Supporting Information for

Synthesis of Carbohydrates in Mineral-Guided Prebiotic Cycles

Hyo-Joong Kim,[a] Alonso Ricardo,[b] Heshan I. Illangkoon,[c] Myong Jung Kim,[a] Matthew A. Carrigan,[a] Fabianne Frye,[d] and Steven A. Benner[a] .*

a) Foundation for Applied Molecular Evolution, 1115 NW 4th Street, Gainesville FL 32601

b) Department of Genetics, Harvard Medical School, Massachusetts General Hospital, Howard Hughes Medical Institute, Simches Research Center, 185 Cambridge Street, Boston MA 02114

c) Department of Chemistry, University of Florida, Gainesville FL 32611

d) Department of Chemistry, University of California, Irvine CA 92697

* sbenner@ffame.org

TABLE OF CONTENTS

Comments on Referee Comments

This system was not explored in the presence of other nucleophiles, such as cyanide, ammonia, amines, or sulfite.
Borate does not inhibit the formation or threose, and strongly binds to and stabilizes threose that is formed.
The light-dependence of the formose process concerns the initial step, the joining of two HCHO molecules to give glycolaldehyde; this is not relevant to the discussion here.
Some of these results were reviewed in: Planetary Organic Chemistry and the Origins of Biomolecules. Steven A. Benner, Hyo-Joong Kim, Myung-Jung Kim, and Alonso Ricardo, Cold Spring Harb Perspect Biol 2010; 2:a003467

Table S1. Estimated half-lifes for decomposition of pentoses and xylulose at saturating calcium deuteroxide (pH ~12.5) in the absence and presence of the boron-containing mineral colemanite.

Carbohydrate	Half-life (min): $Ca(OD)_2$	Half-life (min): $Ca(OD)_2$ + Colemanite
D-ribose	291	2700
D-arabinose	144	259; 331
L-arabinose	124	380; 274
D-lyxose	177	1382
L-lyxose	184	1201
L-xylose	66	572; 258
ribulose	14	2028

Half-lifes for decomposition of the indicated carbohydrates were estimated by measuring the ratio of the signal integral arising from the anomeric proton (for ribulose, the signal from C-5) relative to internal standard (sodium benzoate) over time. In the absence of borate, decomposition products could not be determined (they are "tar"). In the presence of borate, the carbohydrates generally interconverted by borate moderated isomerization (confirmed by GC-MS analysis, data not shown). Multiple isomerization events eventually incorporate deuterium into the anomeric site, causing the loss of its ^{1}H NMR signal. This accounts for several half-life times reported for the same pentose.

Table S2. Percentage of 2'-^{13}C-hydroxymethylerythrose (10 mM) remaining after the indicated time in carbonate buffer (1.1 M, pH 10.4) containing the indicated amount of borate at 65 °C.

	0 mM	5 mM	10 mM	20 mM	40 mM
0 h	100	100	100	100	100
1 h	21	55	60	N.D.	N.D.
2 h	~0	40	48	77	N.D.
4 h	0	21	44	60	N.D.
24h	N.D.	1	12	28	52
48 h	N.D.	N.D.	4	15	34

N.D (not determined). Percentages determined by loss of the nmr signal from the ^{13}C-enriched carbon relative to a carbonate internal standard.

Figure S1. Amount of formaldehyde remaining in the standard formose reaction as a function of time and concentration of HCHO, as determined using chromotropic acid. Absorbance is measured at 580 nm.

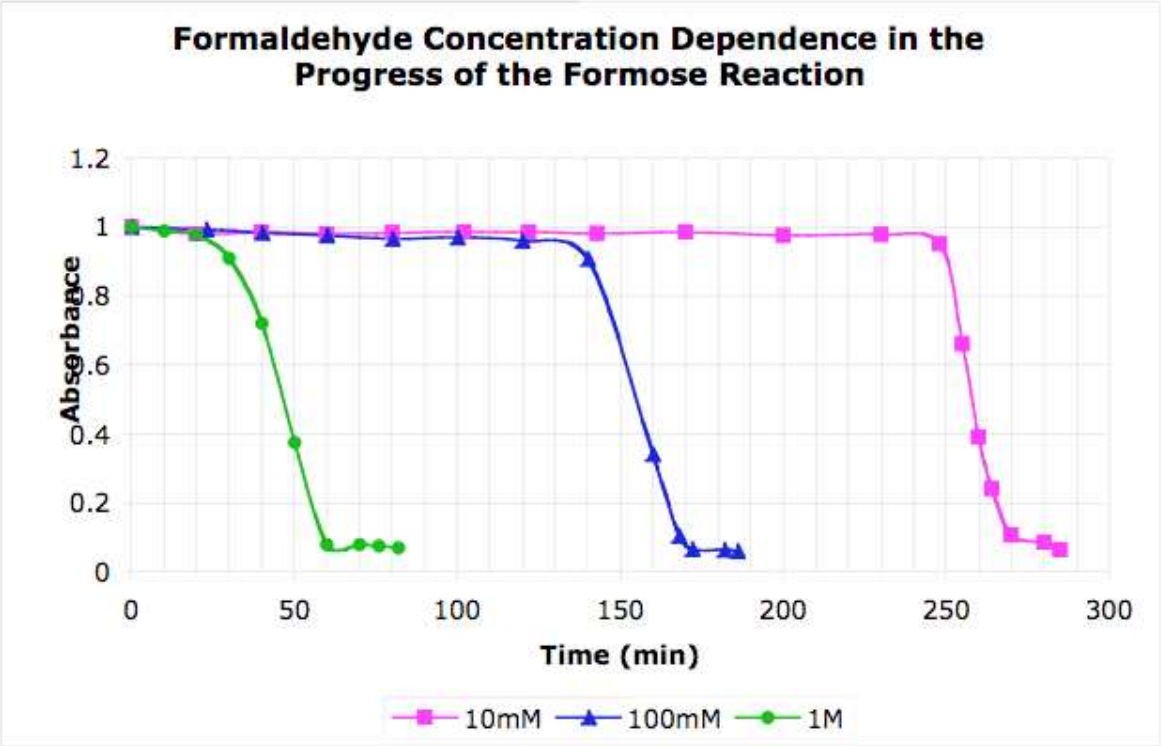
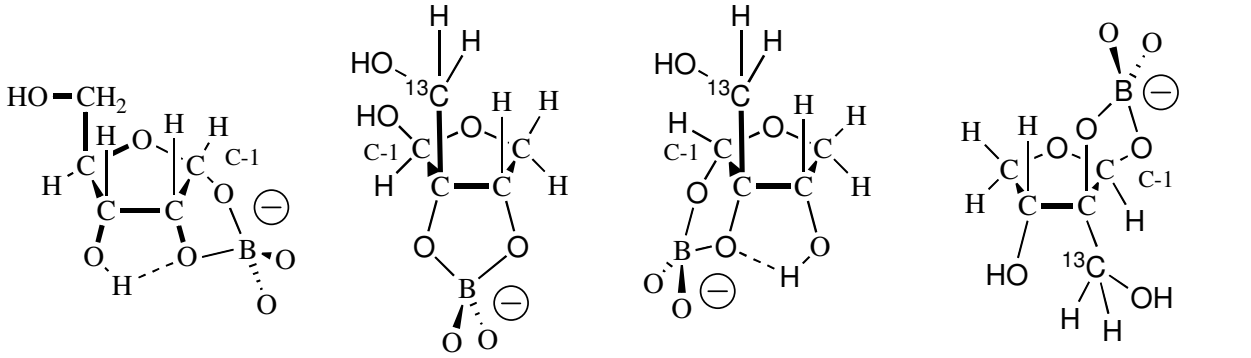
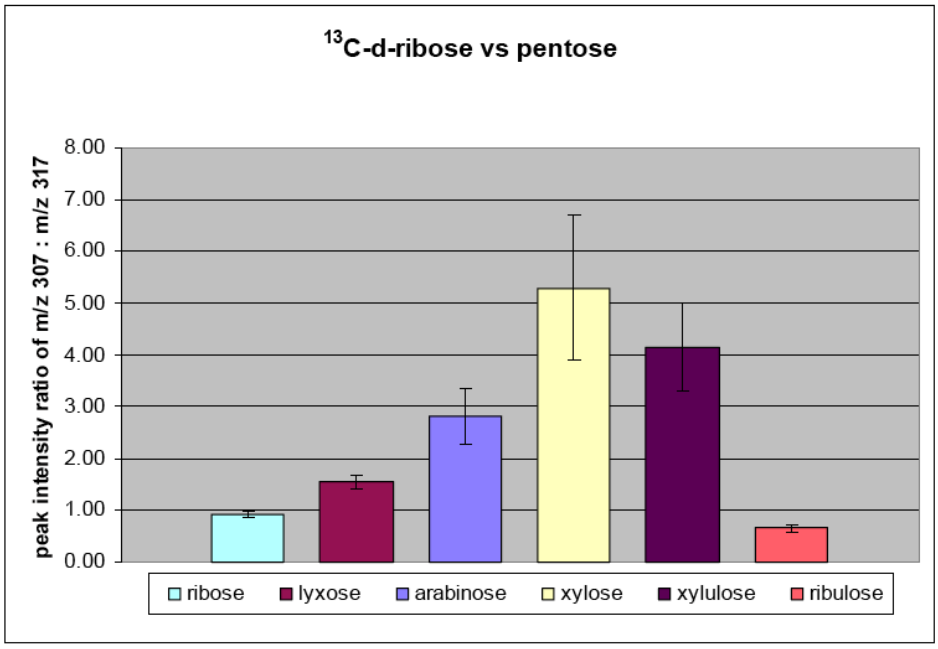


