 50 mM FALD, and FLS (2 U) was initially carried out at 25 °C and 120 rpm for 12 h. Then, FSA^{A129S} (5 U) and FALD (8 mM) were added into the reaction system, and it performed for another 12 h. The supernatant was concentrated under reduced pressure and purified using semi-preparative HPLC to afford the desired product **4a**. Isolated yield: 66%. ¹H NMR and ¹³C spectra were identical to those in previous study. [²¹ ¹H NMR (400 MHz, D₂O) δ 4.62 (d, *J* = 19.4 Hz, 1H), 4.54 (d, *J* = 19.4 Hz, 1H), 4.47 (t, *J* = 4.1 Hz, 1H), 3.89 (d, *J* = 2.3 Hz, 1H). ¹³C NMR (101 MHz, D₂O) δ 212.30, 75.83, 65.80, 62.84.

To synthesise **6a**, one-pot reaction mixture (2 mL) containing 50 mM TEA (pH 7.5), 25 mM FALD, FLS (20 mg, 0.1 U) and FSA (2 mg, 0.32 U) was initially carried out at 25 °C and 120 rpm for 12 h. Then, FSA^{A129S} (0.8 mg, 5 U) were added into the reaction system, and it performed for another 12 h. Samples (100 µl) were captured every one hour, treated with 10 % H₂SO₄, centrifuged (22,000 rpm, 20 min) and analyzed by HPLC. For a preparative scale, the reaction medium (50 mL) containing 50 mM FALD, FLS (2 U), FSA (15 U) in 100 mM tris-ethanolamine (TEA) buffer pH 7.5 was carried out at 25 °C and 100 rpm for 36 h. Then, FSA^{A129S} (5 U) were added into the reaction system, and it performed for another 12 h. The supernatant was concentrated under reduced pressure and purified using semi-preparative HPLC to

afford the desired product **6a**. Isolated yield: 43%. ^1H NMR and ^{13}C spectra were identical to those in previous study. ^{13}C ^1H NMR (400 MHz, D_2O) δ 3.79 – 3.74 (m, 1H), 3.73 (d, J = 5.0 Hz, 1H), 3.71 – 3.69 (m, 1H), 3.68 (s, 1H), 3.66 – 3.63 (m, 1H), 3.54 (d, J = 9.0 Hz, 1H), 3.52 (d, J = 6.5 Hz, 1H). ^{13}C NMR (101 MHz, D_2O) δ 97.71, 73.83, 70.37, 69.40, 63.47, 61.76.

Whole-cell transformation

The recombinant *E. coli* strains were initially cultivated in 500 mL of LB medium to an OD_{600} of 0.6 and then induced with 1 mM IPTG for 20 h at 20°C and 120 rpm. Subsequently, the cells were harvested by centrifugation (8.000×g, 10 min, 4 °C), washed twice and suspended in TEA buffer (50 mM, pH 7.5). The reaction medium containing 30 g cdw l^{-1} (OD_{600} =120) resting cells expressing FSA^{A129S}, 2 M or 3 M substrates (FALD and DHA) were carried out at 30°C and 120 rpm for 4 hours.

For stepwise cascade whole-cell transformation, the cell catalyst overexpressing FSA (5 g cdw l^{-1}), 200 mM FALD and 50 mM GALD were added together in the reaction mixture (2 mL). The reaction system was initially performed at 30 °C and 120 rpm for 4 hours. Then, cell catalyst overexpressing FSA^{A129S} (5 g cdw l^{-1}) and 180 mM DHA supplemented into the system and the reaction process continued for another 4 hours. Samples (100 μl) were captured every 30 min, centrifuged (14,000 rpm, 20 min) and analyzed by HPLC.

Analytical Methods

Protein concentrations were determined by the Bradford method using bovine serum albumin as a standard. Compounds FALD, GALD, DHA, **3a**, **4a**, and **6a** were measured using HPLC system (Agilent 1100 series, Hewlett-Packard), equipped with a refractive index detector. The HPLC was fitted with a Bio-Rad Aminex HPX-87H column. The mobile phase consisted of 10 mM H₂SO₄ at 0.5 mL min⁻¹ (55 °C).

