Electronic Supplementary Material (ESI) for Organic & Biomolecular Chemistry.

Supplementary Information

Glycosylation of a model proto-RNA nucleobase with non-ribose sugars: Implications for the prebiotic synthesis of nucleosides

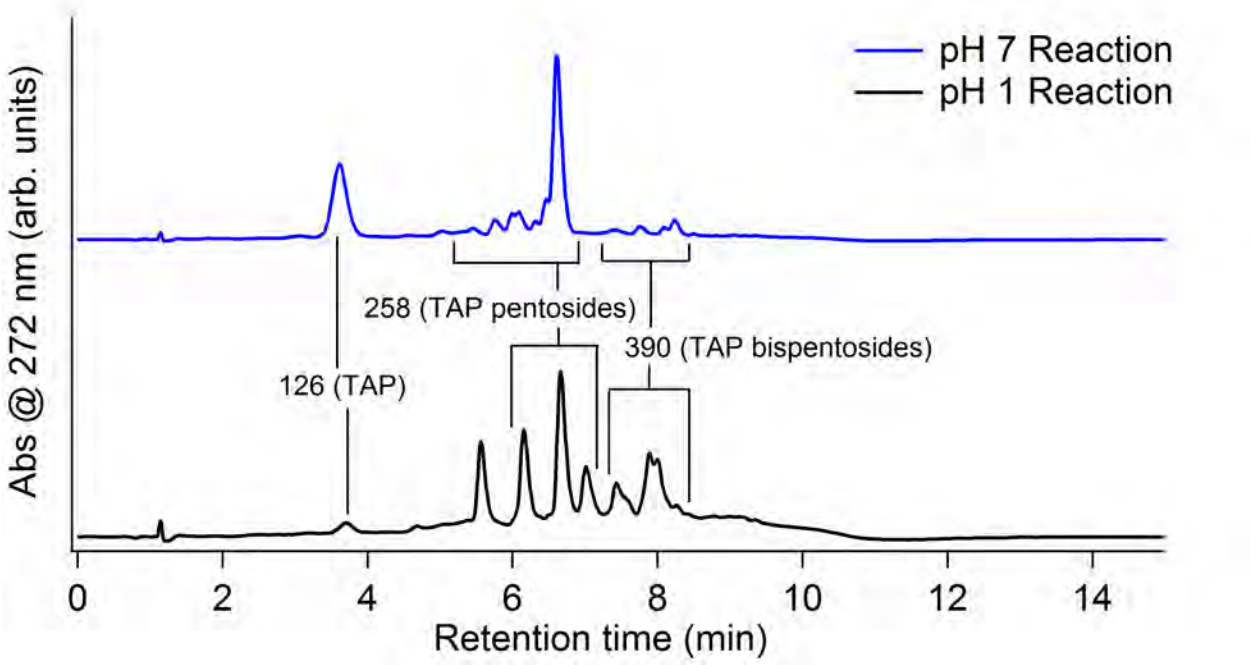
David M. Fialho,[a,b] Kimberly C. Clarke,[a,b] Megan K. Moore,[a,b] Gary B. Schuster,[a,b] Ramanarayanan Krishnamurthy[b,c] and Nicholas V. Hud[a,b]

[a]School of Chemistry and Biochemistry, Georgia Institute of Technology, Atlanta, GA, USA 30033
[b]NSF-NASA Center for Chemical Evolution
[c]Department of Chemistry, The Scripps Research Institute, La Jolla, CA, USA 92037

Table of Contents

I. Analysis of the Crude Reaction Products of TAP with Sugars

A.



B.

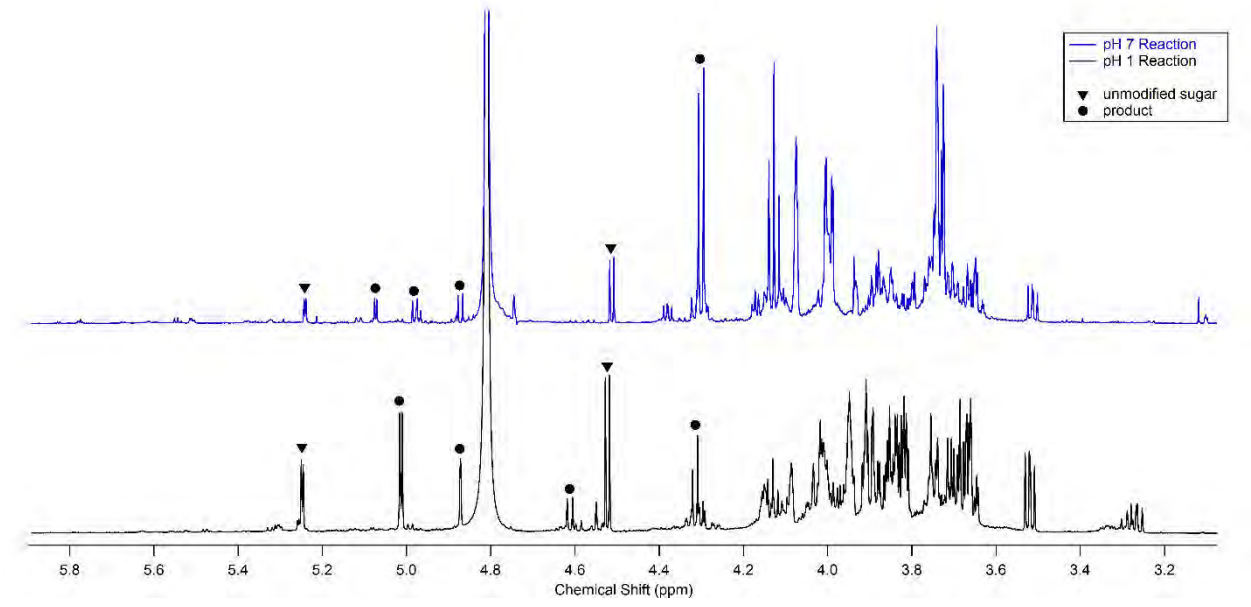
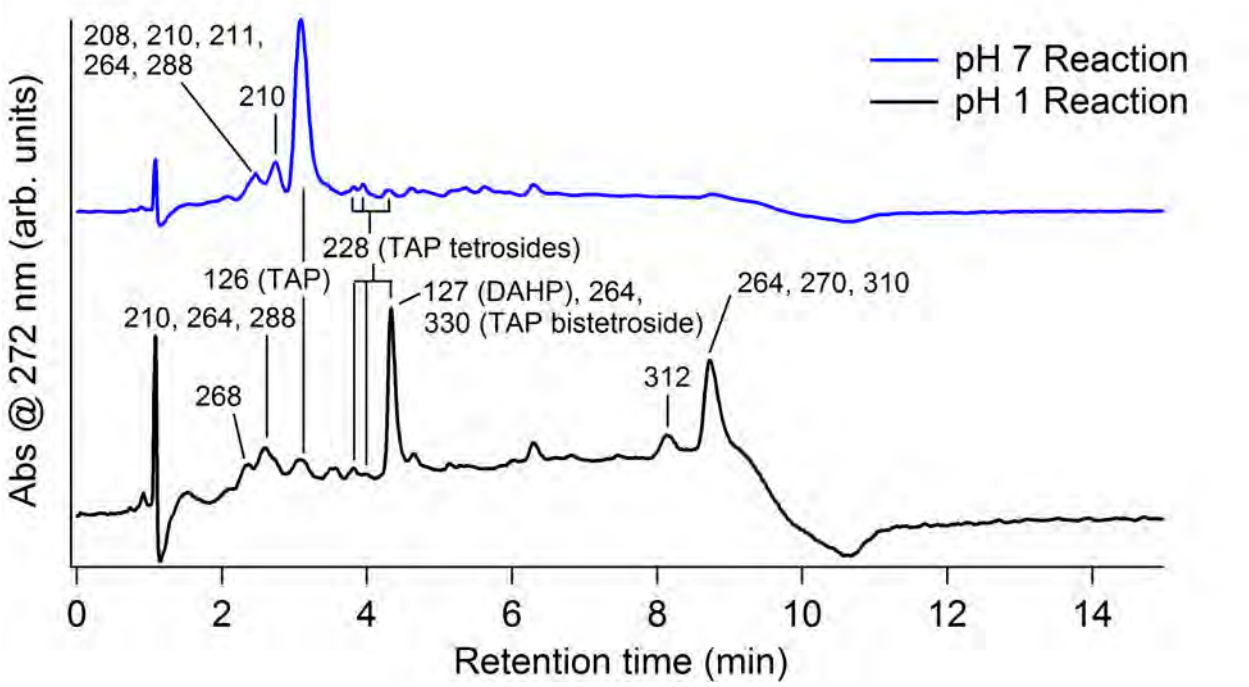
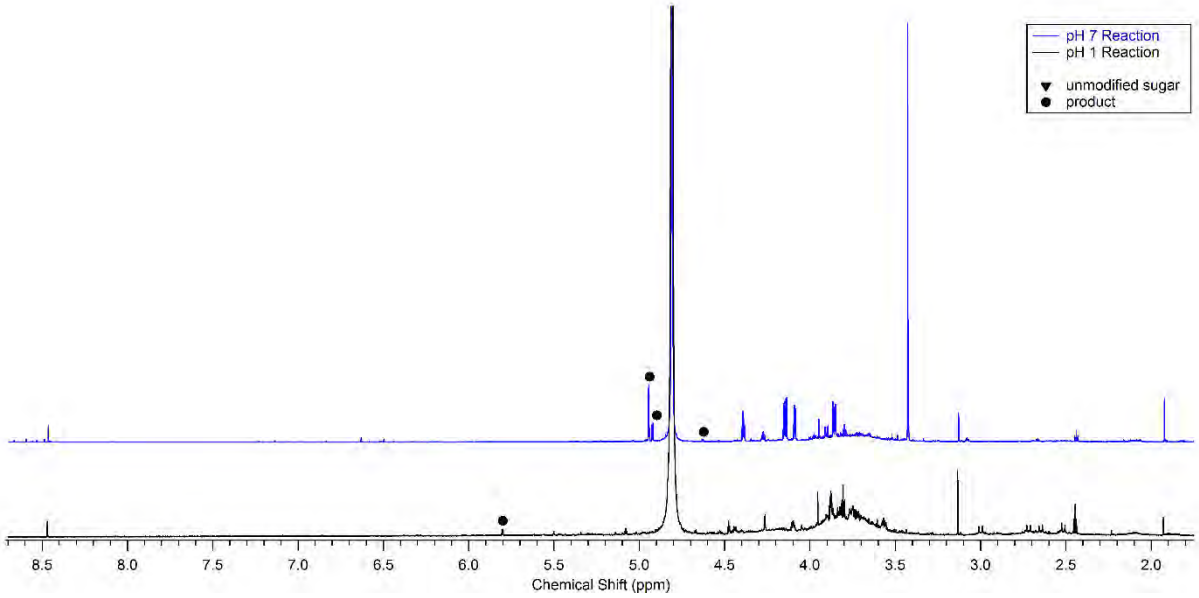


Fig. S1. Analysis of crude products from reactions of TAP with arabinose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.


208, 210, 211,
264, 288
210
pH 7 Reaction
pH 1 Reaction
228 (TAP tetrosides)
126 (TAP)
127 (DAHP), 264,
330 (TAP bistetroside)
264, 270, 310
210, 264, 288
268
312
Abs @ 272 nm (arb. units)
0
2
4
6
8
10
12
14
Retention time (min)

B.


pH 7 Reaction
pH 1 Reaction
unmodified sugar
product
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
Chemical Shift (ppm)

C.