Figure S2. Cyclic forms of ribose (in its "alpha" anomer, where the –OH hydroxyl group at C-1 is down; the "D" enantiomer is shown), 2'-hydroxymethylerythrose (in both its alpha and beta form; enantiomer shown is the one synthesized with a ^{13}C label), and 2'-hydroxymethylthreose (in both its alpha and beta form; enantiomer shown is the one synthesized with a ^{13}C label). The site of boron complexation is known by nmr spectrometry.[12] The site of complexation of boron to the cyclic forms of the branched pentoses is speculative; nmr data suggest that the erythrose forms two borate complexes, while the threose forms only one predominantly.

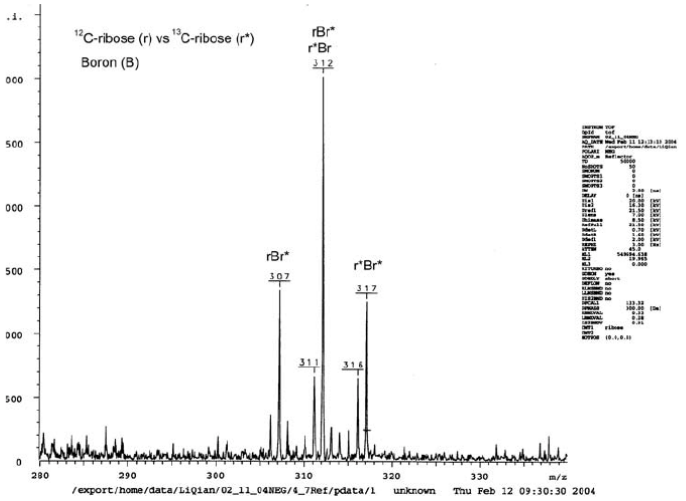


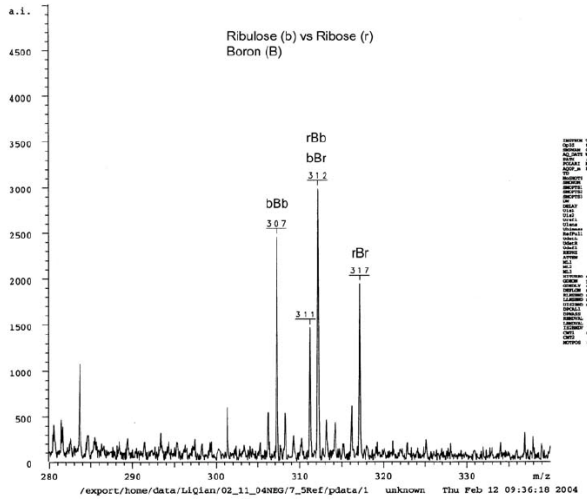
ribose 2'-hydroxymethylerythrose 2'-hydroxymethylthreose

Figure S3. Mass spectrometry competition experiments were used to determine the relative stability of borate complexes of pentoses and pentuloses. The smaller the bar, the more stable the complex.



Fully isotopically labeled ribose (^{13}C-ribose, 98.8% labeled) was used in competition experiments for borate binding versus each of the isomeric pentoses (aldo and keto pentoses) and determined by "Desorption Ionization on Silicon" (DIOS) mass spectrometry. Conditions are as reported in Li, Q., Ricardo. A., Benner, S. A., Winefordner, J. A. & Powell, D. H. Desorption/ionization on porous silicon mass spectrometry (DIOS-MS) studies on pentose-boron complexes. *Anal. Chem*. **77**, 4503-4508 (2005). Each pentose was analyzed separately. Error bars represent the standard deviation calculated from a set of 10 different measurements. Primary data are shown below. ^{12}C-D-ribose was included in the analysis to test the effect of the isotope in ionization-vaporization efficiency; no such effect was observed during the experiment. The slight deviation from the value of 1.0 in D-ribose is due to the presence of 1.1% unlabeled carbon in the 13C-D-ribose used for the competition experiment.


^{12}C-ribose (r) vs ^{13}C-ribose (r*)
Boron (B)
rBr*
r*Br
312
rBr*
307
r*Br*
317
311
316
m/z
/export/home/data/LiQian/02_11_04NEG/4_7Ref/pdata/1 unknown Thu Feb 12 09:30:30 2004


Ribulose (b) vs Ribose (r)
Boron (B)
rBb
bBr
312
bBb
307
rBr
317
311
m/z
/export/home/data/LiQian/02_11_04NEG/7_5Ref/pdata/1 unknown Thu Feb 12 09:36:16 2004

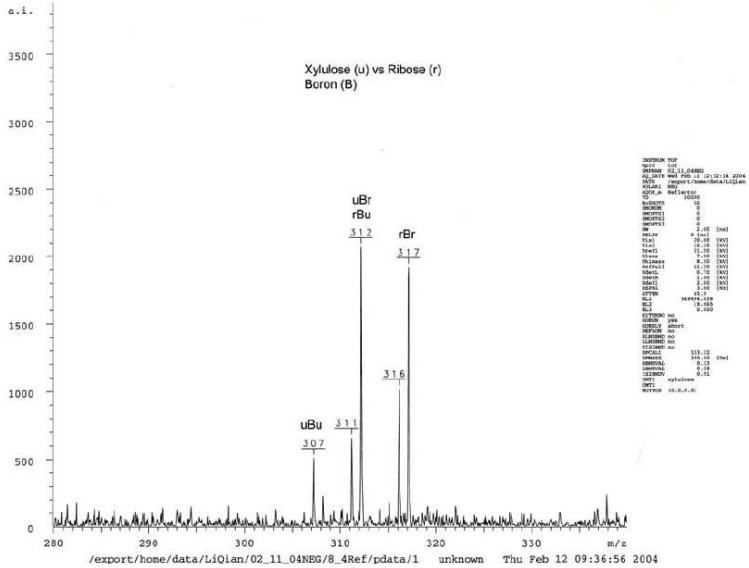
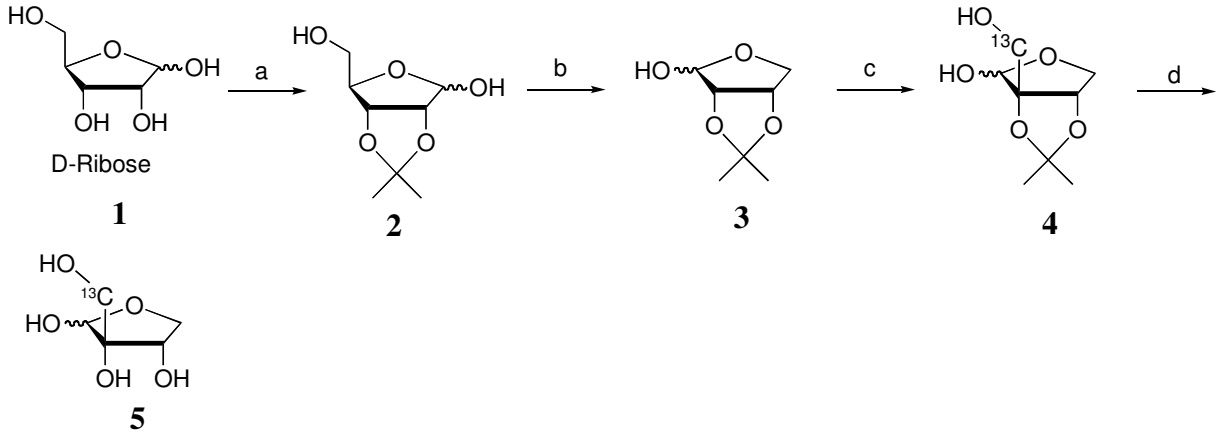