Table S1 Strains and plasmids used in this study.

| Strains and plasmids | Genotype and properties | Source or reference |
|------------------------------------|--|---------------------|
| Strain | | |
| <i>E. coli</i> DH5 α | <i>supE44 ΔlacU169 (φ80 lacZΔM15) hsdR17 recA1 endA1 gyrA96 thi-1 relA1</i> | Novagen |
| <i>E. coli</i> BL21 (DE3) | <i>F-ompT hsdSB (Rb - mB -) gal (λ c I 857 ind1 Sam7 nin5 lacUV5 T7gene1) dcm (DE3)</i> | Novagen |
| Plasmid | | |
| pET21a(+) | Expression vector, <i>Ap</i> ^R | Invitrogen |
| pETDuet | Expression vector, <i>Ap</i> ^R | Invitrogen |
| pET-FLS | pET21a(+) derivative carrying gene <i>FLS</i> which was synthesized from GeneScript | This study |
| pET-FSA | pET21a(+) derivative carrying gene sequence of <i>fsa</i> | This study |
| pET-FSA ^{A129S} | pET21a(+) derivative carrying gene sequence of mutant FSA ^{A129S} | This study |
| pET-T7-FLS-T7-FSA ^{A129S} | pETDuet derivative carrying gene sequence of <i>FLS</i> and FSA ^{A129S} | This study |

Table S2 Primers used in this study.

| Primer | Oligonucleotide sequence |
|-------------------------|---|
| FSA-1 | GGAATTCCATATGcatcatcaccatcaccatGAACTGTATCTGGATACTTCAGAC |
| FSA-2 | ACTCAAGCTTCTATTAAATCGACGTTCTGCCAAA |
| FSA ^{A129S} -1 | TGCGGAATATGTTTCGCCTTACGTTAATCG |
| FSA ^{A129S} -2 | CGATTAACGTAAGGCGAAACATATTCCGCA |

Figure S1 The effect of concentration ratio of FALD to GALD on conversion rate in vitro.

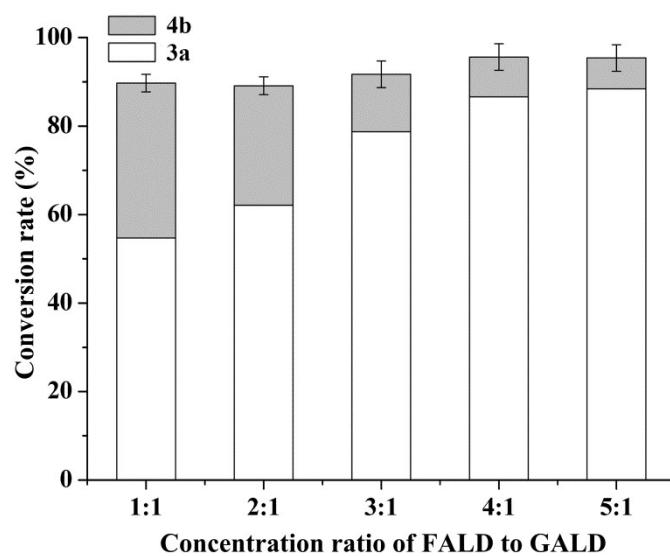


Figure S2 SDS-PAGE analysis of purified FLS, FSA, and FSA^{A129S}. M: Marker; 1: Purified FLS; 2: Purified FSA; 3: Purified FSA^{A129S}.