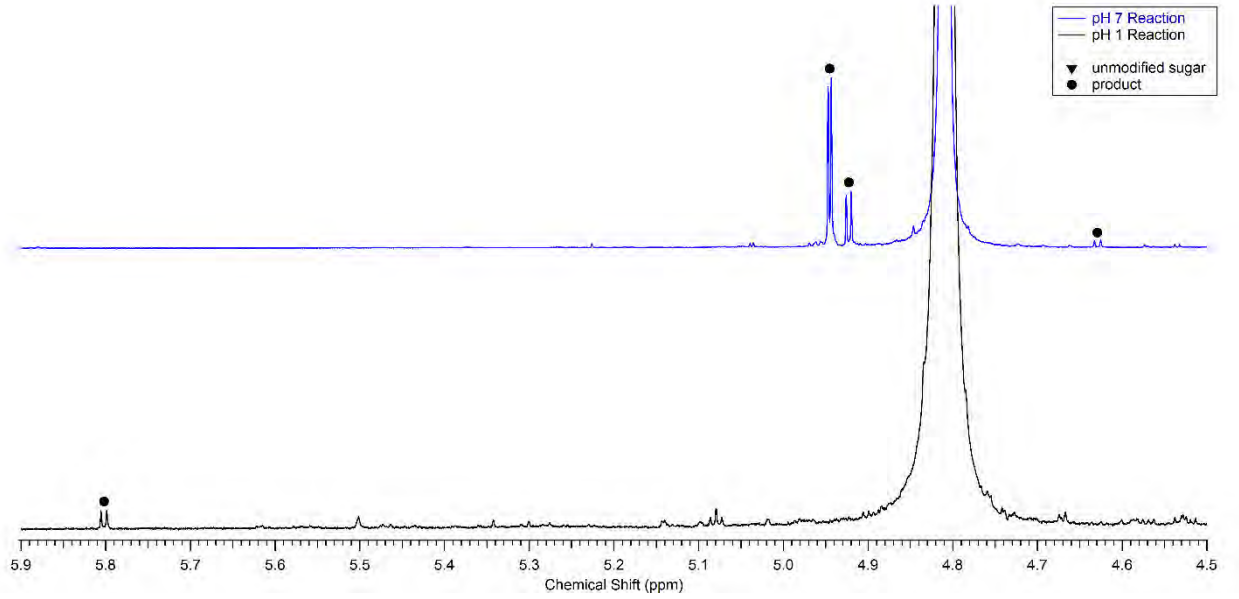
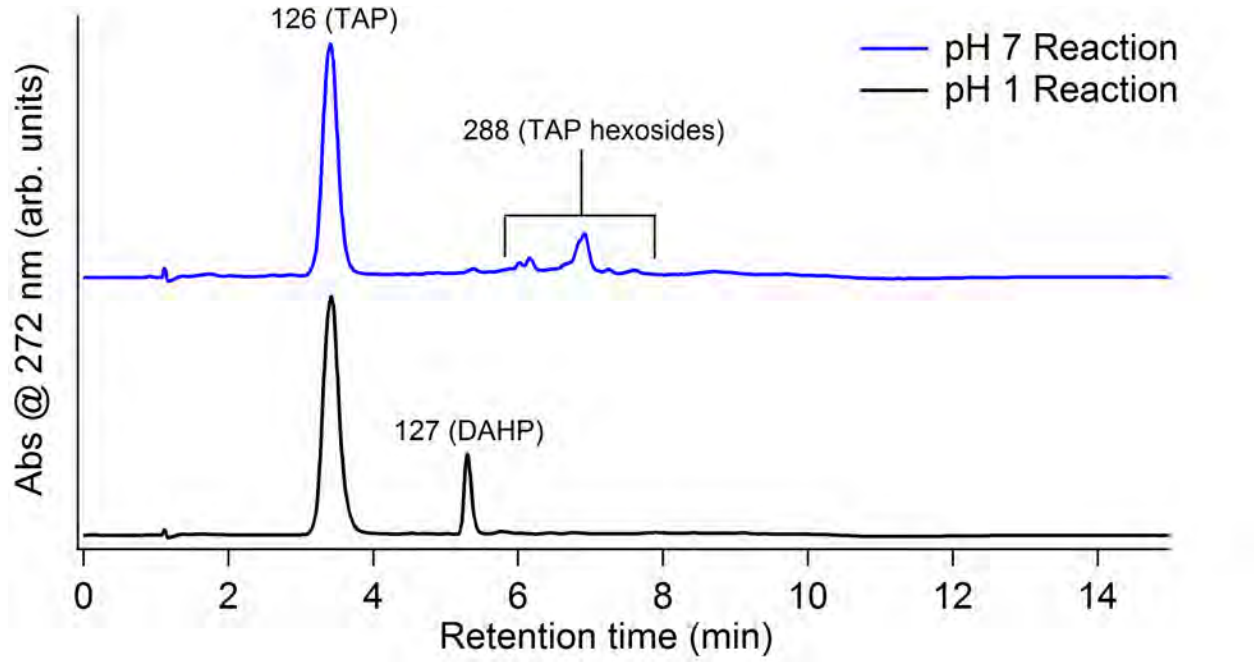


Fig. S2. Analysis of crude products from reactions of TAP with erythrose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. **(C)** Detail of the anomeric region of the ^{1}H NMR spectra.

A.



B.

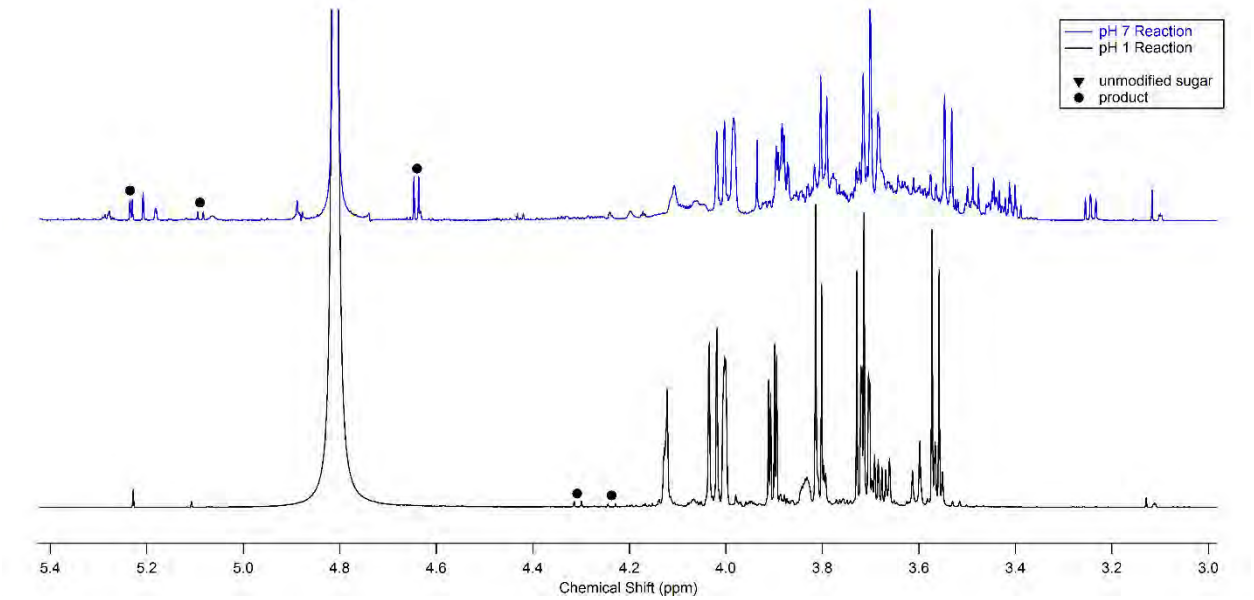
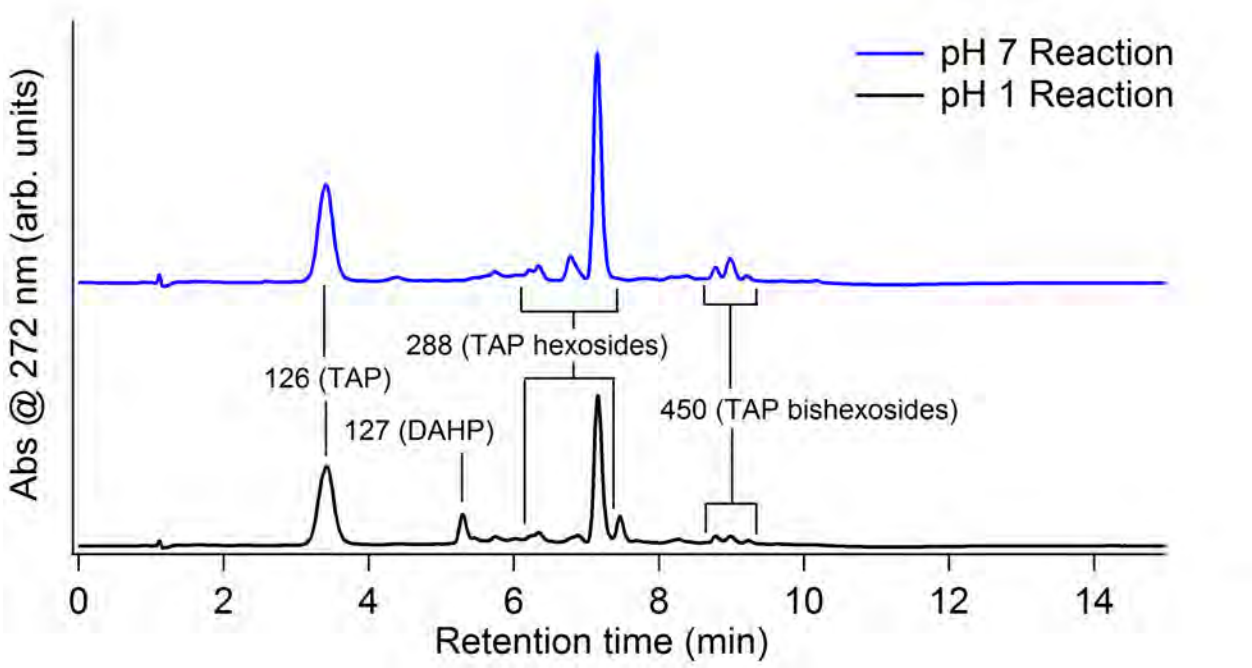


Fig. S3. Analysis of crude products from reactions of TAP with fructose**. (A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Note: Resonances that appear to be associated with TAP-aldose (glucose or mannose) glycosides in the pH 7 spectrum are located at 5.34, 5.20, and 4.75 ppm, and in the pH 1 spectrum are located at 4.42 and 4.35 ppm.

A.



B.

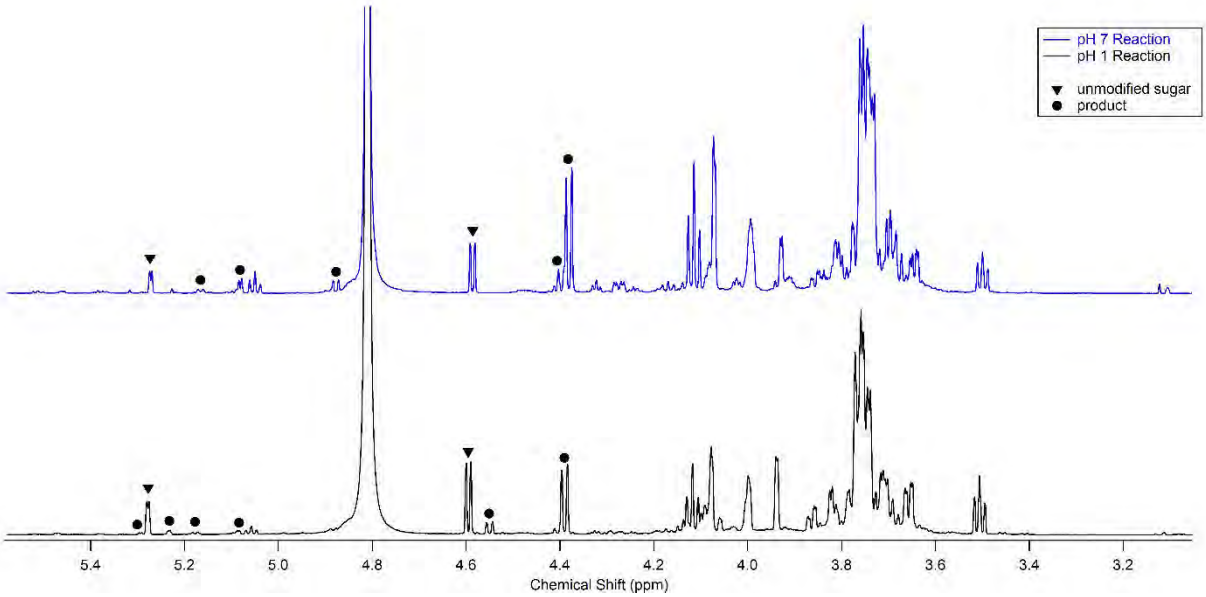
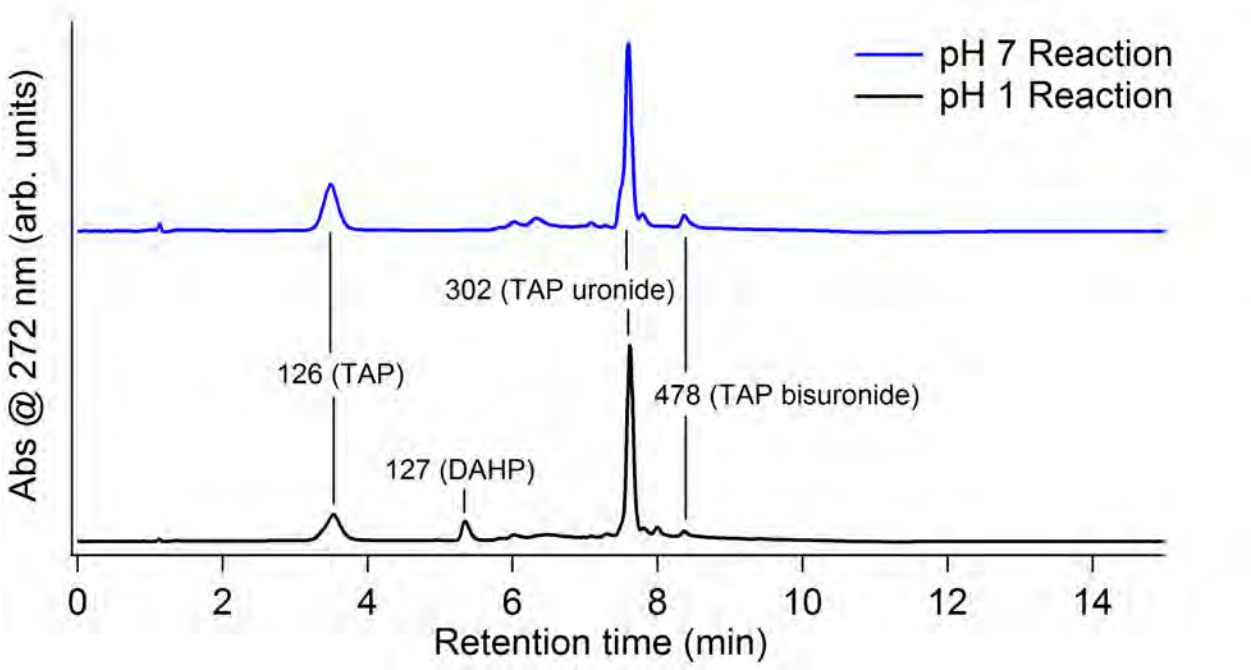


Fig. S4. Analysis of crude products from reactions of TAP with galactose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.



B.