Xylulose (u) vs Ribose (r)
Boron (B)
uBr
rBu
312
rBr
317
316
uBu
307
311
m/z
/export/home/data/LiQian/02_11_04NEG/8_4Ref/pdata/1 unknown Thu Feb 12 09:36:56 2004

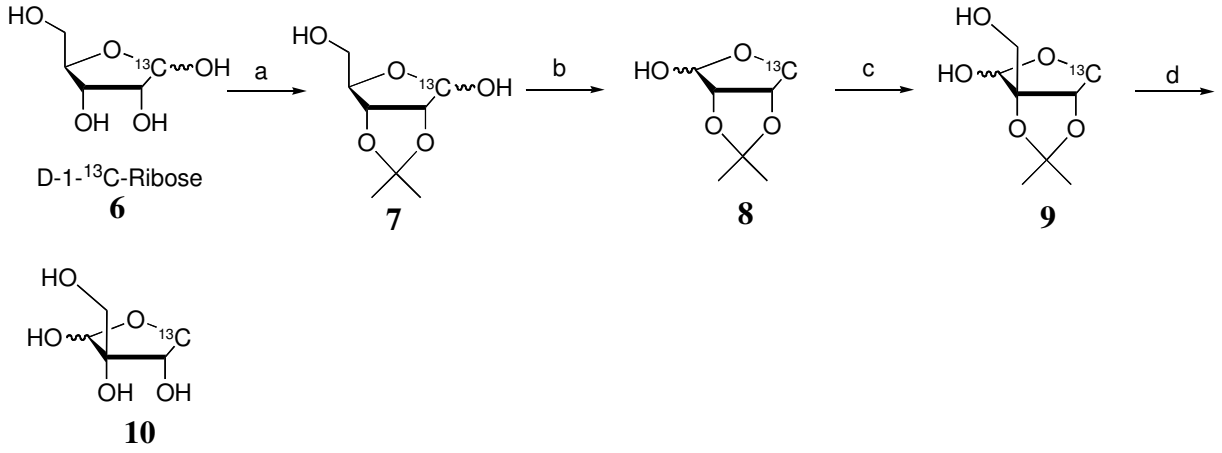
Figure S4. Synthesis of proposed compounds in Figure 1.

Figure S4a.
Synthesis of branched pentoses labeled with carbon-13. 2'-^{13}C-2-hydroxymethylerythrose (erythro **C5b**)



Reagents and conditions: a. 2,2-dimethoxypropane, p-TsOH, acetone, 80 %; b. 1. $NaBH_4$, H_2O. 2. $NaIO_4$, AcOH, H_2O, 71 %; c. $H^{13}CHO$, K_2CO_3, CH_3OH, 65 ^oC, 70 %; d. Dowex 50WX8 (H^+), H_2O, 80^oC, 95 %.

Figure S4b.
Synthesis of branched pentoses labeled with carbon-13. 4-^{13}C-2-hydroxymethylerythrose (erhtyro **C5b**)



Reagents and conditions: a. 2,2-dimethoxypropane, p-TsOH, acetone, 80 %; b. 1. $NaBH_4$, H_2O. 2. $NaIO_4$, AcOH, H_2O, 70 %; c. HCHO, K_2CO_3, CH_3OH, 65 ^oC, 55 %; d. Dowex 50WX8 (H^+), H_2O, 80^oC, 95 %

Representative synthetic procedures for unlabeled 2-hydroxymethylerythrose

Compound **2** (2,3-*O*-Isopropylidene-D-riboose)

Compound **2** was prepared from D-ribose followed by literature method.
Arumugham, B.; Kim, H. J.; Prichard, M. N.; Kern, E. R.; Chu, C. K. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 285. To a mixture of compound **1** (5.0 g, 33.3 mmol) and 2,2-dimethoxypropane (4.54 mL, 36.6 mmol) in acetone (100 mL) was added *p*-toluenesulfonic acid monohydrate (63 mg,

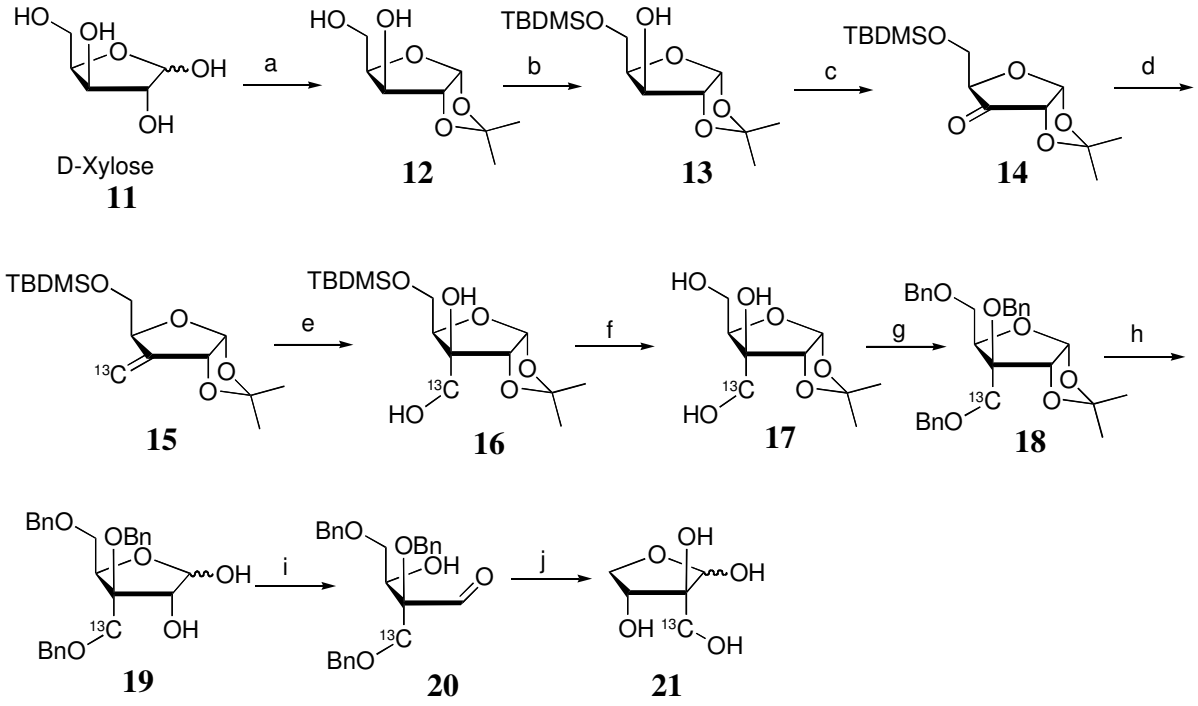
0.33 mmol) and stirred at rt for 1h. The mixture was neutralized with sodium bicarbonate (1 g) and evaporated with silica gel (15 g). The residue was purified by flash chromatography (silica, Hexanes:Ethyl acetate = 1:1 to 1:2) to give a colorless syrup (5.07 g, 80%).

Compound **3** (2,3-*O*-Isopropylidene-L-erythrose)
To a mixture of compound **2** (2.0 g, 10.5 mmol) in water (100 mL) was added sodium borohydride (0.44 g, 11.6 mmol) and stirred at rt for 4h. Acetic acid (4 mL) was added and stirred for 10 min. Then the mixture was treated with sodium periodate (2.48 g, 11.6 mmol) and stirred for 1 h and extracted with dichloromethane. The organic layer was concentrated and purified by flash chromatography (silica, Hexanes:Ethyl acetate = 2:1 to 1:1) to give a colorless liquid (1.2 g, 71%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.40 (d, 1H, *J* = 2.4 Hz), 4.82 (dd, 1H, *J* = 5.7, 3.3 Hz), 4.55 (d, 1H, *J* = 6.0 Hz), 3.9~4.1 (m, 2H), 3.23 (d, 1H, *J* = 2.1 Hz), 1.45 (s, 3H), 1.30 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 112.49, 101.95, 85.28, 80.09, 72.07, 26.35, 24.87.

Compound **4** (2,3-*O*-Isopropylidene-2-hydroxymethyl-L-erythrose)
A mixture of compound **3** (1.2 g, 7.49 mmol), formaldehyde (37%, 6 mL) and potassium carbonate (400 mg) in methanol (40 mL) was refluxed for 4 h. The mixture was poured into water (100 mL) and extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, filtered, concentrated and purified by flash chromatography (silica, Hexanes:Ethyl acetate = 2:1 to 1:1) to give a colorless liquid (1.0 g, 70%) as a mixture of two compounds. ^{13}C NMR (75 MHz, D_2O) δ 114.57, 114.46, 102.55, 98.53, 94.91, 89.68, 82.83, 82.07, 72.05, 67.75, 62.39, 61.72, 27.29, 27.20, 26.59, 26.43.