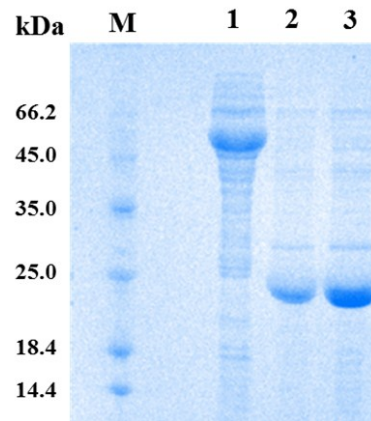
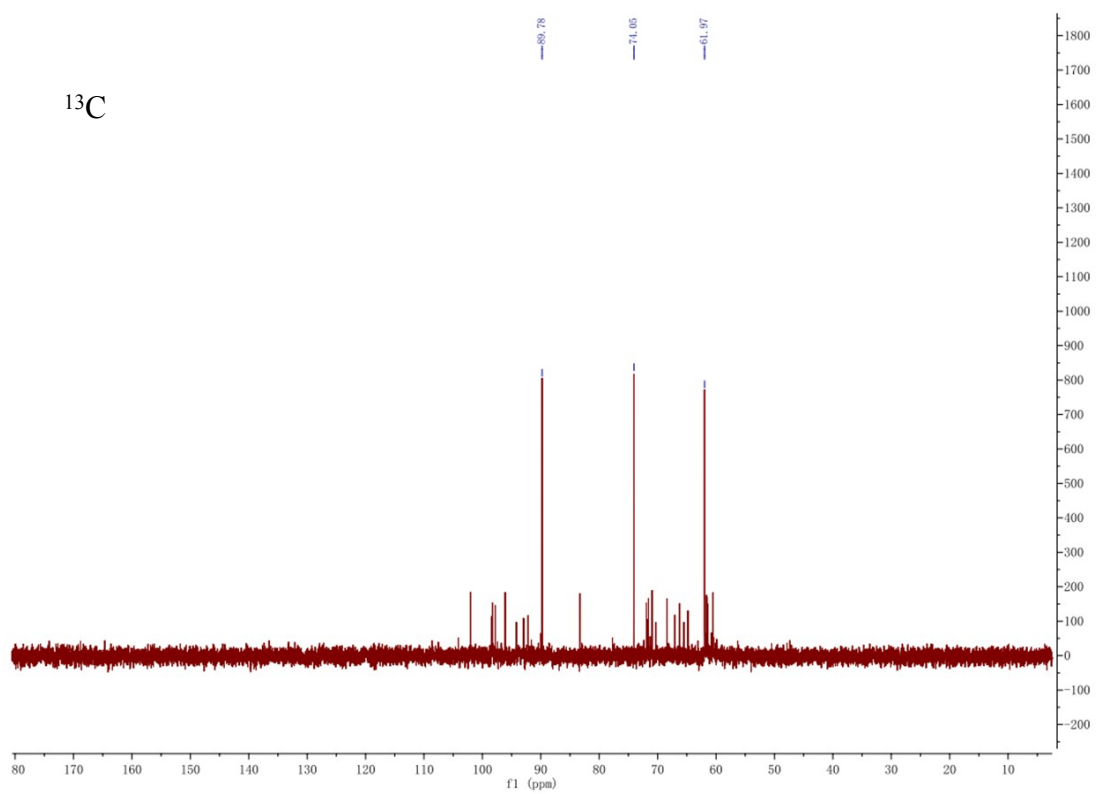