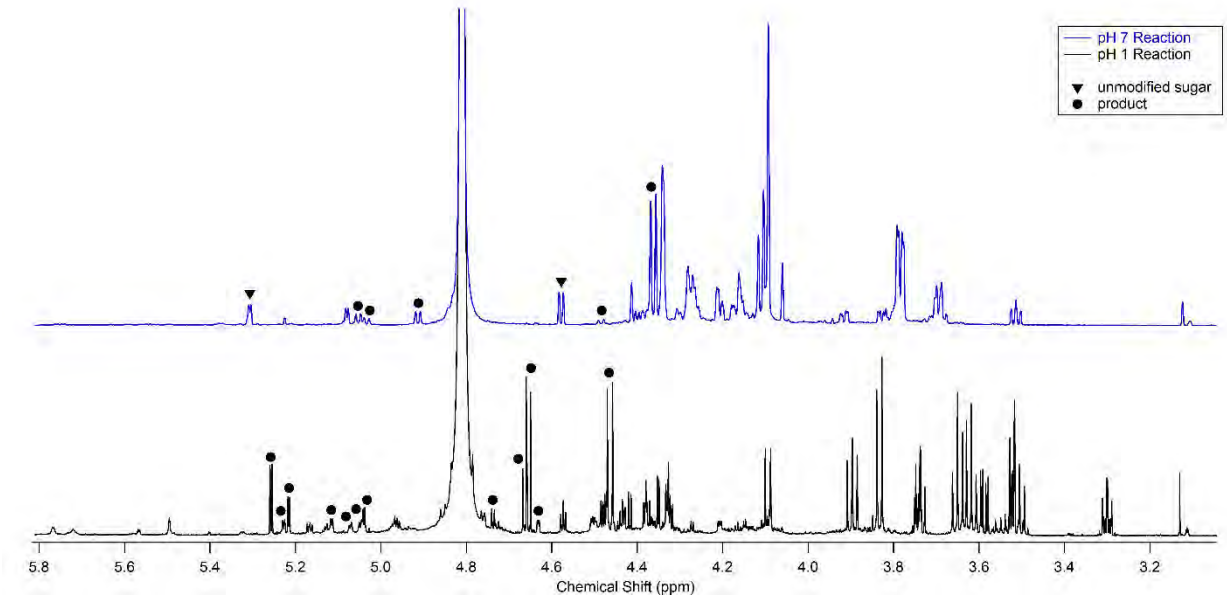
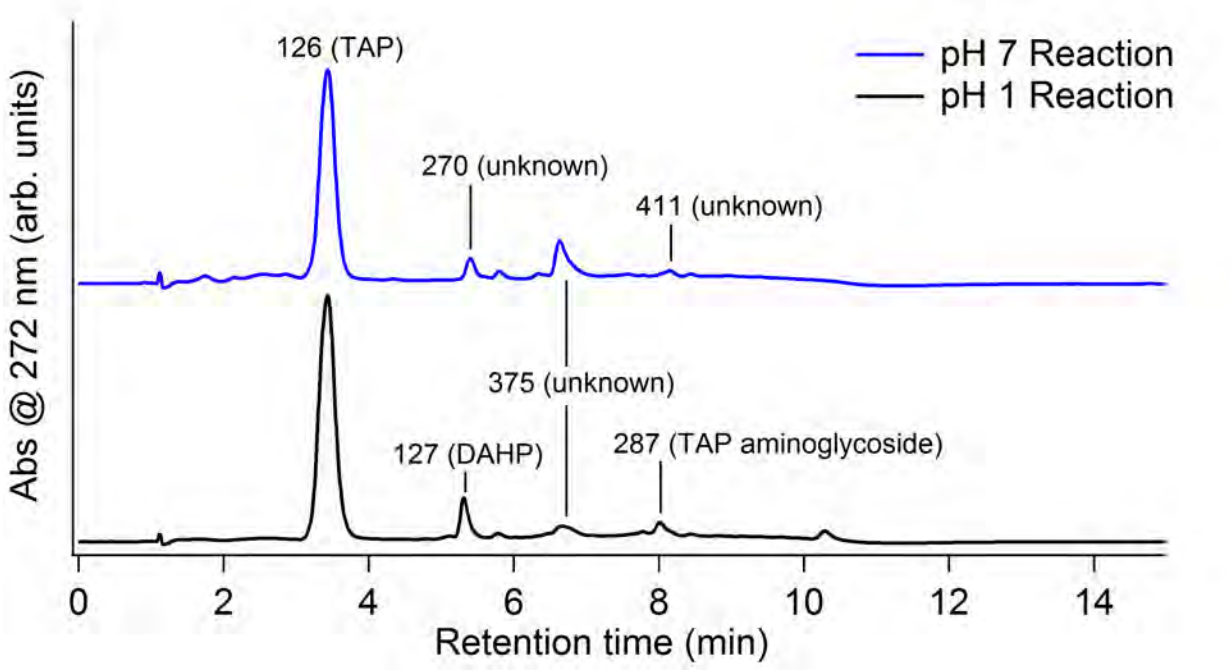
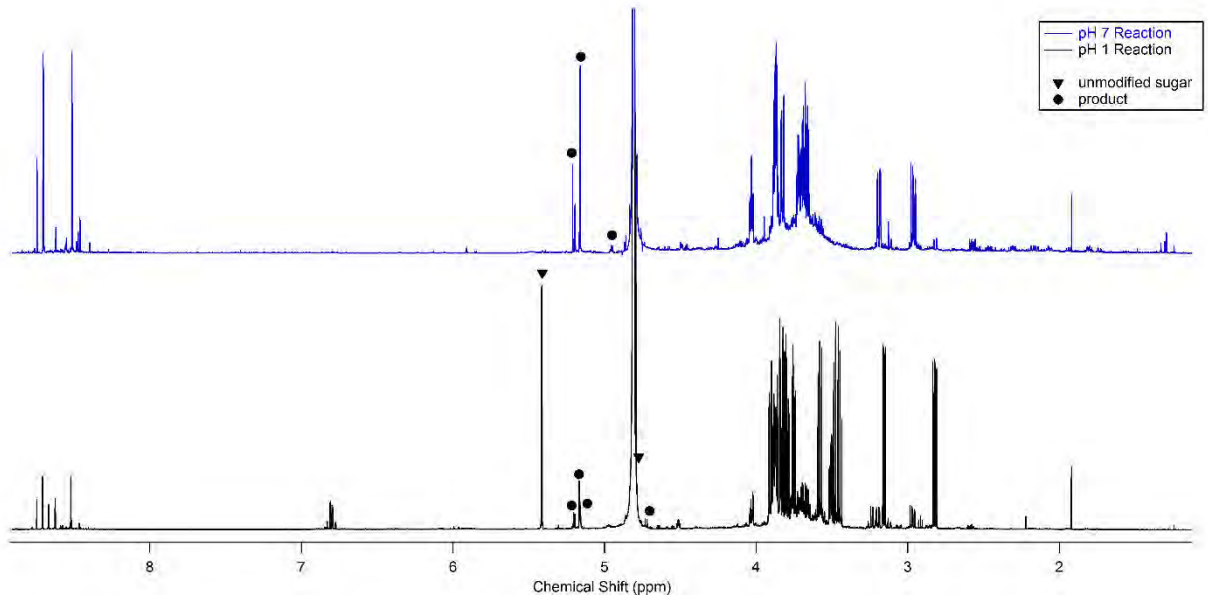


Fig. S5. Analysis of crude products from reactions of TAP with galacturonic acid. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.


126 (TAP)
270 (unknown)
411 (unknown)
375 (unknown)
127 (DAHP)
287 (TAP aminoglycoside)
pH 7 Reaction
pH 1 Reaction
Abs @ 272 nm (arb. units)
0
2
4
6
8
10
12
14
Retention time (min)

B.


pH 7 Reaction
pH 1 Reaction
unmodified sugar
product
8
7
6
5
4
3
2
Chemical Shift (ppm)

C.

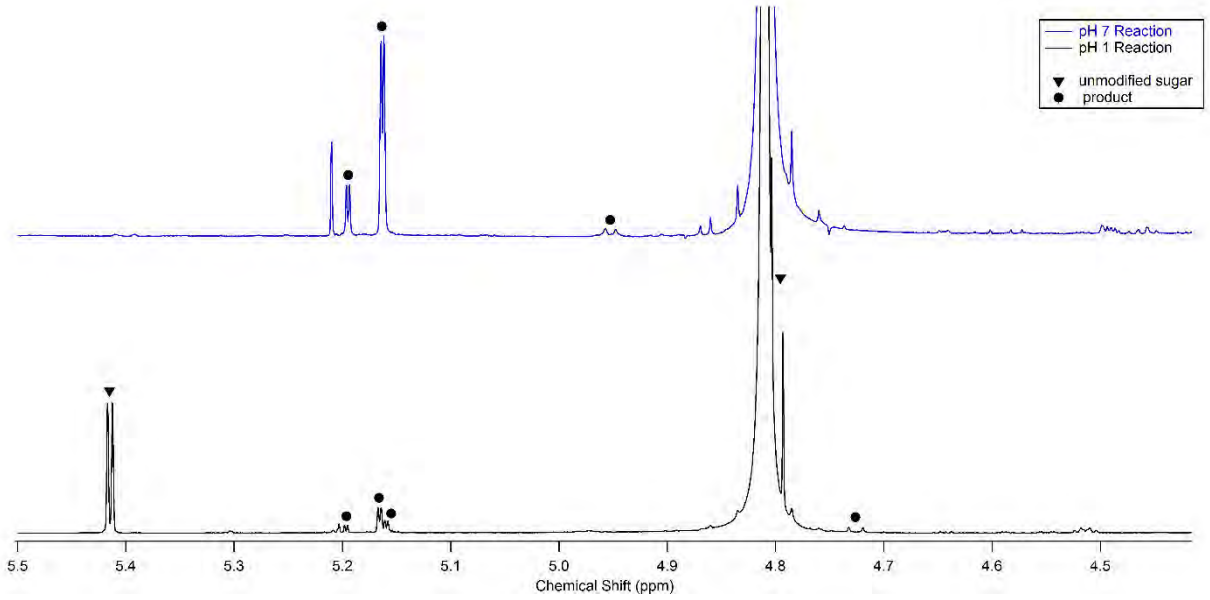
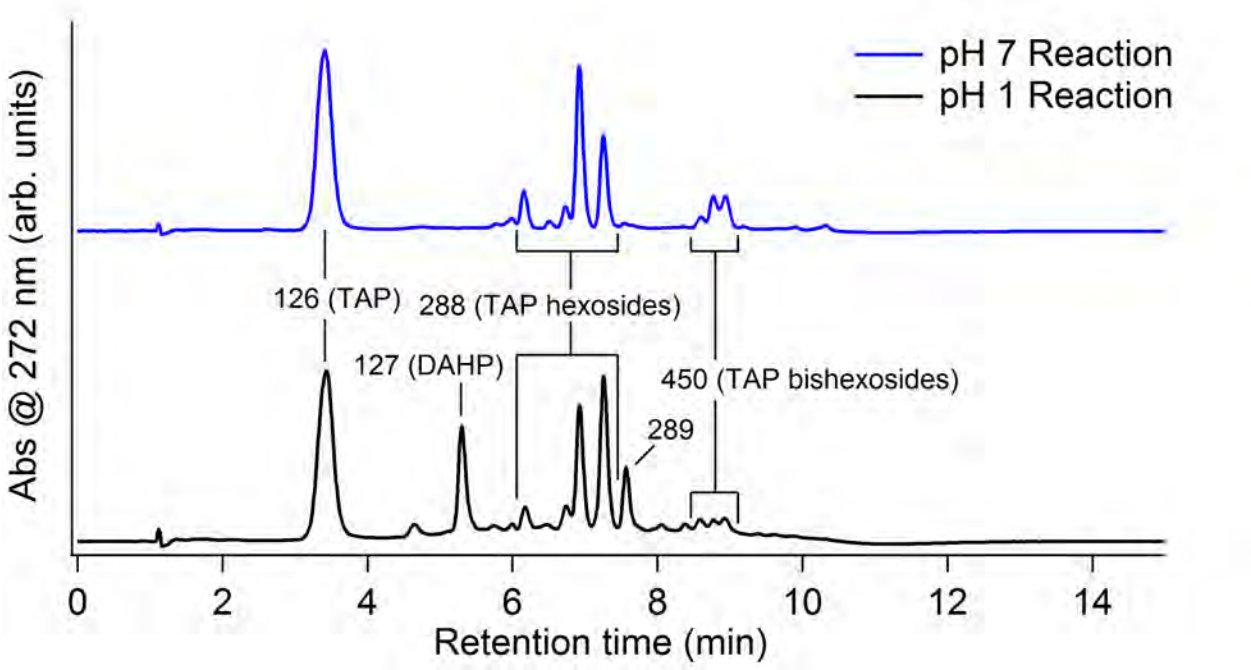


Fig. S6. Analysis of crude products from reactions of TAP with glucosamine. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. **(C)** Detail of the anomeric region of the ^{1}H NMR spectra.

A.



B.