Compound **5** (2-Hydroxymethyl-L-erythrose)
A mixture of compound **4** (0.9 g, 4.7 mmol) and Dowex 50WX8 (3.0 g) in water (30 mL) was heated to 80 °C for 1 h. The mixture was filtered over Celite, concentrated and purified by flash chromatography (silica, CH_2Cl_2:MeOH = 5:1 to 2:1) to give a colorless syrup (670 mg, 95%) as a mixture of two compounds. ^{13}C NMR (75 MHz, D_2O) δ 102.18, 97.69, 81.13, 79.30, 71.51, 71.15, 70.82, 70.47, 63.82, 63.07.

Figure S4c.
Synthesis of branched pentoses labeled with carbon-13.
2'-^{13}C-2-hydroxy-^{13}C-methylthreose (threo **C5b**)



Reagents and conditions: a. 1. 2,2-dimethoxypropane, p-TsOH, acetone 2. 0.2% HCl in H_2O, 50 °C, 54 %; b. TBDMSCl, imidazole, CH_2Cl_2, 84 %; c. pyridinium dichromate, acetic anhydride, CH_2Cl_2, 83 %; d.$^{13}CH_3PPh_3Br$, KO^tBu, THF, 72 % ; e. OsO_4, NMO, THF/H_2O, 80 %; f. TBAF, THF, 85 %; g. BnBr, NaH, DMF, 90 %; h. 80% HCOOH, 50 °C, 2h, 50 %; i. 1. $NaBH_4$, EtOH/H_2O 2. $NaIO_4$, 90 %; j. Pd/C, MeOH, rt, 3days, 56 %

The spectral data shown below are from unlabeled 2-hydroxymethylthreose.

Compound **15** was prepared from D-xylose followed by literature method.

Jeong, L. S.; Yoo, S. J. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 847.
Yoo, S. J.; Kim, H. O.; Lim, Y.; Kim, J.; Jeong, L. S. *Bioorg. Med. Chem.* **2002**, *10*, 215.
Ma, T.; Pai, S. B.; Zhu, Y. L.; Lin, J. S.; Shanmuganathan, K.; Du, J.; Wang, C.; Kim, H.; Newton, M. G.; Cheng, Y. C.; Chu, C. K. *J. Med. Chem.* **1996**, *39*, 2835.

Compound **16**
To a mixture of compound **15** (150 mg, 0.50 mmol) and NMO (88 mg, 0.75 mmol) in THF/H_2O (5 mL/ 5 mL) was added OsO_4 (4% solution in H_2O, 0.1 mL, 0.015 mmol) and stirred at rt for 6h. The mixture was diluted with brine (20 mL) and extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Hexanes:EtOAc = 2:1) to give a colorless syrup (150 mg, 90%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.91 (d, 1H, *J* = 3.6 Hz), 4.57 (s, 1H), 4.39 (d, 1H, *J* = 3.6 Hz), 3.96~4.04 (m, 3H), 3.79 (d, 2H, *J* = 6.0 Hz), 2.80 (t, 1H, *J* = 6.0 Hz), 1.53 (s, 3H), 1.32 (s, 3H), 0.89 (s, 9H), 0.11 (s, 6H).

Compound **17**

A mixture of compound **16** (820 mg, 2.45 mmol) and TBAF (1M solution in THF, 3.5 mL, 3.5 mmol) in THF (20 mL) was stirred at rt for 1h. The mixture was concentrated and purified by flash chromatography (silica, Hexanes:EtOAc = 1:1 to EtOAc 100%) to give a light brown solid (405 mg, 75%). ^{1}H NMR (300 MHz, CD_3OD) δ 5.85 (d, 1H, *J* = 3.6 Hz), 4.37 (d, 1H, *J* = 3.6 Hz), 3.97 (t, 1H, *J* = 5.4 Hz), 3.65~3.85 (m, 4H), 1.47 (s, 3H), 1.30 (s, 3H); ^{13}C NMR (75 MHz, CD_3OD) δ 113.61, 106.01, 87.14, 83.27, 83.25, 63.50, 61.19, 27.45, 26.74.

Compound **18**

To a solution of compound **17** (405 mg, 2.83 mmol) in DMF (10 mL) was added NaH (60% dispersed in mineral oil, 330 mg, 8.24 mmol) at 0 °C and stirred for 10 min. The mixture was treated with benzyl bromide (0.76 mL, 6.4 mmol) and stirred for 30 min. The excess NaH was quenched with slow addition of water. The mixture was poured into water (30 mL) and extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, filtered, concentrated and purified by flash chromatography (silica, Hexanes:EtOAc = 10:1 to 5:1) to give a colorless liquid (756 mg, 84%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.2~7.4 (m, 15H), 5.90 (d, 1H, *J* = 3.9 Hz), 4.4~4.8 (m, 7H), 4.29 (dd, 1H, *J* = 6.6, 4.8 Hz), 3.7~4.0 (m, 4H), 1.50 (s, 3H), 1.32 (s, 3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 139.19, 138.45, 138.19, 128.50, 127.92, 127.85, 127.70, 127.43, 126.90, 112.44, 105.23, 86.50, 82.99, 82.28, 73.86, 73.67, 69.08, 66.40, 27.33, 26.79.

Compound **19**

A solution of compound **18** (390 mg, 0.80 mmol) in 80% formic acid (10 mL) was heated to 50 °C for 2 hours. The mixture was diluted with water (50 mL) and extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, filtered, concentrated and purified by flash chromatography (silica, Hexanes:EtOAc = 2.5:1 to 1:1) to give a colorless syrup (180 mg, 50%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.2~7.4 (m, 15H), 5.41 (m, 0.7H), 5.10 (d, 0.3H, *J* = 11.1 Hz), 4.3~4.7 (m, 9H), 3.6~4.0 (m, 5H), 2.83 (d, 0.7H, 6.9 Hz), 2.64 (br s, 0.3H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 148.45, 138.94, 138.41, 137.00, 128.90, 128.85, 128.72, 128.66, 128.50, 128.47, 128.38, 128.16, 128.00, 127.90, 127.73, 127.59, 127.56, 126.94, 103.85, 98.68, 96.33, 85.84, 84.60, 82.61, 81.20, 80.18, 76.95, 74.19, 73.87, 73.64, 69.54, 69.26, 68.99, 67.31, 66.58, 66.15.

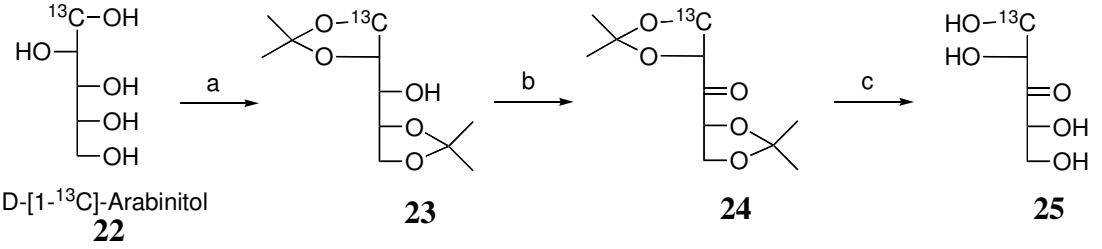
Compound **20**

To a mixture of compound **19** (170 mg, 0.38 mmol) in MeOH (5 mL) and water (7 mL) was added $NaBH_4$ (22 mg, 0.57 mmol) and stirred at rt for 4h. The mixture was treated with acetic acid (1 mL) and stirred 10 min. It was treated with $NaIO_4$ (122mg, 0.57 mmol), water (10 mL) and ethyl acetate (10 mL) and stirred at rt for 1h. The mixture was diluted with water (20 mL) and extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, filtered, concentrated and purified by flash chromatography (silica, Hexanes:EtOAc = 10:1 to 6:1 to 4:1) to give a colorless syrup (88 mg, 55%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.75 (s, 1H), 8.04 (s, 1H), 7.2~7.4 (m, 15H), 5.62 (m, 1H), 4.69 (dd, 2H, *J* = 17.1, 11.1 Hz), 4.4~4.6 (m, 4H), 3.7~3.9 (m, 4H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 200.68, 160.00, 148.46, 138.06, 137.61, 137.33, 128.70, 128.64, 128.61, 128.24, 128.21, 128.00, 127.89, 127.63, 83.77, 73.92, 73.60, 72.54, 67.84, 67.36, 67.28.