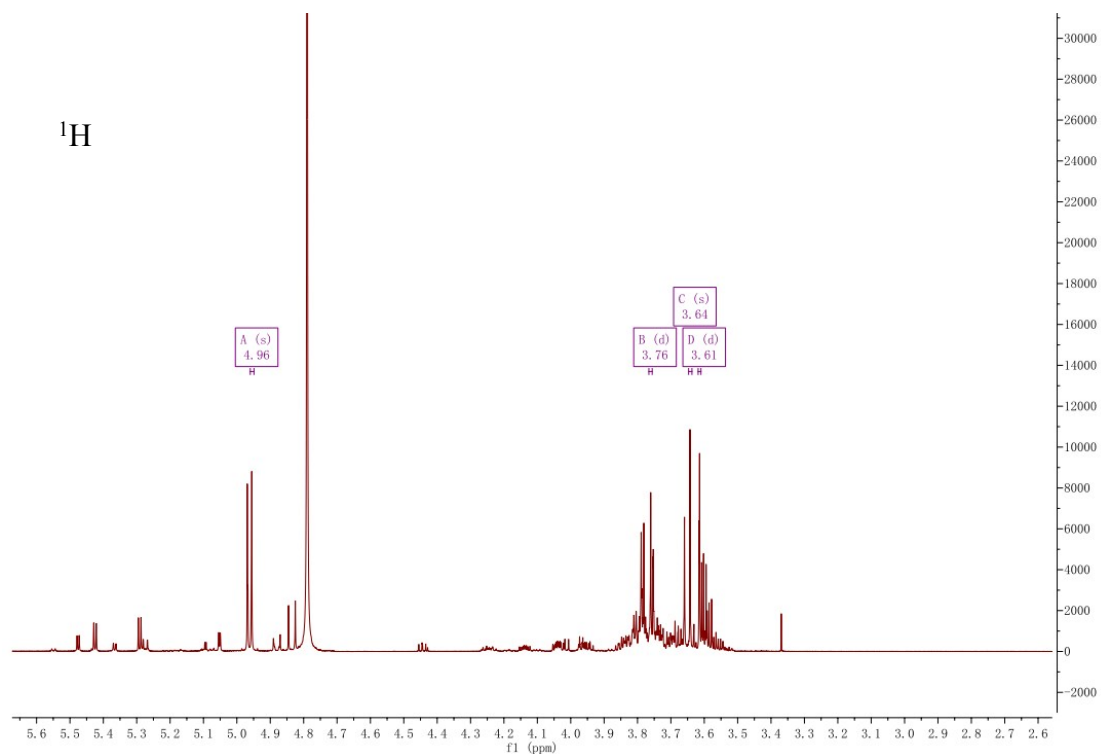