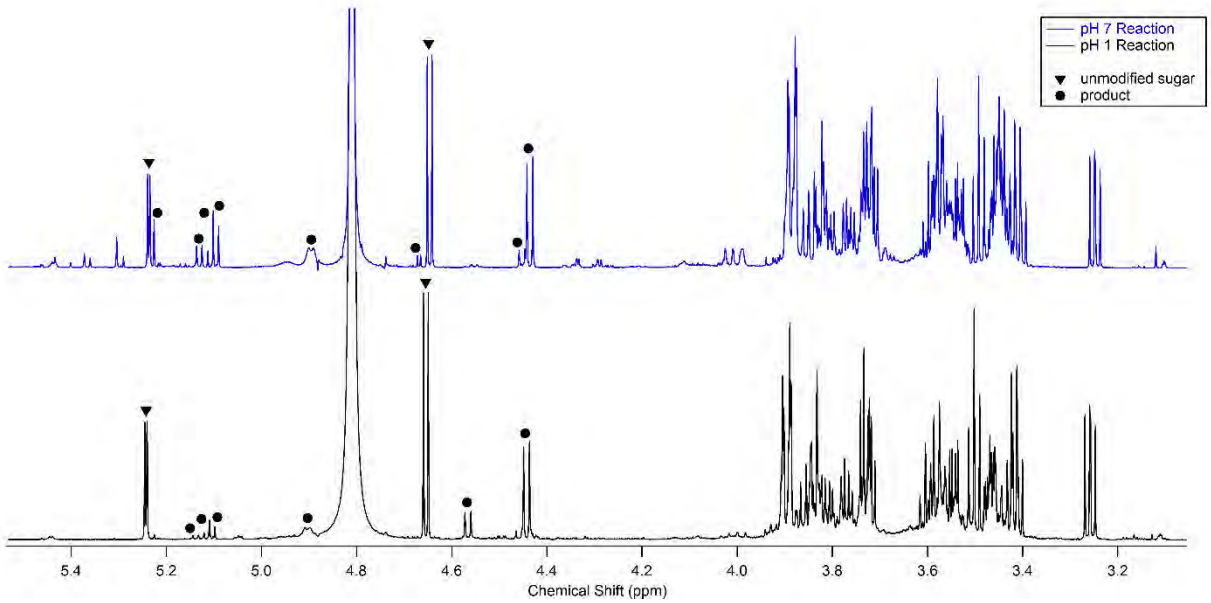
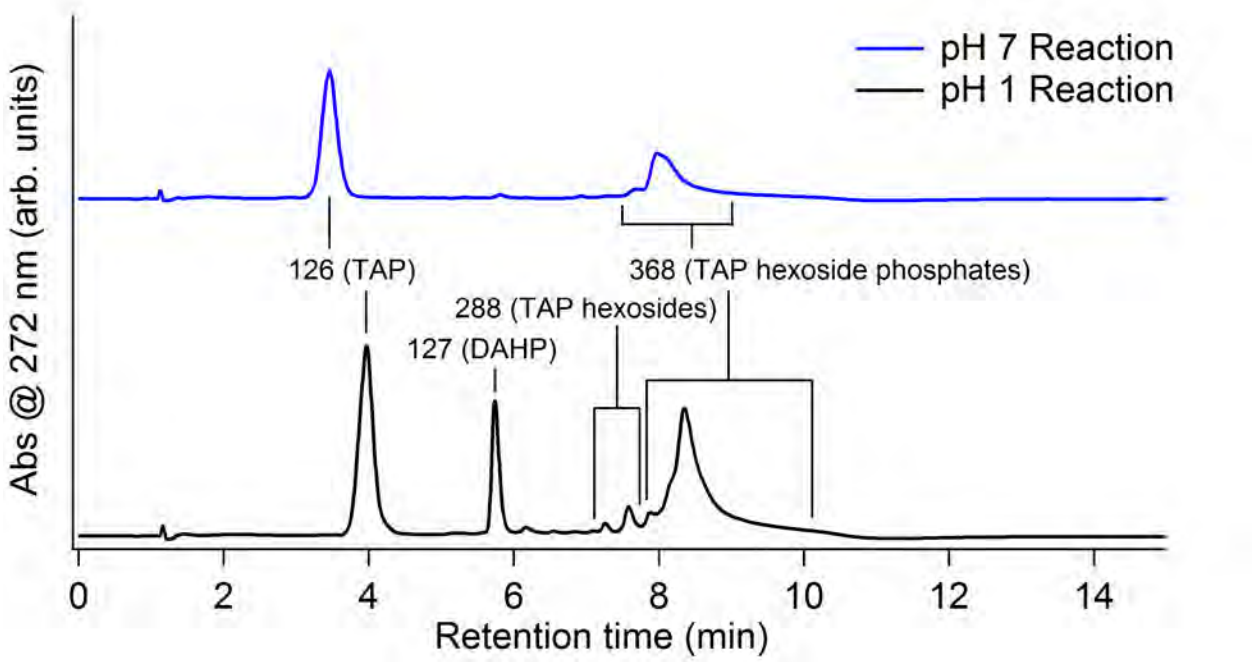


Fig. S7. Analysis of crude products from reactions of TAP with glucose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.



B.

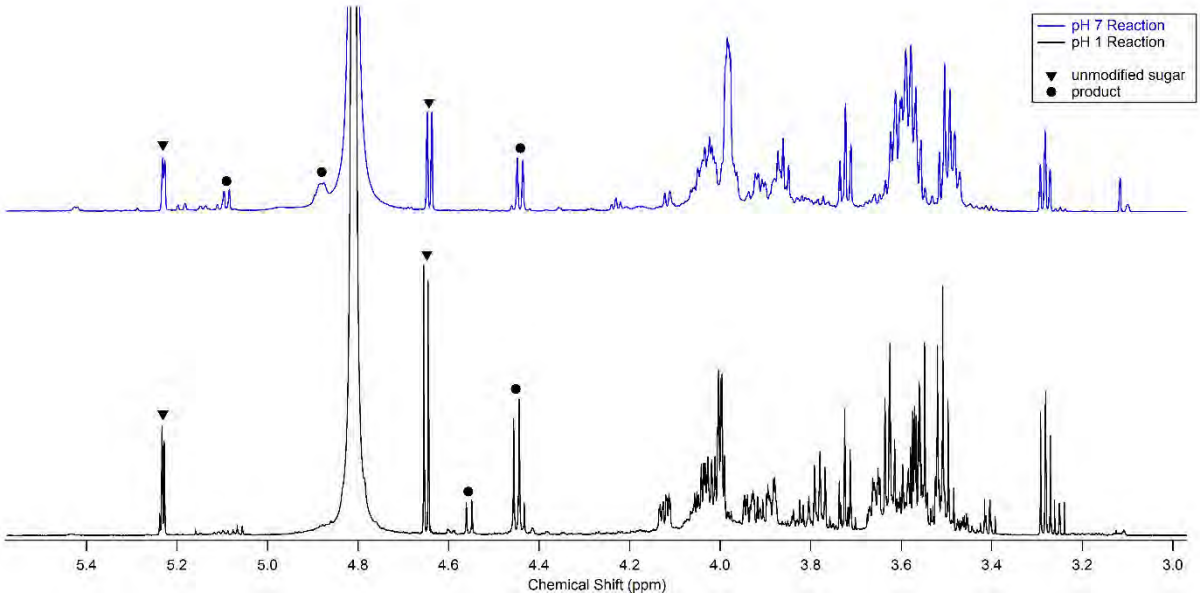
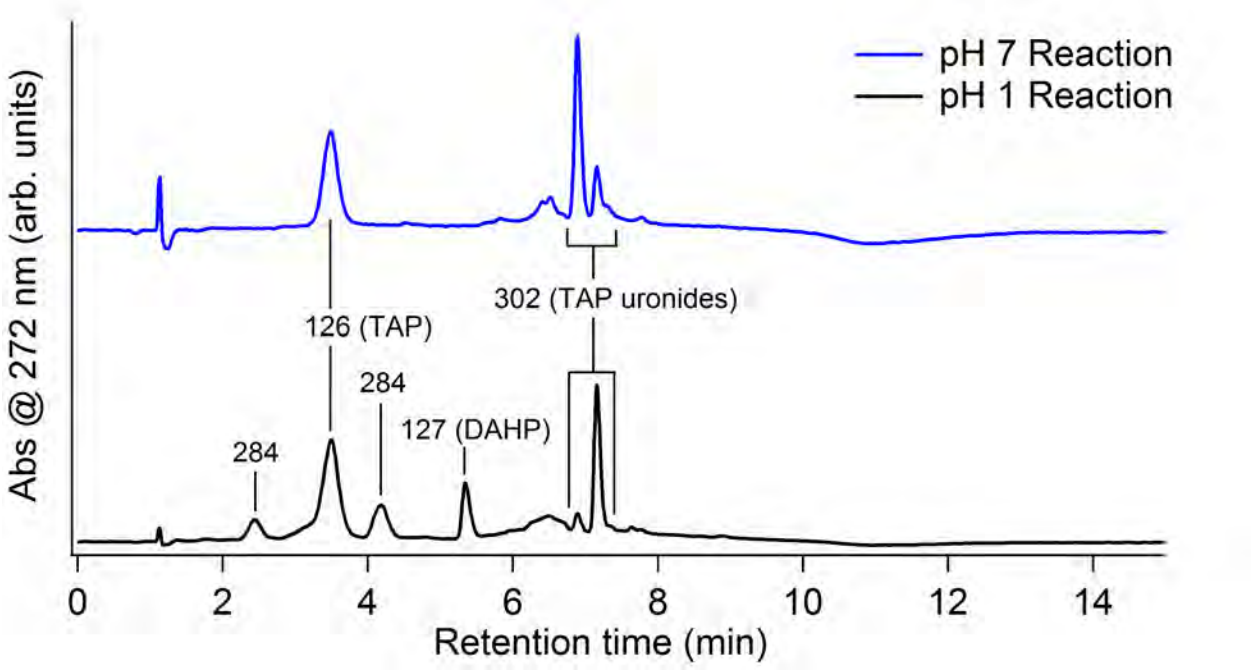


Fig. S8. Analysis of crude products from reactions of TAP with glucose-6-phosphate (G6P). **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at pH 1 or pH 7, sampled at 24 hours.

A.



B.

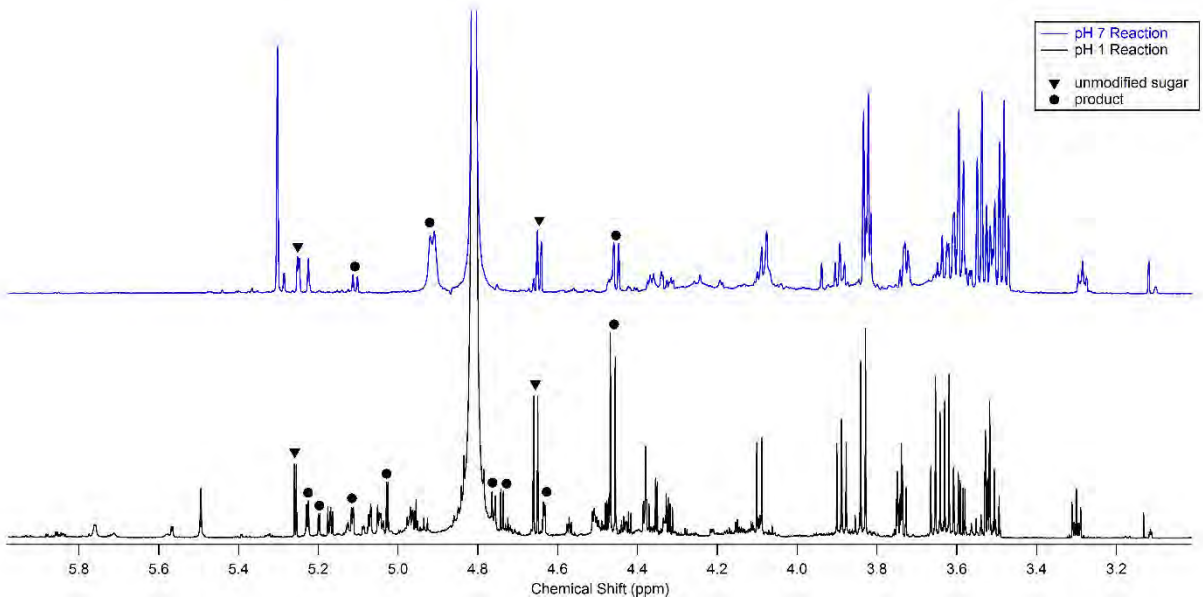
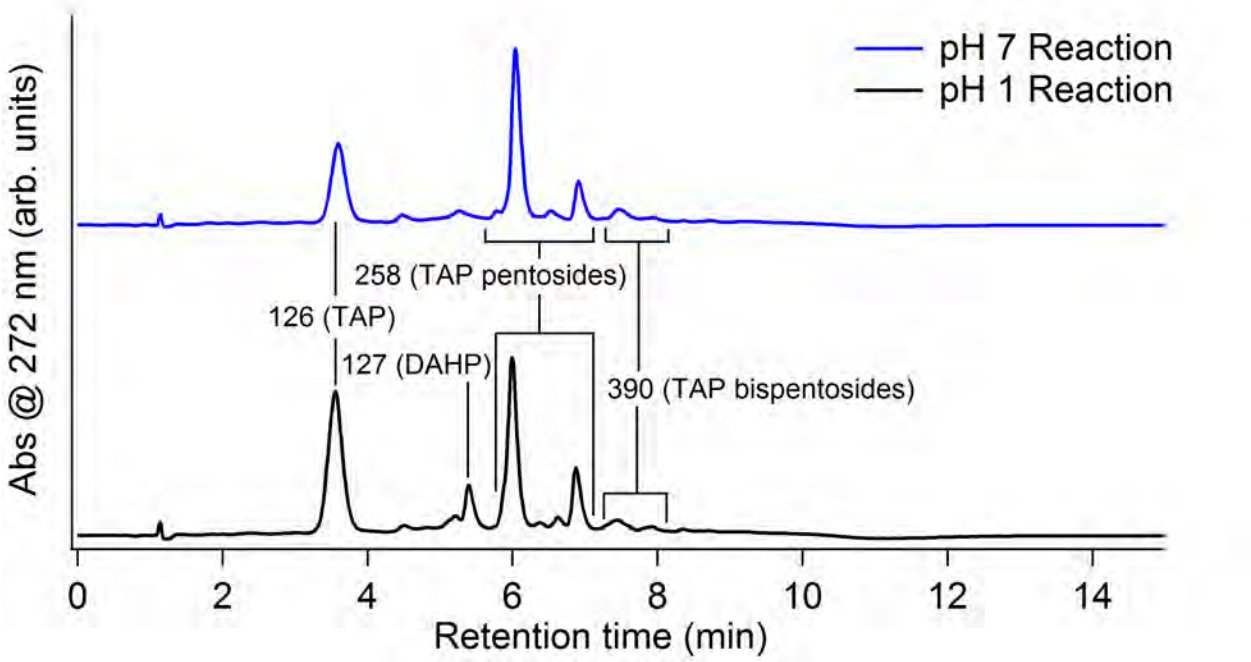


Fig. S9. Analysis of crude products from reactions of TAP with glucuronic acid. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at pH 1 or pH 7, sampled at 24 hours.