Compound **21**

To a mixture of compound **20** (88 mg, 0.21 mmol) in MeOH (10 mL) was added Pd/C (10 %) (300 mg) and stirred under H_2 gas at rt for 48h. The mixture was filtered and the filtrate was concentrated and purified by flash chromatography (silica, CH_2Cl_2:MeOH = 8:1 to 5:1 to 2.5:1) to give a colorless syrup (28 mg, 89%). ^{13}C NMR (75 MHz, D_2O) δ 102.15, 98.70, 83.76, 81.46, 76.13, 75.27, 75.02, 72.55, 62.18, 61.07.

Figure S4d.
Synthesis of linear pentulose 1,2,4,5-tetrahydroxypentan-3-one labeled with carbon-13. 1-^{13}C-1,2,4,5-tetrahydroxypentan-3-one (**C5l**) (Linclau, B.; Boydell, A. J.; Clarke, P. J.; Horan, R.; Jacquet, C. *J. Org. Chem*. **2003,** *68,* 1821.)



D-[1-^{13}C]-Arabinitol **22** **23** **24** **25**

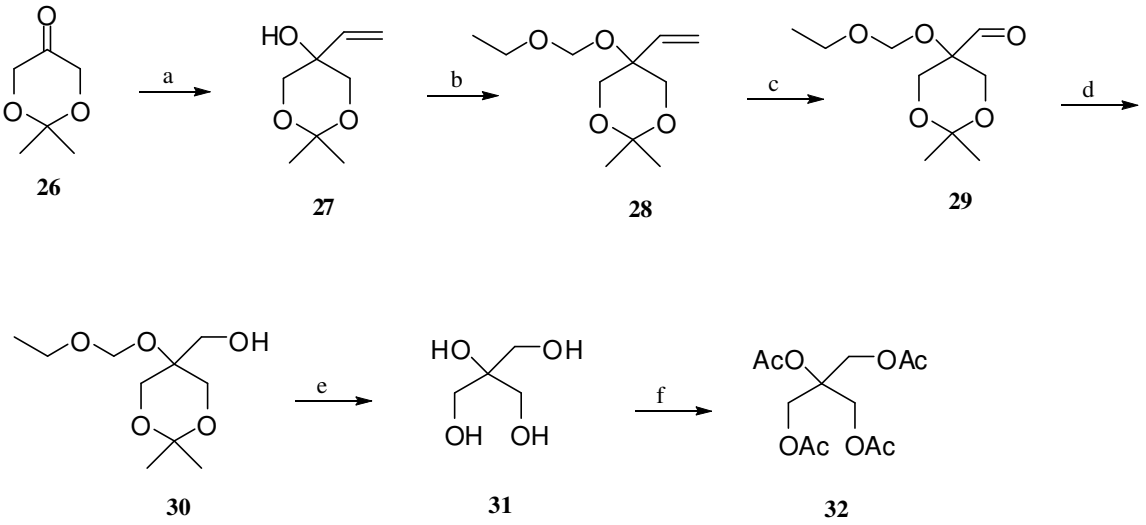
Reagents and conditions: a. 2,2-dimethoxypropane, CSA, THF, reflux; b. PDC, Ac_2O, CH_2Cl_2, reflux, 42 % for two steps; c. MeOH, H_2O, CSA, 50 ^oC, 86 %.

Compound **23** (1,2:4,5-Di-*O*-(2,2-isopropylidene)-arabinitol)
A mixture of arabinitol (0.25 g, 1.63 mmol), 2,2-dimethoxypropane (0.80 mL) and camphosulfonic acid (113 mg) in THF (10 mL) was refluxed for 15 min. The mixture was poured into water and extracted with ethylether. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Hexanes:Ethyl acetate = 5:1 to 2:1) to give a colorless syrup as a inseparable mixture (246 mg, 65%) of the desired product (^{13}C NMR (75 MHz, $CDCl_3$) δ 66.44) and side products.

Compound **24** (1,2:4,5-Bis(2,2-propylidenedioxy)-3-pentanone)
A mixture of the protected arabinitol mixture from above procedure, pyridinium dichromate (0.29 g, 0.77 mmol) and acetic anhydride (0.36 mL, 3.84 mmol) in dichloromethane was refluxed for 3h. The mixture was filtered over Celite and purified by flash chromatography (silica, Hexanes:Ethyl acetate = 5:1 to 3.5:1) to give a colorless liquid (160 mg, 43% for two steps). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.75~4.85 (m, 2H), 4.52 (t, 0.5H, *J* = 8.4 Hz), 4.23~4.33 (m, 2H), 3.98~4.08 (m, 2H), 3.79 (dd, 0.5H, *J* = 8.7, 6.0 Hz), 1.47 (s, 6H), 1.40 (s, 6H); ^{13}C NMR (75 MHz, $CDCl_3$) δ 66.27 (only labeled carbon was shown).

Compound **25** (1,2,4,5-tetrahydroxy-3-pentanone)
A mixture of the compound **24** (160 mg, 0.70 mmol) and camphosulfonic acid (25 mg) in methanol (7 mL) and water (0.5 mL) was heated to 50 °C for 3h and stirred at rt for 16h. Sodium bicarbonate was added to the mixture, evaporated with silica gel and purified by flash chromatography (silica, CH_2Cl_2:MeOH = 5:1 to 3:1) to give a colorless liquid (90 mg, 85%). ^{13}C NMR (75 MHz, D2O) δ 63.39 (only labeled carbon was shown).

Figure S4e.
Synthesis of reduced branched tetrose (**C4b**).



a. vinylmagnesium bromide, THF, 74%; b. chloromethyl ethyl ether, DIPEA, CH_2Cl_2, 70%; c. OsO_4, NMO, THF/H_2O then $NaIO_4$, CH_2Cl_2/H_2O, 74%; d. Dowex 50WX8-400, H_2O, 66% ; f. Ac_2O, DMAP, pyridine.

Compound **27** (2,2-Dimethyl-5-vinyl-1,3-dioxan-5-ol)
To a mixture of vinylmagnesium bromide (1M in THF, 17.3 mL, 17.3 mmol) in THF (50 mL) was added compound **26** (750 mg, 5.77 mmol) dissolved in THF (4 mL) at 0 °C. After 30 min, the mixture was quenched with saturated aqueous NH_4Cl and extracted with ethylacetate. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:6 to 1:2) to give a white solid (680 mg, 74%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.6~5.8 (m, 1H), 5.46 (dd, 1H, *J* = 17.4, 1.2 Hz), 5.27 (dd, 1H, *J* = 10.8, 1.2 Hz), 3.91 (d, 2H, *J* = 12 Hz), 3.53 (d, 2H, *J* = 11.7 Hz), 3.30 (s, 1H), 1.47 (s, 3H), 1.46 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 135.82, 117.23, 98.32, 68.61, 68.37, 29.07, 18.26.

Compound **28**
To a mixture of compound **27** (81 mg, 0.51 mmol) and diisopropylethylamine (2 mL) in CH_2Cl_2 (2 mL) was added chloromethyl ethyl ether (0.47 mL, 5.1 mmol) at rt and stirred overnight. The mixture was poured into brine and extracted with ethylacetate. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:6) to give a light brown liquid (77 mg, 70%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.78 (dd, 1H, *J* = 18, 11.1 Hz), 5.42 (dd, 1H, *J* = 9.3, 1.2 Hz), 5.37 (dd, 1H, *J* = 2.4, 1.2 Hz), 4.80 (s, 2H), 3.8~4.0 (m, 4H), 3.69 (q, 2H, *J* = 7.2 Hz), 1.46 (s, 3H), 1.41 (s, 3H), 1.20 (t, 3H, *J* = 7.2 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ 136.82, 118.62, 98.17, 90.70, 72.85, 66.61, 63.72, 24.57, 22.97, 15.26.

Compound **29**
A mixture of compound **28** (216 mg, 1.0 mmol), NMO (176 mg, 1.5 mmol) and OsO4 (2.5% solution in *t*-BuOH, 0.30 mL, 0.03 mmol) in THF/H_2O (5 mL/5 mL) was stirred at rt overnight.

The mixture was poured into brine and extracted with ethylacetate. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was treated with $NaIO_4$ (0.32 g, 1.5 mmol) in CH_2Cl_2/H_2O (10 mL/10 mL) and stirred for 1h. The mixture was poured into brine and extracted with ethylacetate. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:4) to give a colorless liquid (162 mg, 74%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 9.66 (s, 1H), 4.86 (s, 2H), 4.10 (d, 2H, *J* = 12.3 Hz), 3.92 (d, 2H, *J* = 12.3 Hz), 3.66 (q, 2H, *J* = 6.9 Hz), 1.43 (s, 3H), 1.41 (s, 3H), 1.17 (t, 3H, *J* = 6.9 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ 201.69, 98.79, 91.38, 77.15, 64.35, 62.22, 24.79, 22.35, 15.10.