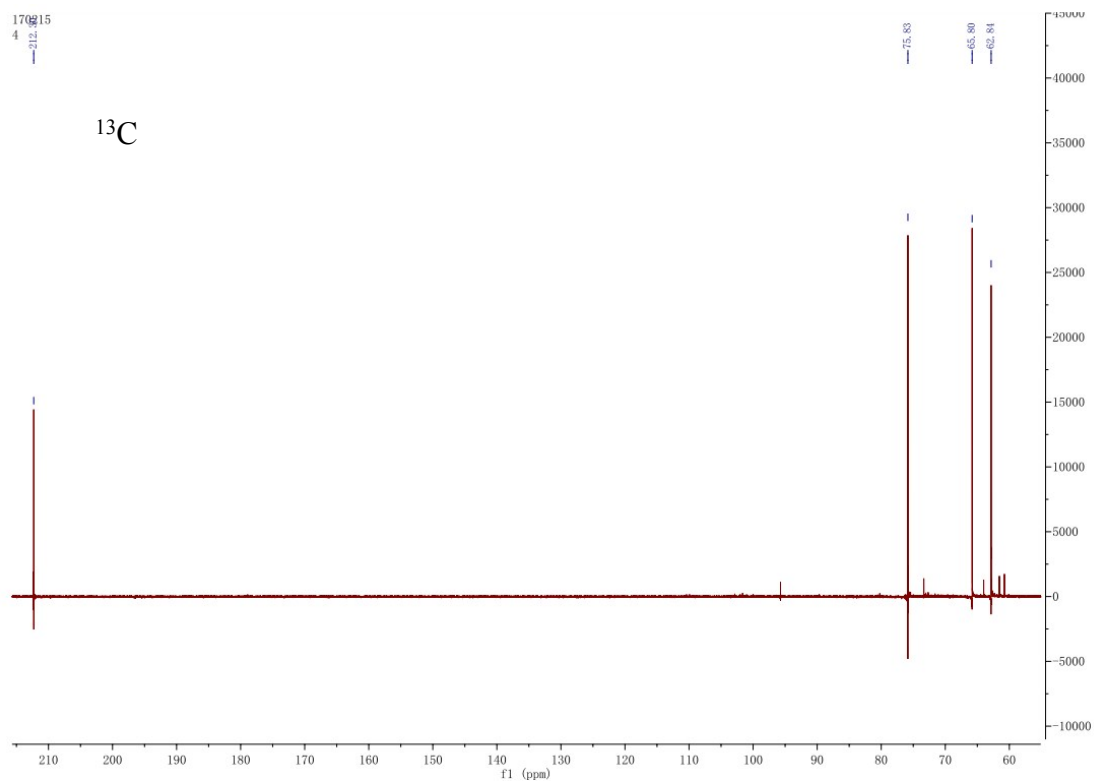
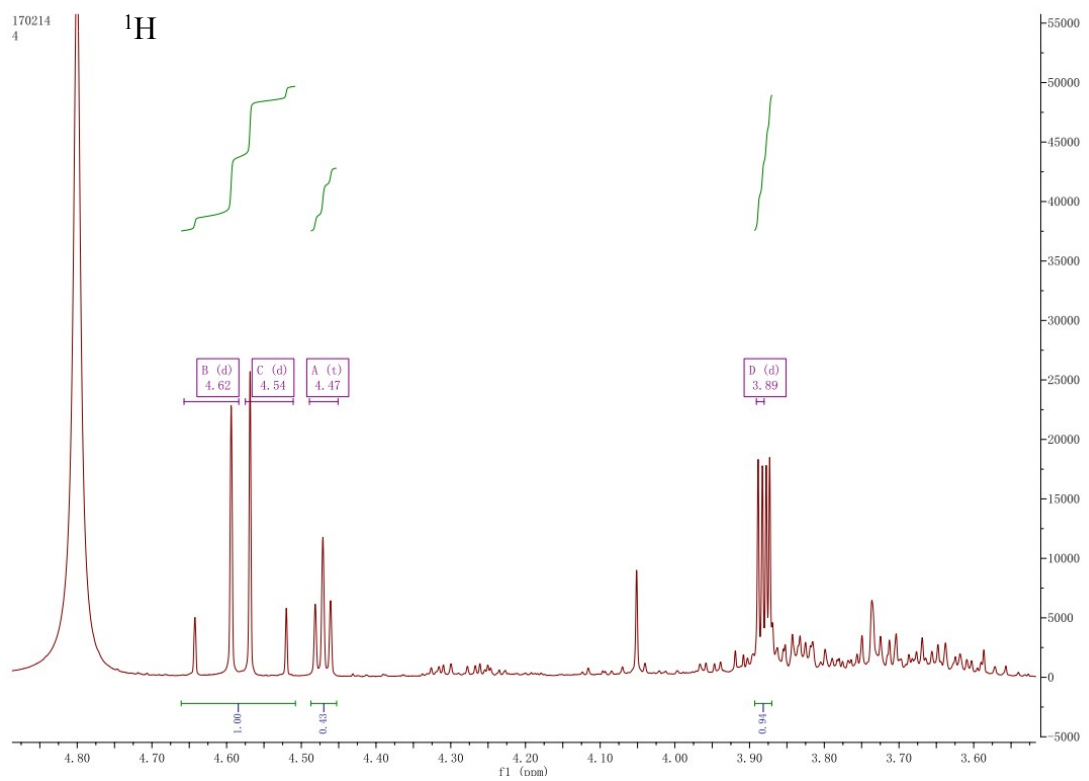