A.



B.

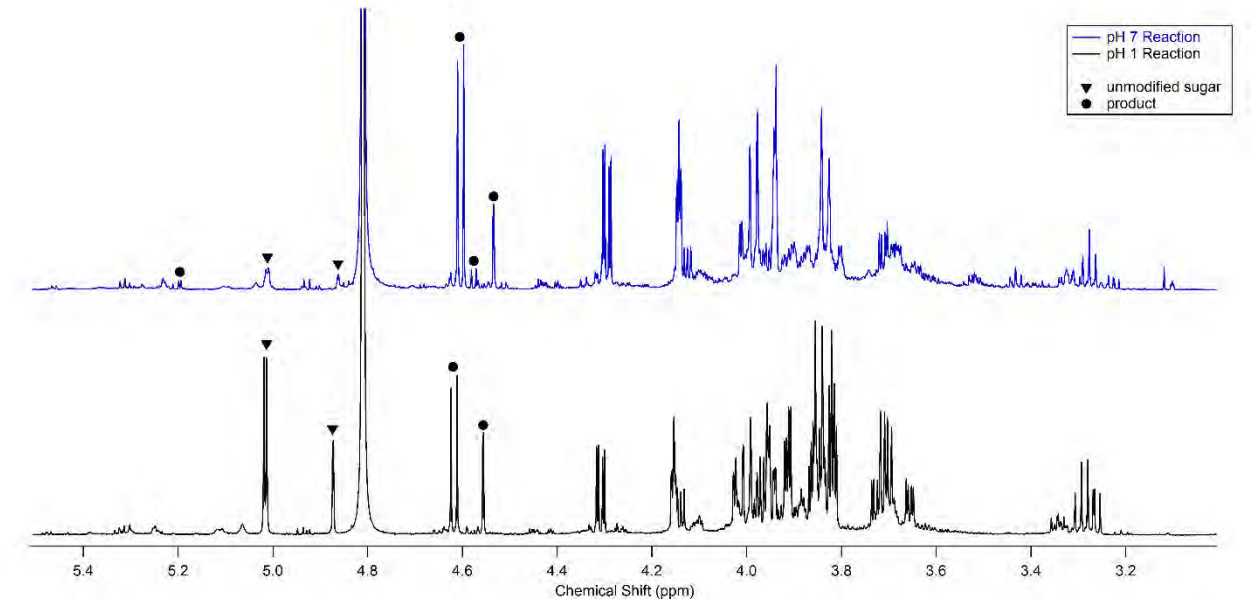
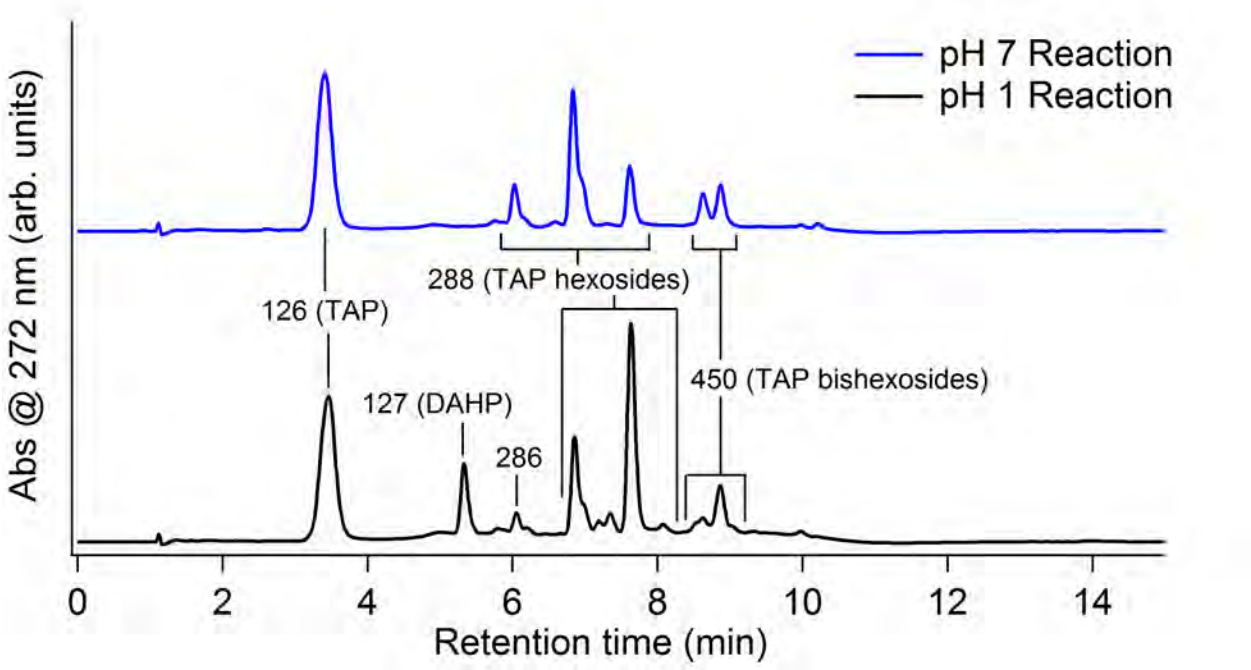


Fig. S10. Analysis of crude products from reactions of TAP with lyxose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.



B.

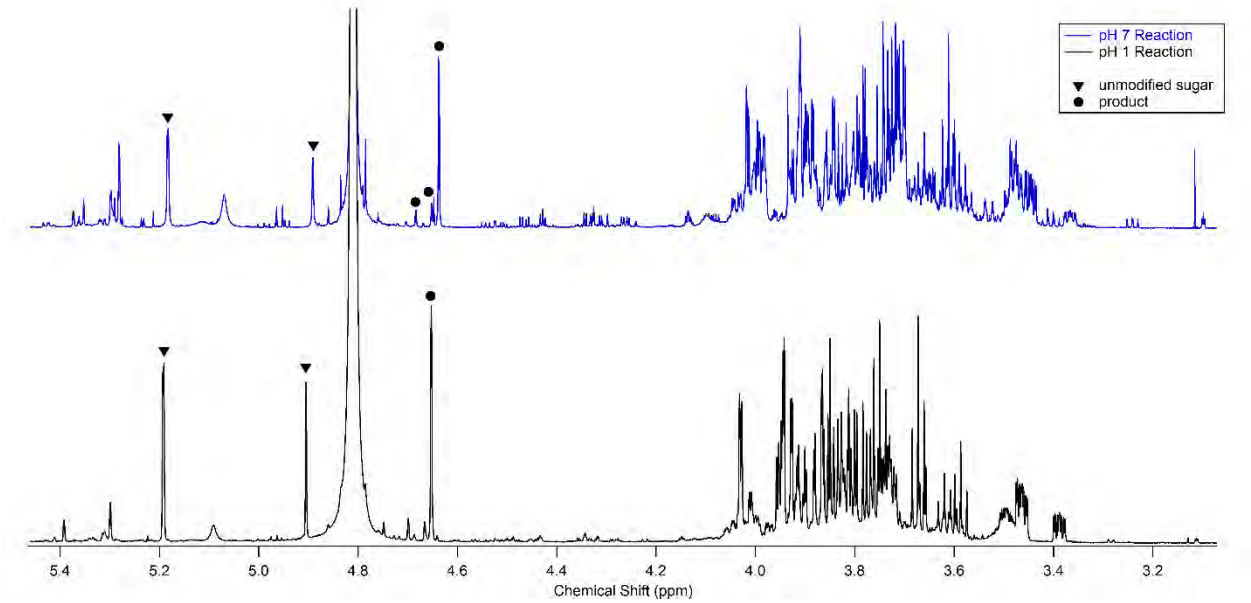
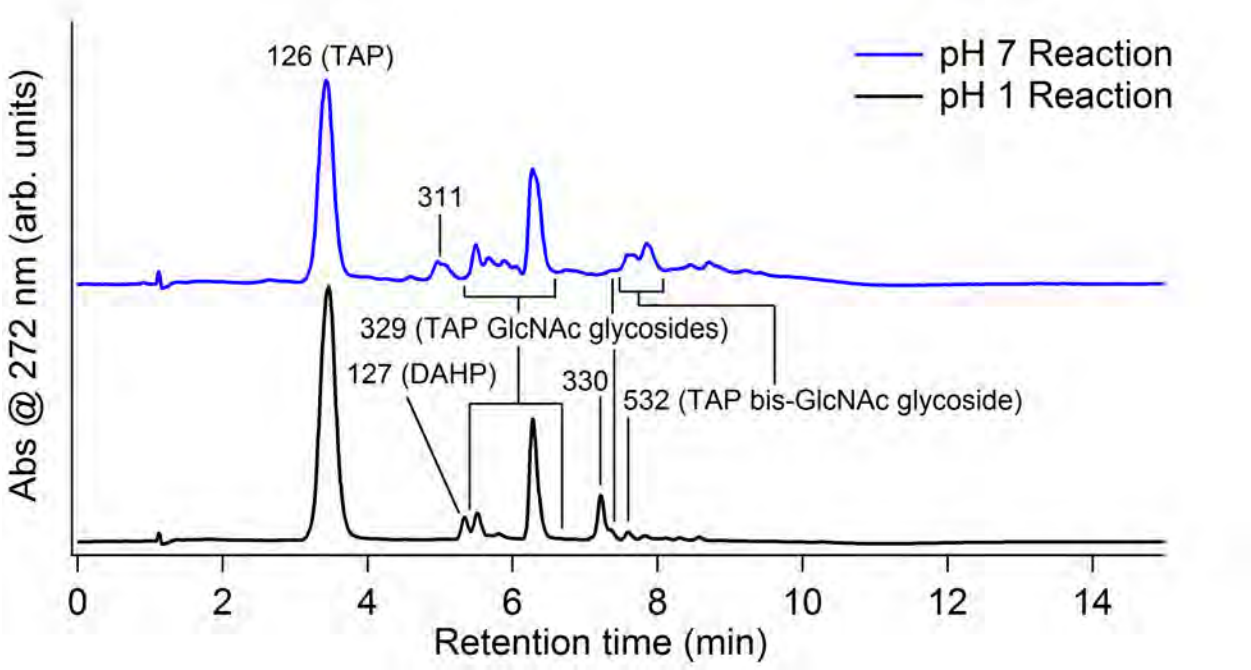


Fig. S11. Analysis of crude products from reactions of TAP with mannose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.



B.

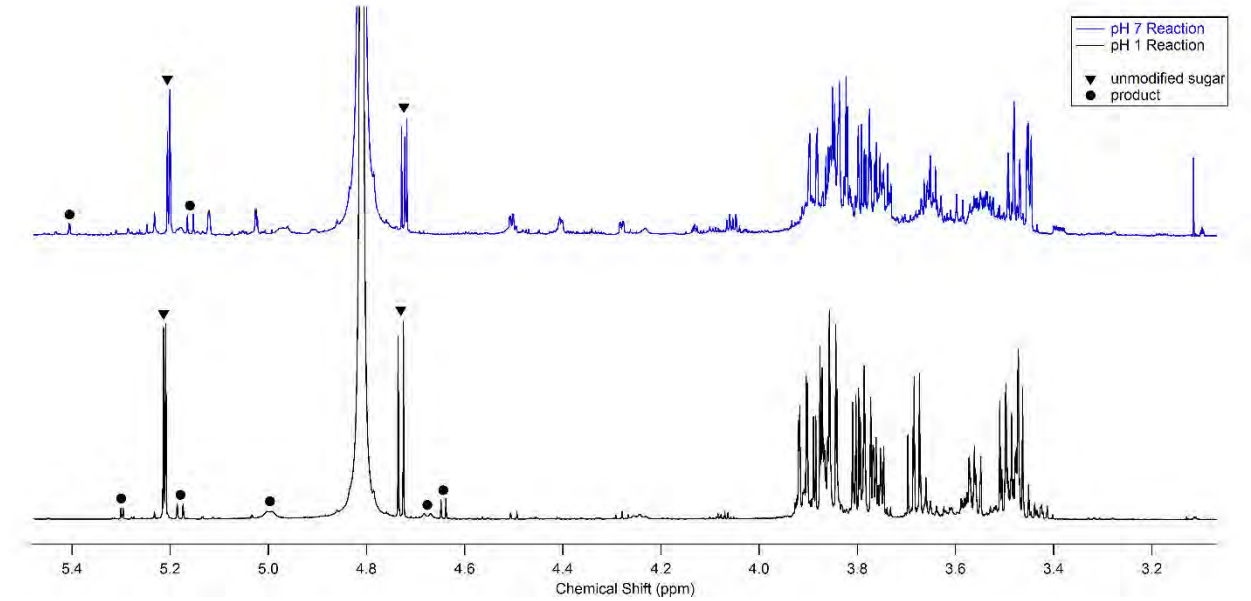
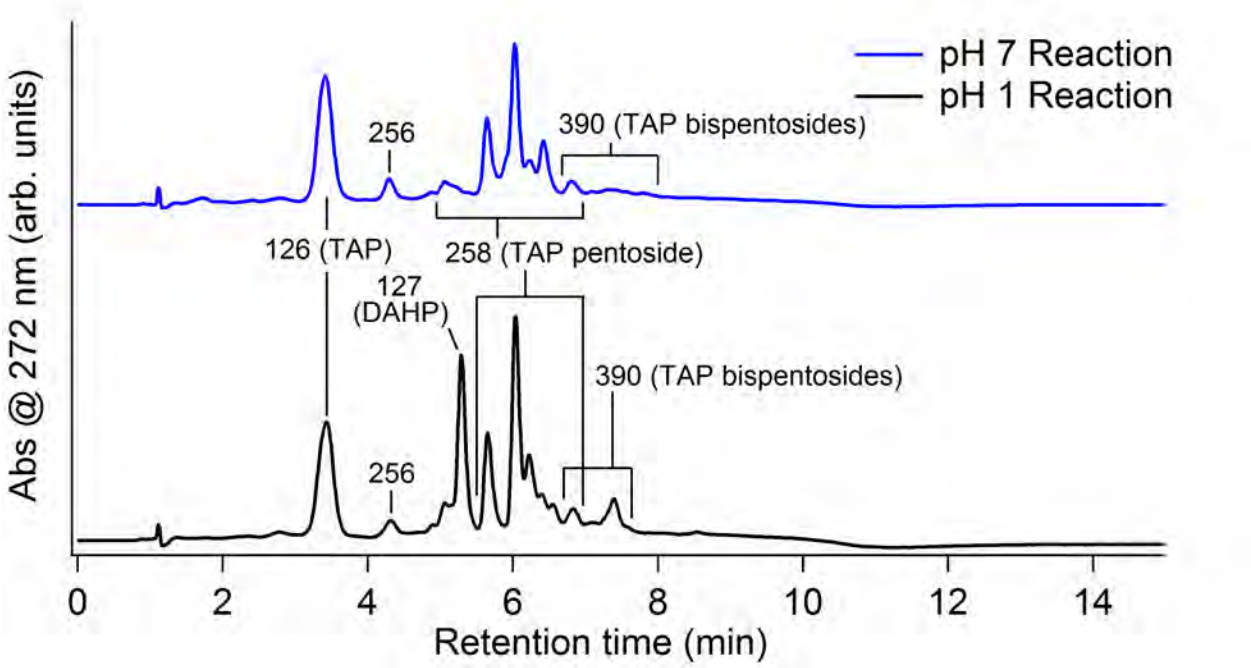