Compound **30**

To a suspension of $LiAlH_4$ (70 mg, 1.84 mmol) in THF (5 mL) was added a solution of compound **29** (270 mg, 1.24 mmol) in THF (2 mL) at 0 °C and stirred for 1 hour. The mixture was quenched with aqueous saturated NH_4Cl solution and extracted with ethyl acetate. The organic layer was dried over sodium sulfate, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:1.5) to give a colorless syrup (230 mg, 84%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.83 (s, 2H), 3.78 (s, 4H), 3.68 (dd, 4H, *J* = 14.0, 7.0 Hz), 3.22 (t, 1H, *J* = 7.0 Hz), 1.43 (s, 3H), 1.39 (s, 3H), 1.23 (t, 3H, *J* = 7.0 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ 98.93, 89.83, 74.07, 64.27, 63.76, 63.36, 24.30, 23.09, 15.27.

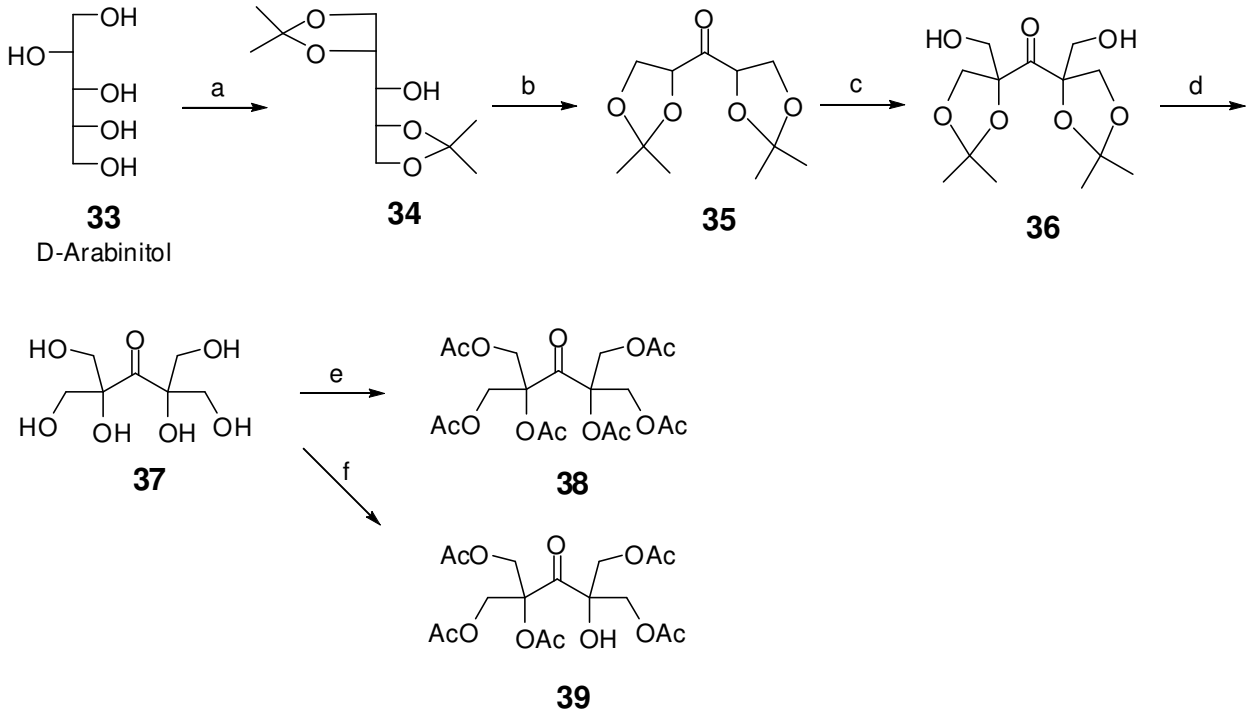
Compound **31**

A mixture of compound **30** (200 mg, 0.91 mmol) and Dowex 500WX8-400 (300 mg) in H_2O (10 mL) was stirred at 70 °C for 1h. The mixture was filtered through Celite and concentrated. The residue was purified by flash chromatography (silica, MeOH:CH_2Cl_2 = 1:5 to 1:2) to give a colorless syrup (110 mg, 99%). ^{1}H NMR (300 MHz, D_2O) δ 3.43 (s, 6H). ^{13}C NMR (75 MHz, D_2O) δ 75.59, 62.80.

Compound **32**

A mixture of compound **31** (100 mg, 0.83 mmol), acetic anhydride (0.3 mL) and DMAP (10 mg) in pyridine (3 mL) was stirred at rt for 24h. The mixture was evaporated and purified by flash chromatography (silica, EA:Hx = 1:2) to give a colorless syrup (190 mg, 78%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.44 (s, 6H), 2.07 (s, 9H), 2.05 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.35, 169.98, 80.33, 61.59, 21.85, 20.87.

Figure S4f.
Synthesis of **C7k** and reduced **C7k**.



a. 2,2-dimethoxypropane, CSA, THF, reflux; b. PDC, Ac_2O, CH_2Cl_2, reflux, 38 % for two steps; c. CH_2O, MeOH, reflux, 38%; d. Dowex 50WX8-400, H_2O, 66%; e. Ac_2O, pyridine, DMAP; f. Ac_2O, pyridine

Compound **35**
A mixture of arabinitol (2.5 g, 16.4 mmol), dimethoxypropane (8 mL, 65 mmol) and camphorsulfonic acid (1.13 g, 4.9 mmol) in THF (100 mL) was refluxed for 15 min. It was poured into water and extracted with ether. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, ethyl acetate:hexanes = 1:5 to 1:2) to give a colorless liquid. It was refluxed with PDC (2.9 g, 7.7 mmol) and acetic anhydride (3.6 mL, 38 mmol) in dichloromethane (100 mL) for 2h and filtered through Celite. The filtrate was evaporated and purified by flash chromatography (silica, ethyl acetate:hexanes = 1:10 to 1:3.5) to give a colorless liquid (1.45 g, 38%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.80 (dd, 2H, *J* = 8.0, 5.5 Hz), 4.26 (dd, 2H, *J* = 9.0, 8.0 Hz), 4.04 (dd, 2H, *J* = 9.0, 6.0 Hz). ^{13}C NMR (75 MHz, $CDCl_3$) δ 206.51, 111.09, 78.68, 66.24, 25.99, 25.00.

Compound **36**
A mixture of compound **35** (866 mg, 3.76 mmol), formaldehyde (37% solution, 1.8 mL, 22.6 mmol) and potassium carbonate (300 mg) in methanol (50 mL) was refluxed for 4h. It was neutralized by addition of acetic acid (0.27 mL) and poured into water. The product was extracted with ethyl acetate and purified by flash chromatography (silica, ethyl acetate:hexanes = 1:1 to EA 100%) to give a colorless liquid (411 mg, 38%) as a mixture of two compounds. ^{13}C NMR (75 MHz, $CDCl_3$) δ 209.73, 209.47, 111.97, 111.89, 90.31, 89.60, 69.06, 68.43, 66.93, 66.85, 27.32, 27.30, 25.69, 25.67.

Compound **37**

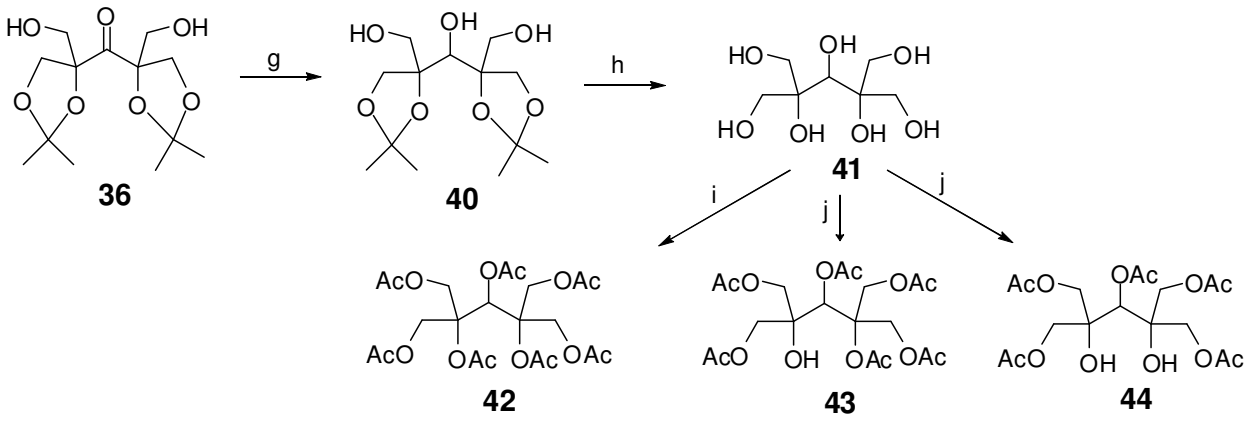
A mixture of compound **36** (200 mg) and Dowex 500WX8-400 (300 mg) in H_2O (10 mL) was stirred at 70 °C for 1h. The mixture was filtered through Celite and concentrated to give a colorless syrup (120 mg, 83%). ^{1}H NMR (300 MHz, D_2O) δ 3.80 (d, 4H, *J* = 11.7 Hz), 3.51 (d, 4H, *J* = 11.7 Hz). ^{13}C NMR (75 MHz, D_2O) δ 215.11, 86.14, 64.62.