Figure S3 NMR analysis of **3a**, **4a**, and **6a**.

(1) ^1H and ^{13}C spectra of **3a**



(2) ^1H and ^{13}C spectra of **4a**



(3) ^1H and ^{13}C spectra of **6a**

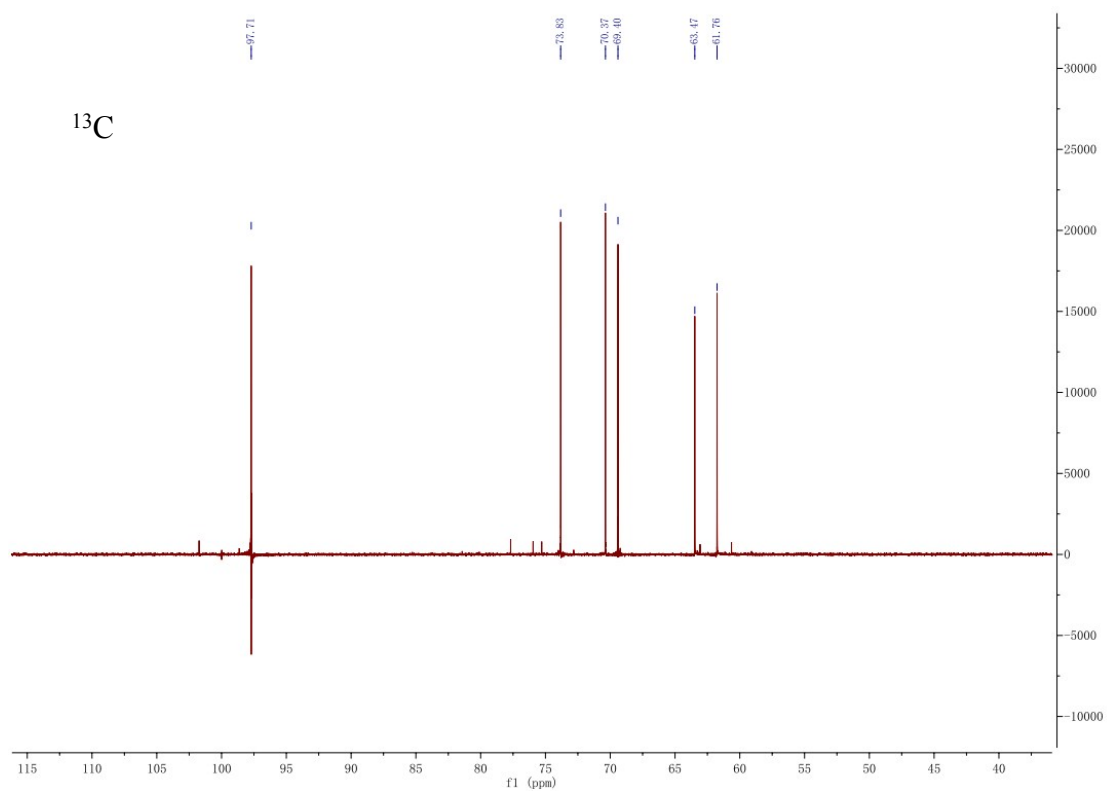
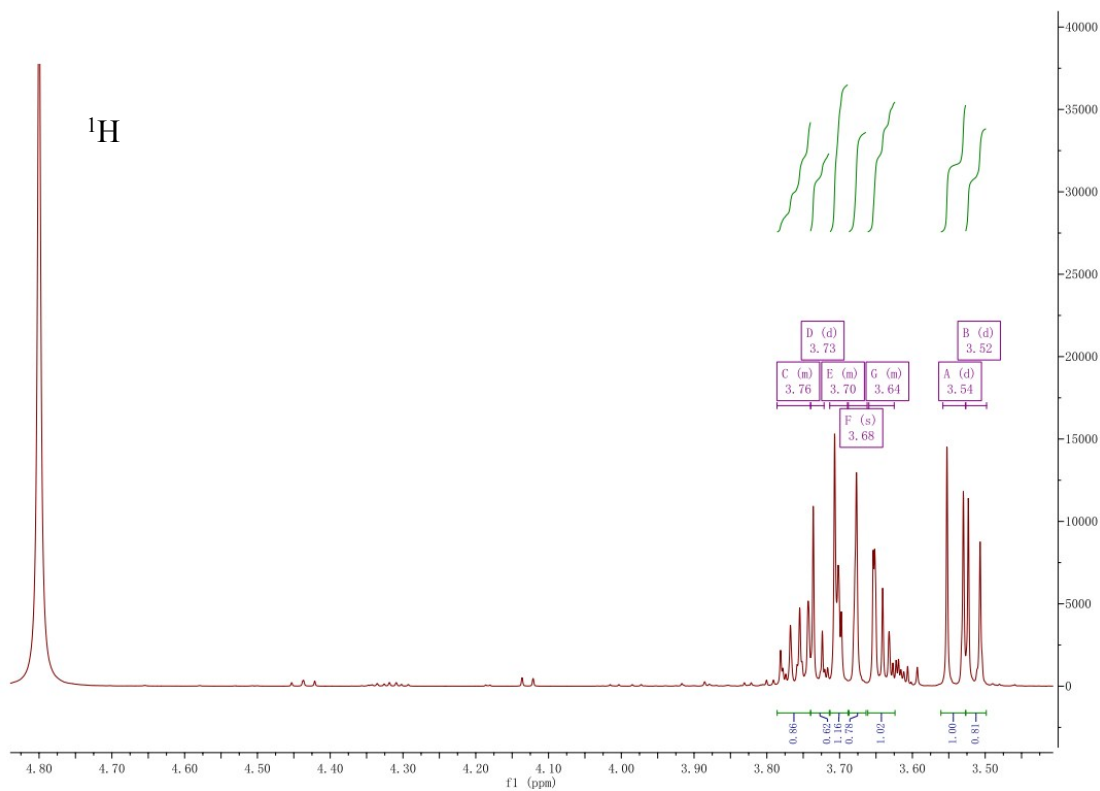
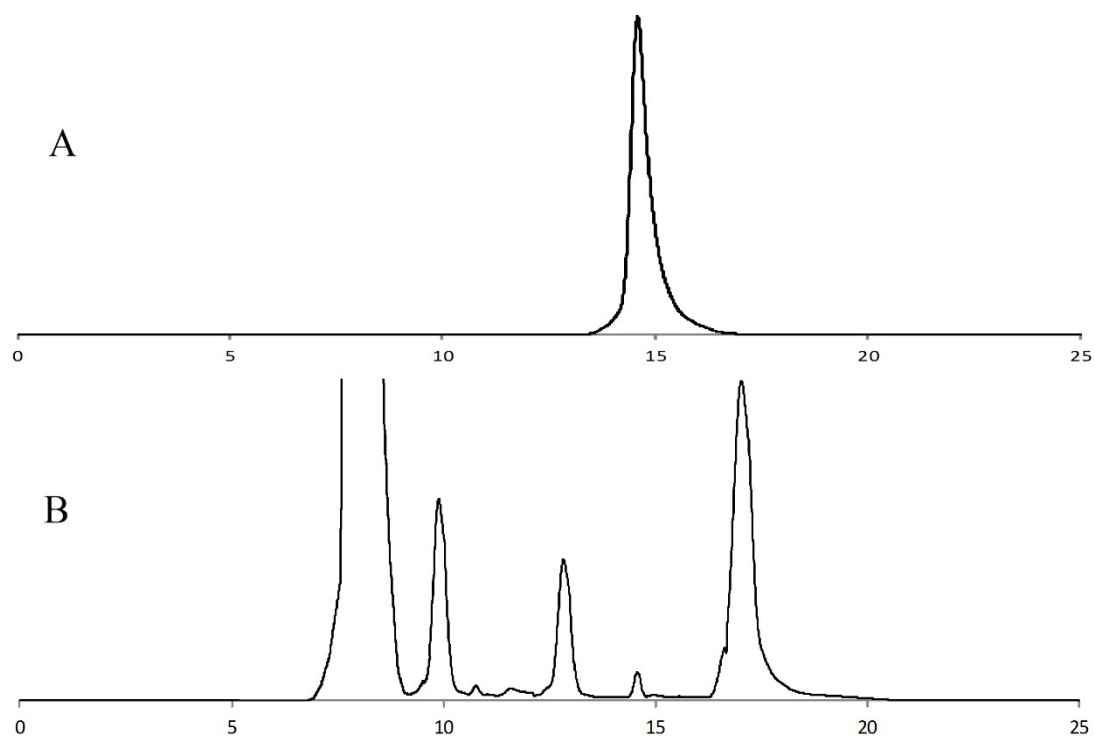


Fig. S4 Whole-cell transformation of FALD to synthesise **4a** using recombinant strain expressing both of FLS and FSA^{A129S}. (A) The standard of **4a**; (B) Whole-cell transformation of FALD to synthesise **4a**.



Reference

- [1] X. Garrabou, J.A. Castillo, C. Guérard-Hélaine, T. Parella, J. Joglar, M. Lemaire, P. Clapés, *Angew Chem Int Ed*, **2009**, *121*, 5629–5633.
- [2] J. A. Castillo, C. Guérard-Hélaine, M. Gutiérrez, X. Garrabou, M. Sancelme, M. Schürmann, T. Inoue, V. Hélaine, F. Charmantray, T. Gefflaut, L. Hecquet, J. Joglar, P. Clapés, G. A. Sprenger, M. Lemaire, *Adv. Synth. Catal*, **2010**, *352*, 1039–1046.
- [3] A. Blaskó, C.A. Bunton, E. Moroga, S. Bunel, C. Ibarra. *Carbohydr. Res*, **1995**, *278*, 315.