Fig. S12. Analysis of crude products from reactions of TAP with *N*-acetylglucosamine. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.



B.

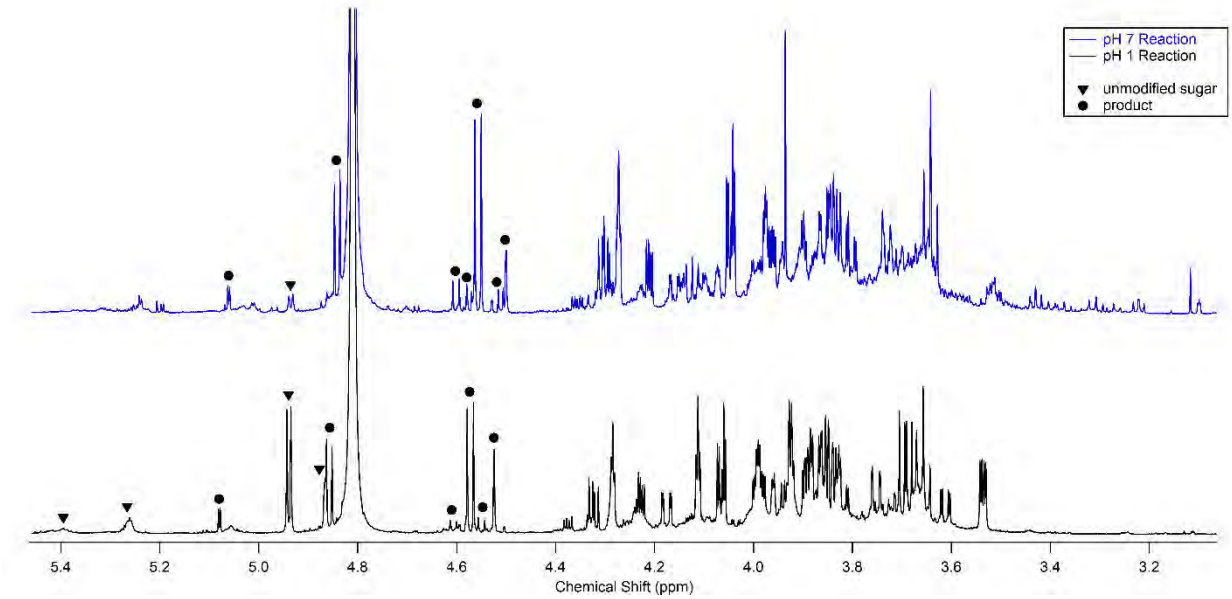
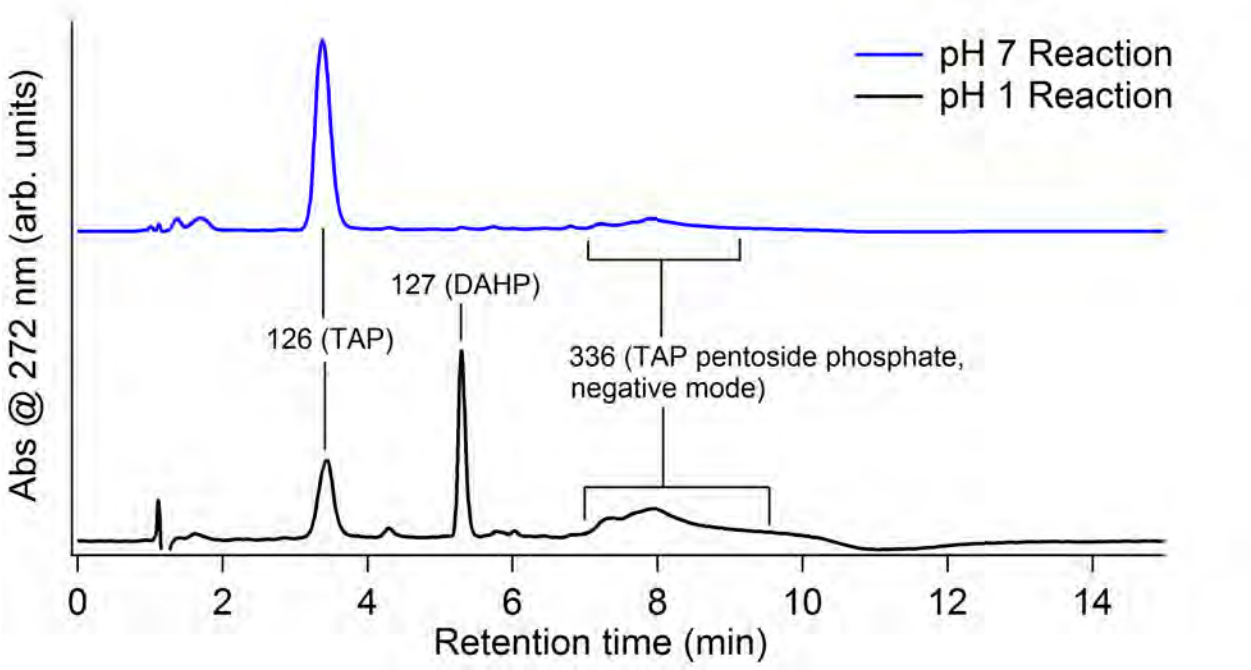
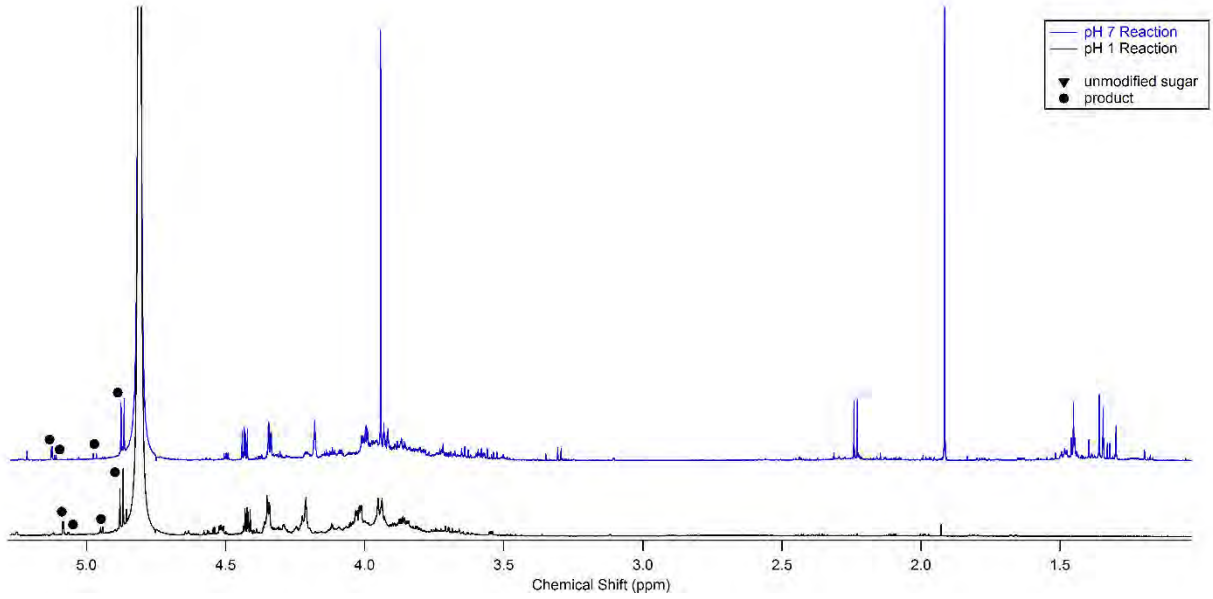


Fig. S13. Analysis of crude products from reactions of TAP with ribose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.


pH 7 Reaction
pH 1 Reaction
Abs @ 272 nm (arb. units)
126 (TAP)
127 (DAHP)
336 (TAP pentoside phosphate, negative mode)
0
2
4
6
8
10
12
14
Retention time (min)

B.


pH 7 Reaction
pH 1 Reaction
unmodified sugar
product
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
Chemical Shift (ppm)

C.

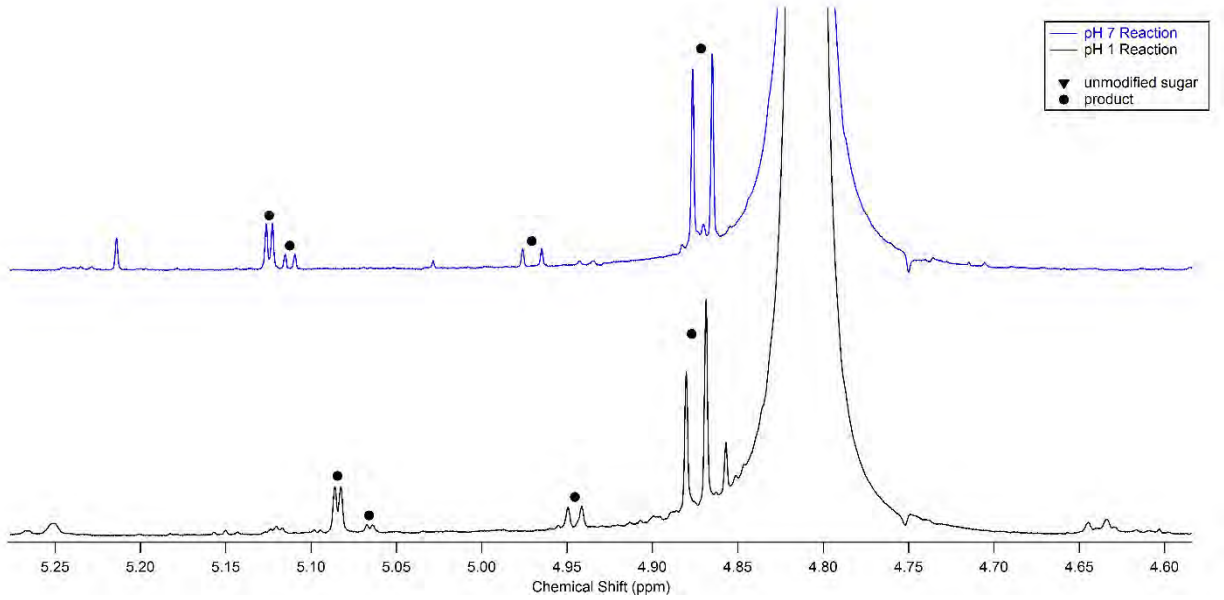
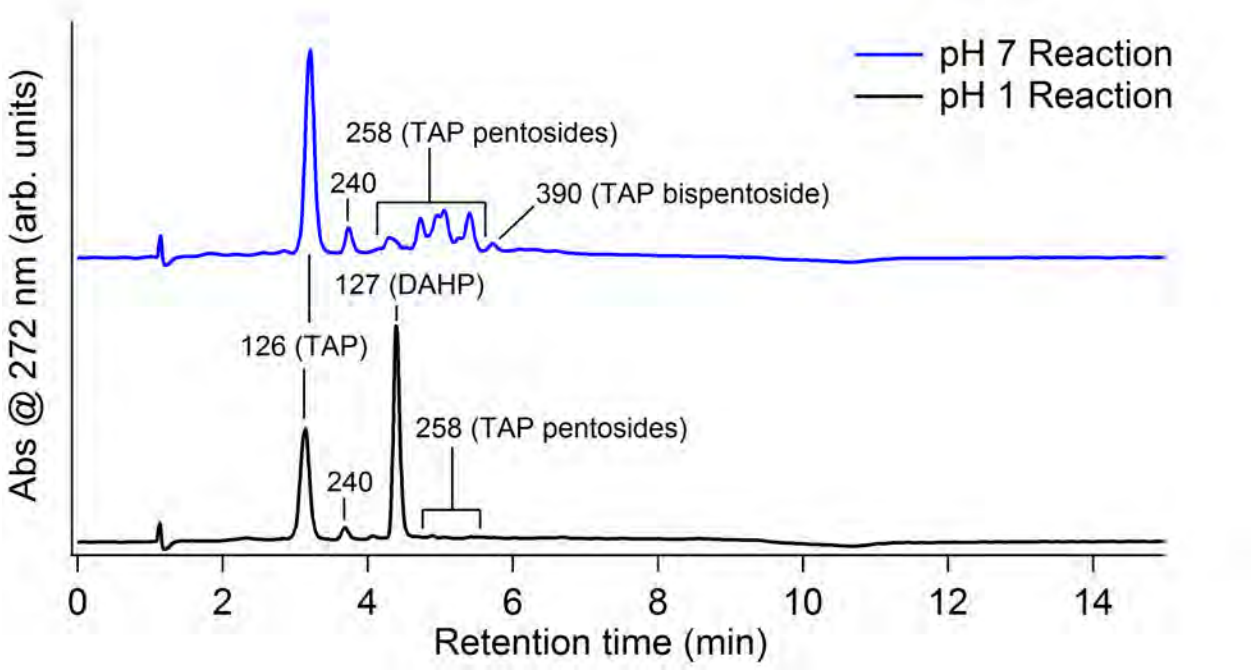


Fig. S14. Analysis of crude products from reactions of TAP with ribose-5-phosphate. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode, except for TAP nucleotides, which are displayed as deprotonated species as detected in negative mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. **(C)** Detail of the anomeric region of the ^{1}H NMR spectra.