Compound **38**

A mixture of compound **37** (40 mg, 0.19 mmol), acetic anhydride (0.35 mL) and DMAP (10 mg) in pyridine (3.5 mL) was stirred at rt for 48h. The mixture was evaporated and purified by flash chromatography (silica, ethyl acetate:hexanes = 1:1) to give a colorless syrup (79 mg, 90%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.76 (d, 4H, *J* = 11.7 Hz), 4.54 (d, 4H, *J* = 12 Hz), 2.14 (s, 6H), 2.04 (s, 12H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 199.32, 170.09, 168.96, 87.19, 62.84, 21.42, 20.77.

Compound **39**

A mixture of compound **37** (40 mg, 0.19 mmol) and acetic anhydride (0.35 mL) in pyridine (3.5 mL) was stirred at rt for 24h. The mixture was evaporated and purified by flash chromatography (silica, ethyl acetate:hexanes = 1:1) to give a colorless syrup (70 mg, 88%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.67 (dd, 4H, *J* = 18.3, 12 Hz), 4.33 (s, 4H), 2.10 (br s, 9H), 2.06 (s, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 202.89, 171.59, 170.32, 84.72, 82.20, 66.74, 66.07, 62.22, 60.62, 20.89.



g. $LiAlH_4$, THF, 78%; h. Dowex 50WX8-400, H_2O, 66%; i. Ac_2O, pyridine, DMAP; j. Ac_2O, pyridine

Compound **40**

To a suspension of $LiAlH_4$ (70 mg) in THF (5 mL) was added compound **36** (150 mg, 0.52 mmol) in THF (2 mL) at 0 °C and stirred for 1h. An ammonium chloride solution was added into the mixture and extracted with ethyl acetate. The organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, ethyl acetate:hexanes = 1.5:1 to 3:1) to give a colorless liquid (117 mg, 78%) as a mixture of three isomers. ^{13}C NMR (75 MHz, $CDCl_3$) δ 111.38, 111.09, 110.36, 110.18, 85.02, 84.82, 84.66, 82.62, 76.90, 75.15, 74.13, 72.59, 69.93, 68.78, 68.65, 65.07, 64.26, 64.21, 64.18, 28.11, 27.79, 27.54, 26.48, 26.11.

Compound **41**

A mixture of compound **40** (100 mg) and Dowex 500WX8-400 (200 mg) in H_2O (6 mL) was stirred at 70 °C for 1h. The mixture was filtered through Celite and concentrated to give a colorless syrup (70 mg, 96%). ^{13}C NMR (75 MHz, D_2O) δ 77.45, 72.34, 64.34, 62.69.

Compound **42**

A mixture of compound **41** (23 mg), acetic anhydride (0.3 mL) and DMAP (10 mg) in pyridine (3 mL) was stirred at rt for 72h. The mixture was evaporated and purified by flash chromatography (silica, ethyl acetate:hexanes = 1:1) to give a colorless syrup (52 mg, 95%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 6.01 (s, 1H), 4.68 (d, 2H, *J* = 12 Hz), 4.58 (dd, 4H, *J* = 15, 12.3 Hz), 2.09 (s, 3H), 2.07 (br s, 12H), 2.06 (s, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.26, 170.20, 169.24, 169.10, 83.27, 70.13, 62.62, 62.52, 21.95, 20.88, 20.74.

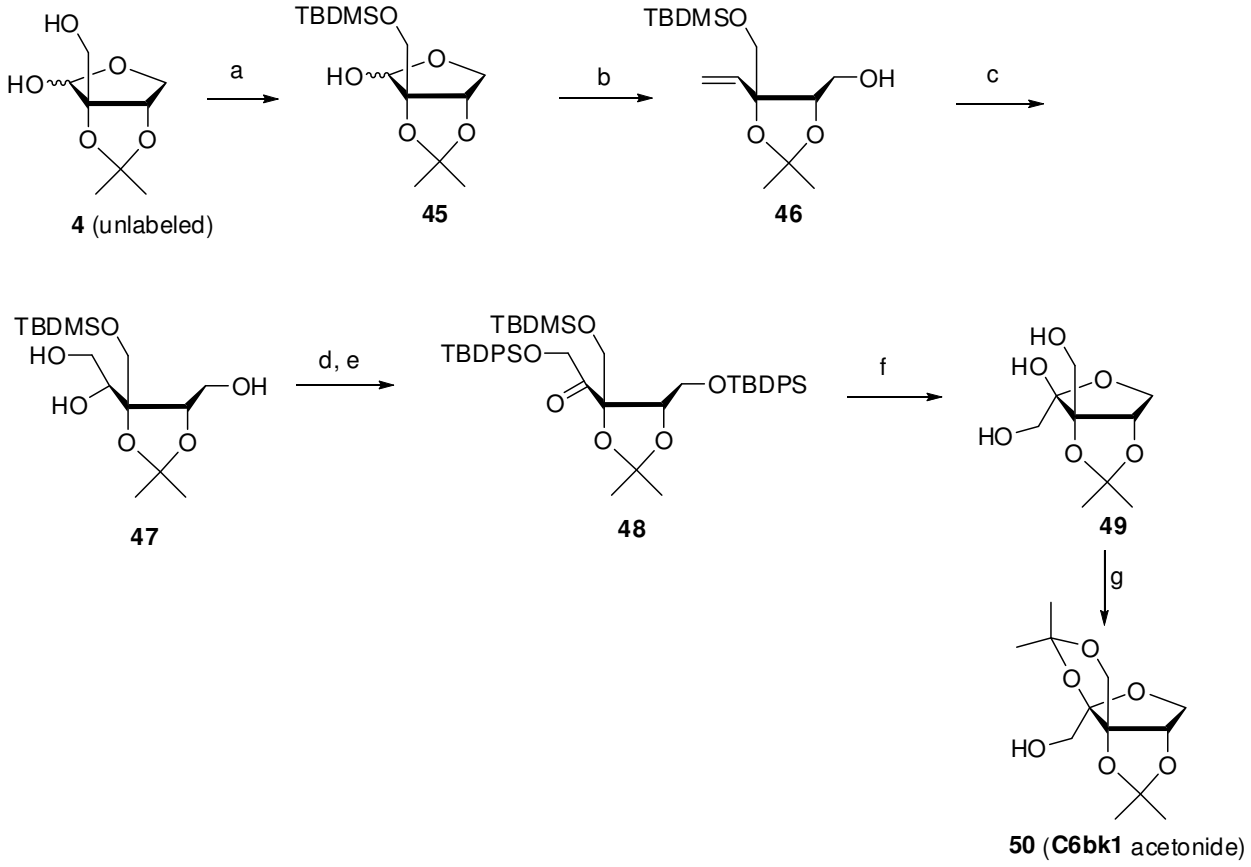
Compound **43**

A mixture of compound **41** (20 mg) and acetic anhydride (0.15 mL) in pyridine (2 mL) was stirred at rt for 7days. The mixture was evaporated and purified by flash chromatography (silica, ethyl acetate:hexanes = 2:1) to give a colorless syrup (33 mg, 75%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.64 (s, 1H), 4.0~4.8 (m, 8H), 2.11 (s, 6H), 2.10 (s, 3H), 2.09 (s, 3H), 2.07 (s, 3H), 2.06 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.23, 170.21, 169.34, 83.19, 75.66, 70.71, 64.99, 64.66, 62.67, 62.16, 22.03, 20.99.