A.



B.

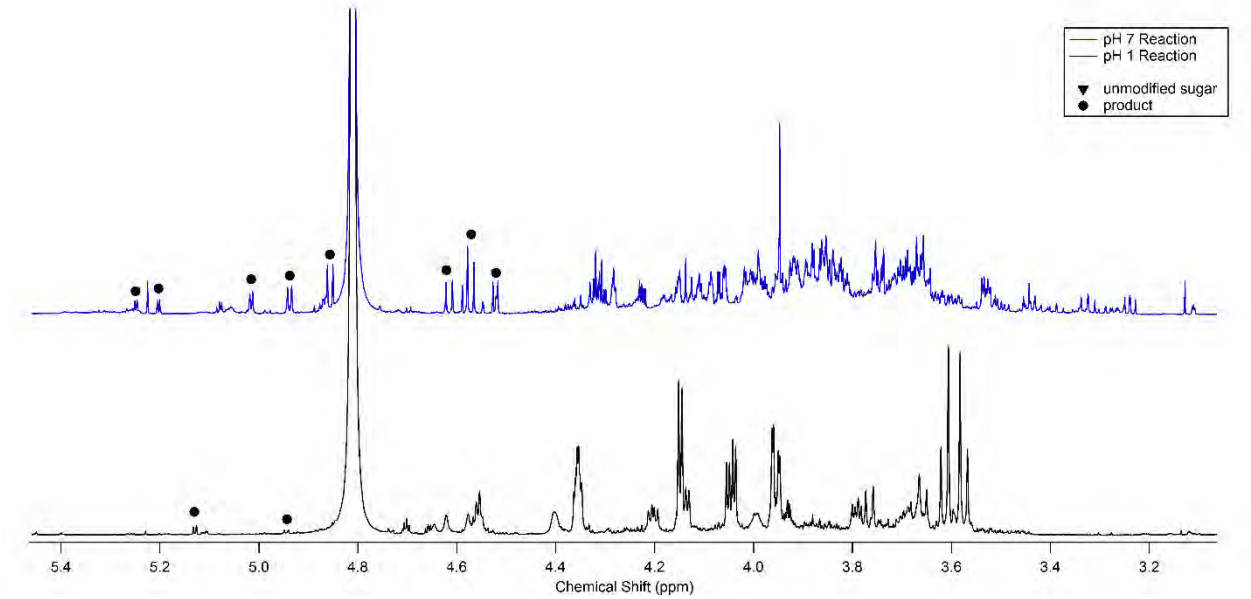
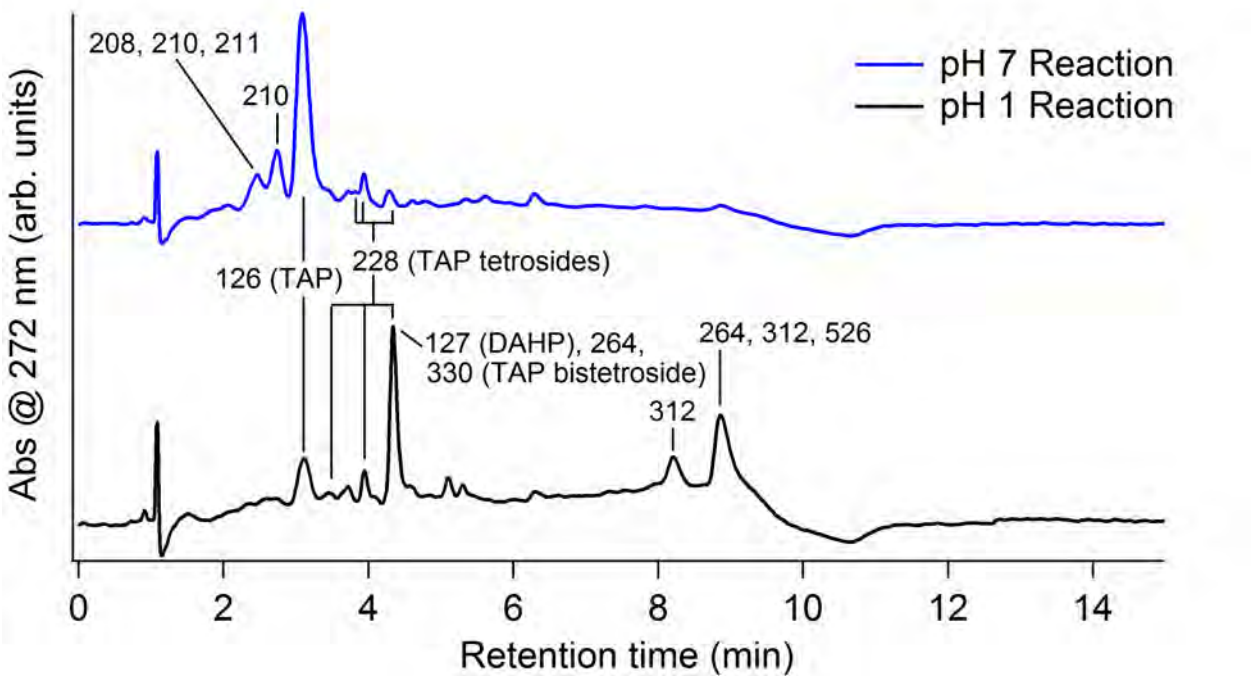
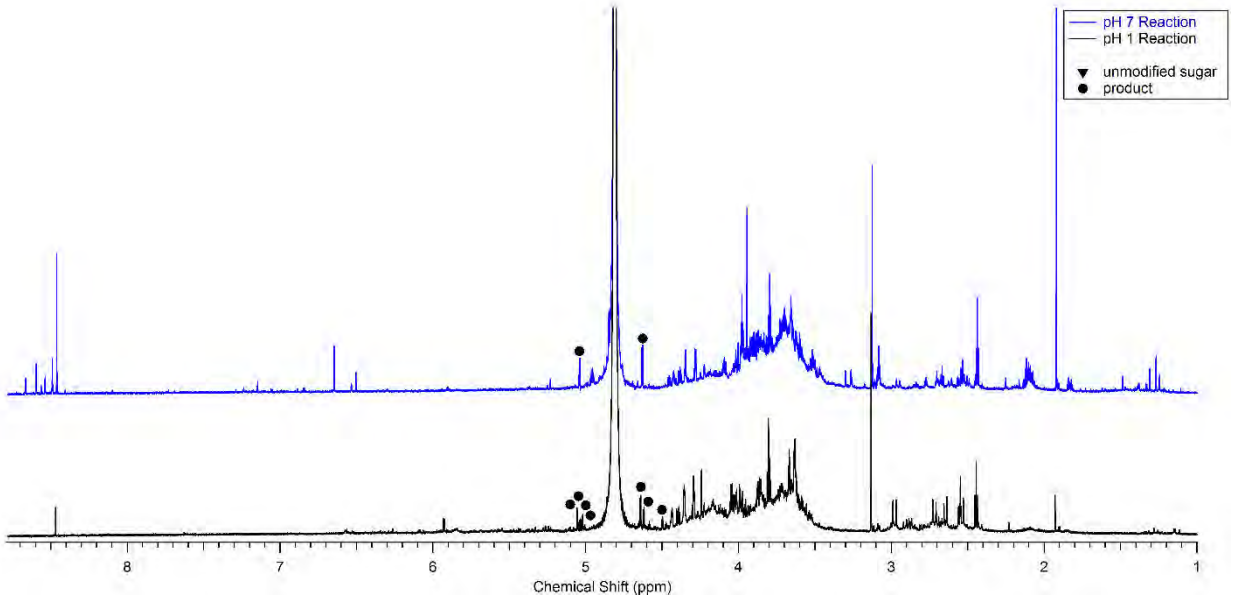


Fig. S15. Analysis of crude products from reactions of TAP with ribulose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Note: Resonances that appear to be associated with TAP-aldose (ribose or arabinose) glycosides in the pH 7 spectrum are located at 5.25, 5.20, 5.01, 4.94, 4.86, 4.61, 4.57 and 4.52 ppm, and in the pH 1 spectrum are located at 5.13 and 4.95 ppm.

A.


208, 210, 211
210
pH 7 Reaction
pH 1 Reaction
126 (TAP)
228 (TAP tetrosides)
127 (DAHP), 264,
330 (TAP bistetroside)
264, 312, 526
312
Abs @ 272 nm (arb. units)
0
2
4
6
8
10
12
14
Retention time (min)

B.


pH 7 Reaction
pH 1 Reaction
unmodified sugar
product
8
7
6
5
4
3
2
1
Chemical Shift (ppm)

C.

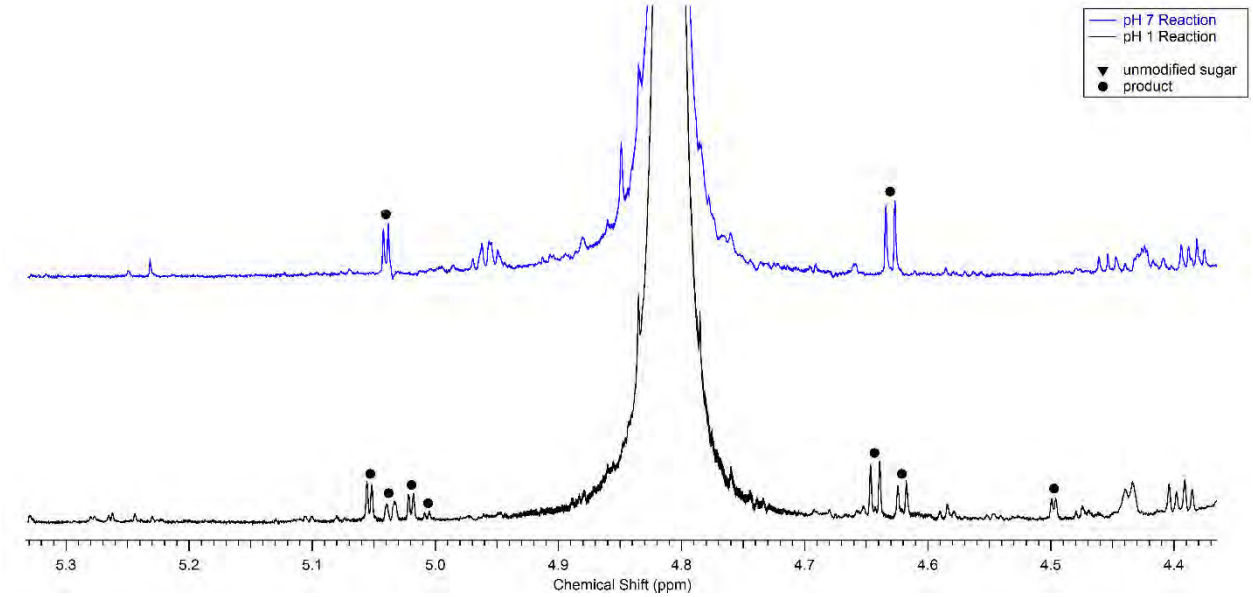
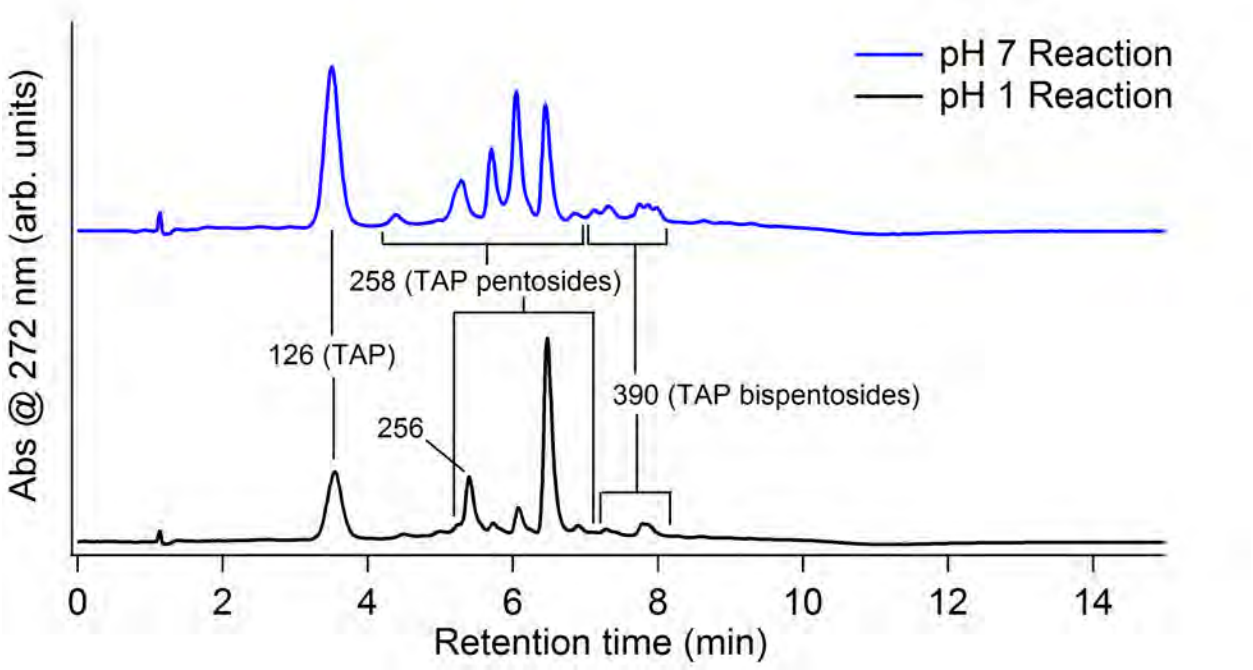


Fig. S16. Analysis of crude products from reactions of TAP with threose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. **(C)** Detail of the anomeric region of the ^{1}H NMR spectra.