Compound **44**

A mixture of compound **41** (20 mg) and acetic anhydride (0.15 mL) in pyridine (2 mL) was stirred at rt for 24h. The mixture was evaporated and purified by flash chromatography (silica, ethyl acetate:hexanes = 2:1) to give a colorless syrup (28 mg, 70%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.28 (s, 1H), 4.38 (d, 2H, *J* = 11.7 Hz), 4.26 (dd, 4H, *J* = 20.7, 12 Hz), 4.14 (d, 2H, *J* = 12 Hz), 2.10 (s, 9H), 2.09 (s, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.30, 171.19, 169.52, 75.98, 71.33, 65.83, 64.75, 20.98, 20.85.

Figure S4g.
Synthesis of **C6bk1**



a. TBDMSCl, Imidazole, CH_2Cl_2, 66%; b. CH_3PPh_3Br, *n*-BuLi, THF, 82%; c. OsO_4, NMO, THF/H_2O, 53%; d. TBDPSCl, Imidazole, CH_2Cl_2; e. PDC, Ac_2O, CH_2Cl_2, 58% for two steps; f. TBAF, THF, 59%; g. Acetone, H_2SO_4.

Compound **45**
To a mixture of compound **4** (649 mg, 3.41 mmol) and imidazole (464 mg, 6.82 mmol) in dichloromethane (10 mL) was added TBDMSCl (540 mg, 3.58 mmol). The mixture was stirred at rt for 2h and quenched with water and extracted with dichloromethane. The combined organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:5 to 1:3) to give a colorless syrup (α/β mixture, 683 mg, 66%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.31 (d, 0.5H, J = 7.2 Hz), 4.95 (d, 0.5H, J = 11.7 Hz), 4.66 (m, 1H), 4.38 (dd, 0.5H, J = 7.2, 0.9 Hz), 3.7~4.1 (m, 4H), 3.52 (dd, 0.5H, J = 10.8, 3.3 Hz), 1.54 (s, 1.5H), 1.47 (s, 1.5H), 1.42 (s, 1.5H), 1.37 (s, 1.5H), 0.90 (s, 4.5H), 0.88 (s, 4.5H), 0.123 (s, 1.5H), 0.119 (s, 1.5H), 0.074 (s, 1.5H), 0.072 (s, 1.5H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 113.72, 113.61, 105.05, 97.91, 92.22, 88.47, 83.59, 82.98, 72.29, 68.23, 64.46, 63.28, 27.66, 27.26, 26.96, 25.97, 25.94, 18.36, 18.33, -5.44 (broad peak).

Compound **46**
To a suspension of CH_3PPh_3Br (3.98 g, 11.15 mmol) in THF (100 mL) was added *n*-BuLi (1.6M solution, 6.3 mL, 10.04 mmol) at 0 °C and stirred for 20 min. A solution of compound **45** (680 mg, 2.23 mmol) in THF (5 mL) was added to the mixture and stirred at rt for 3h. The reaction

mixture was quenched with brine and extracted with ethyl acetate. The combined organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:5 to 1:2) to give a colorless syrup (550 mg, 82%). ^{1}H NMR (300 MHz, $CDCl_3$) δ 5.92 (dd, 1H, *J* = 17.4, 10.8 Hz), 5.48 (dd, 1H, *J* = 17.1, 2.1 Hz), 5.26 (dd, 1H, *J* = 10.8, 1.8 Hz), 4.28 (t, 1H, *J* = 6.0 Hz), 3.70 (td, 2H, *J* = 6.0, 1.8 Hz), 3.63 (dd, 2H, *J* = 13.2, 9.6 Hz), 2.37 (t, 1H, *J* = 5.7 Hz), 1.51 (s, 3H), 1.43 (s, 3H), 0.90 (s, 9H), 0.08 (s, 3H), 0.07 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 135.74, 116.46, 109.15, 83.98, 81.99, 68.85, 62.32, 28.05, 26.83, 26.05, 18.50, -5.32, -5.45.

Compound **47**

A mixture of compound **46** (680 mg, 2.25 mmol), NMO (395 mg, 3.38 mmol) and OsO4 (2.5% solution in *t*-BuOH, 1.15 mL, 0.1125 mmol) in THF/H_2O (20 mL/20 mL) was stirred at rt for 2 days. The mixture was poured into brine and extracted with ethyl acetate. The combined organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:1) to give a colorless solid (403 mg, 53%) as a mixture of two diastereomers (ratio, ~5.5:1). Spectral data shown below are for major compounds only. ^{1}H NMR (300 MHz, $CDCl_3$) δ 3.7~4.1 (m, 10H), 3.43 (dd, 1H, *J* = 9.6, 3.9 Hz), 3.25 (dd, 1H, *J* = 8.7, 6.0 Hz), 1.39 (s, 3H), 1.37 (s, 3H), 0.91 (s, 9H), 0.13 (s, 3H), 0.12 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 108.90, 83.97, 81.18, 71.27, 65.75, 61.47, 60.79, 28.62, 26.88, 26.03, 18.55, -5.27, -5.37.

Compound **48**

To a mixture of compound **47** (200 mg, 0.59 mmol) and imidazole (121 mg, 1.78 mmol) in dichloromethane (5 mL) was added TBDPSCl (327 mg, 1.19 mmol). The mixture was stirred at rt for 2h and quenched with water and extracted with dichloromethane. The combined organic layer was dried over Na_2SO_4, filtered and concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:15) to give a colorless syrup. It was mixed with acetic anhydride (116 mg, 1.14 mmol) and PDC (86 mg, 0.23 mmol) in dichloromethane (15 mL) and refluxed for 3h. After cooling to rt, the mixture was filtered through Celite and the filtrate was concentrated. The residue was purified by flash chromatography (silica, Ethylacetate:Hexanes = 1:15) to give a colorless syrup (279 mg, 58% for two steps). ^{1}H NMR (300 MHz, $CDCl_3$) δ 7.2~7.8 (m, 20H), 4.67 (d, 1H, *J* = 19.5 Hz), 4.53 (d, 1H, *J* = 19.5 Hz), 4.14 (dd, 1H, *J* = 6.0, 3.3 Hz), 3.80 (dd, 1H, *J* = 11.7, 3.6 Hz), 3.4~3.6 (m, 3H), 1.34 (s, 3H), 1.31 (s, 3H), 1.061 (s, 9H), 1.058 (s, 9H), -0.017 (s, 6H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 208.65, 135.97, 135.84, 135.73, 133.66, 133.45, 133.39, 129.89, 127.85, 110.06, 89.53, 79.79, 69.79, 65.41, 62.32, 27.04, 26.94, 26.56, 26.14, 19.55, 19.42, 18.64, -5.2, -5.3.

Compound **49**

A mixture of compound **48** (279 mg, 0.344 mmol) and TBAF (1M solution in THF, 1.38 mL, 1.38 mmol) in THF (20 mL) was stirred at rt for 3h. The mixture was concentrated and the residue was purified by flash chromatography (silica, Ethylacetate 100% to Ethylacetate:MeOH = 20:1) to give a white solid (45 mg, 59%). ^{1}H NMR (300 MHz, CD_3OD) δ 4.63 (d, 1H, *J* = 3.6 Hz), 4.02 (dd, 1H, *J* = 10.2, 3.9 Hz), 3.7~3.9 (m, 4H), 3.59 (d, 1H, *J* = 11.1 Hz), 1.44 (s, 3H), 1.41 (s, 3H). ^{13}C NMR (75 MHz, CD_3OD) δ 114.84, 107.43, 95.69, 85.27, 71.97, 64.64, 63.72, 28.35, 27.82.

Compound **50**

A solution of compound **49** (22 mg, 0.10 mmol) in acetone (5 mL) was treated with sulfuric acetic (0.1 mL) and stirred at rt for 1.5h. It was neutralized by sodium bicarbonate and

evaporated and purified by silica gel column chromatography (EA:Hx=1:1) to give a colorless syrup (21 mg, 80%). The ^{1}H and ^{13}C NMR of compound **50** were identical to the spectral data obtained from the reaction product of DHA and HCHO after acetonide derivatization. ^{1}H NMR (300 MHz, $CDCl_3$) δ 4.83 (d, 1H, *J* = 3.6 Hz), 4.25 (d, 1H, *J* = 9.9 Hz), 4.06 (d, 1H, *J* = 10.2 Hz), 3.6~4.0 (m, 4H), 2.51 (t, 1H, *J* = 6.8 Hz), 1.48 (s, 3H), 1.42 (s, 6H), 1.40 (s, 3H). ^{13}C NMR (75 MHz, $CDCl_3$) δ 114.18, 113.77, 112.32, 91.92, 83.06, 71.63, 70.06, 63.33, 27.85, 27.45, 26.55, 26.50.