A.



B.

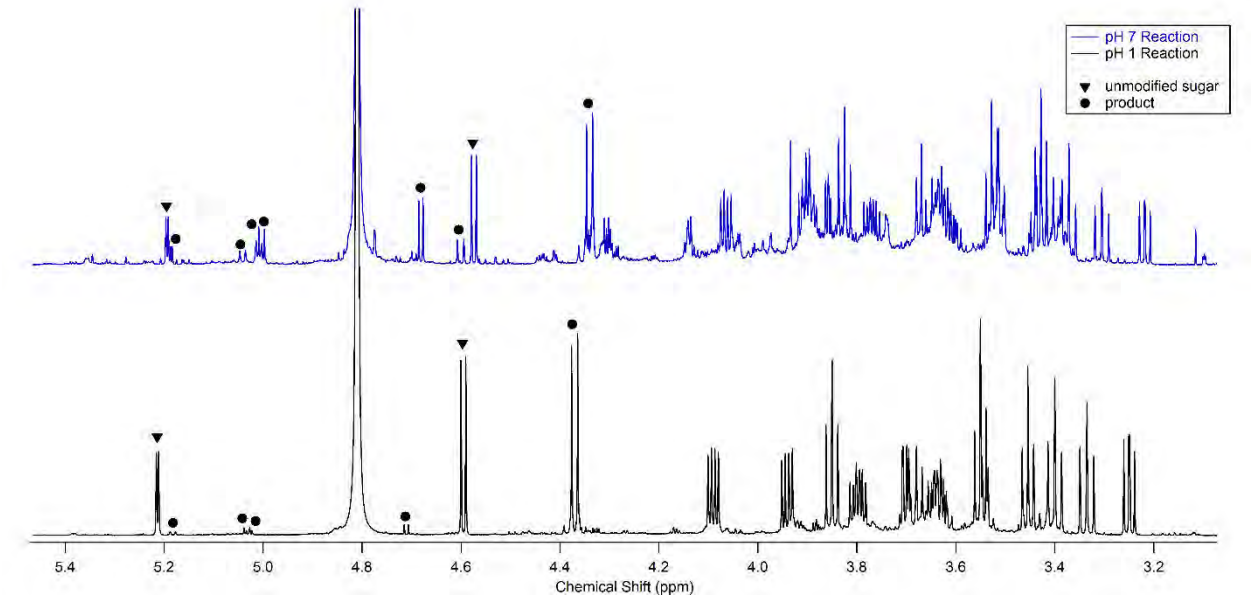
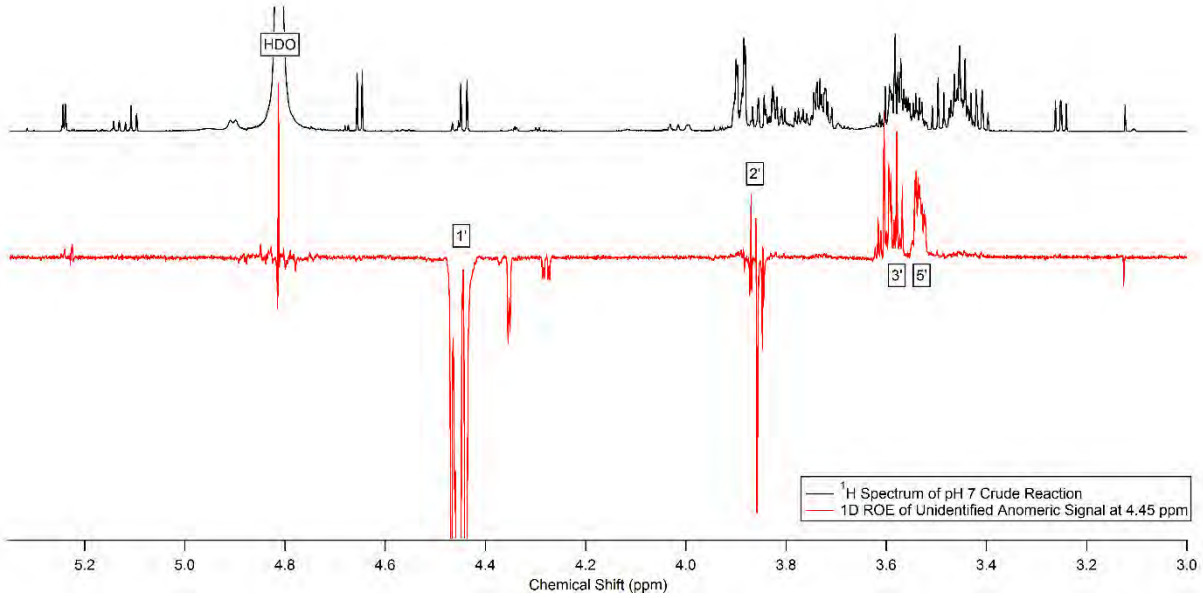


Fig. S17. Analysis of crude products from reactions of TAP with xylose. **(A)** LC-MS/UV chromatograms of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours. Masses shown are protonated species in positive mode. **(B)** ^{1}H NMR spectra of products resulting from the reactions at either pH 1 or pH 7, 85 °C, sampled at 24 hours.

A.



B.

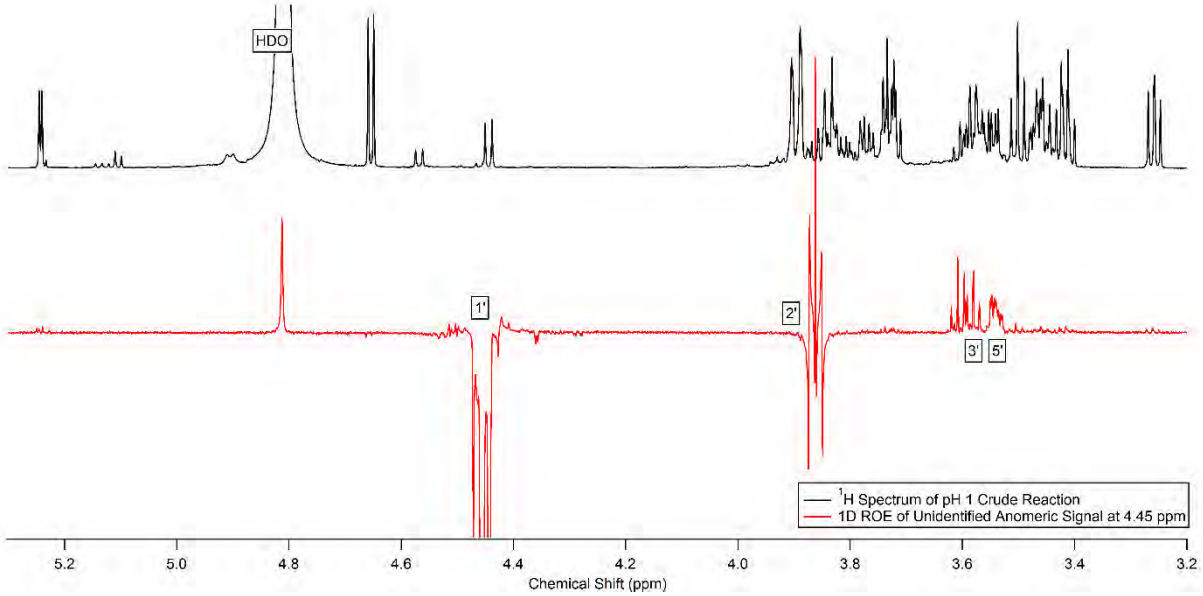
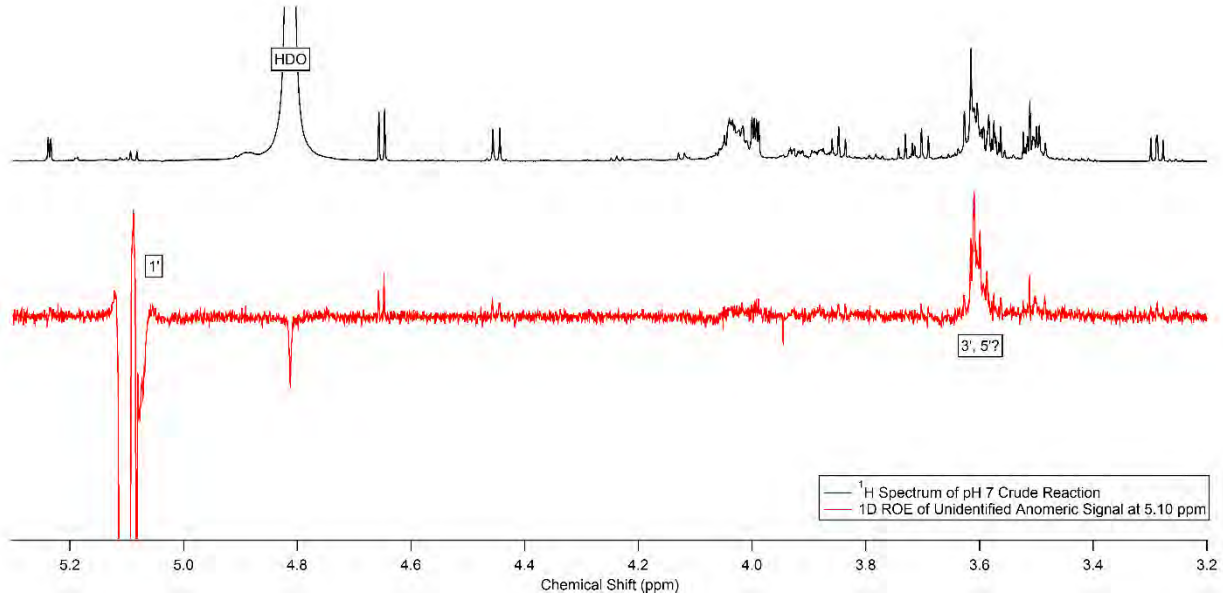
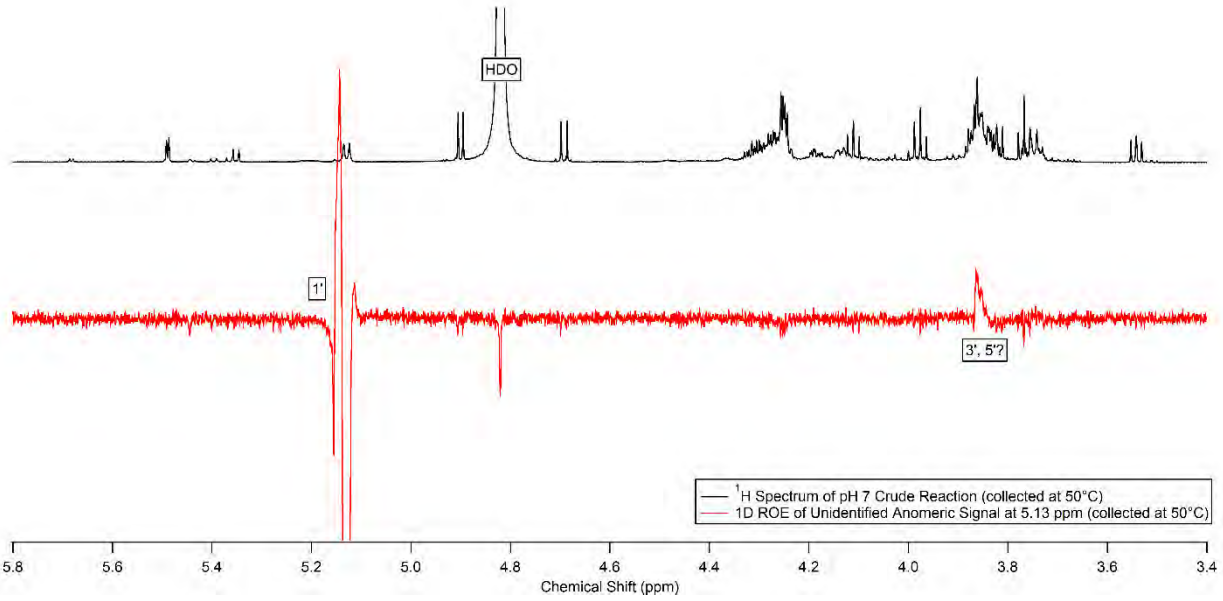


Fig. S18. 1D ROE NMR analysis of the crude reactions of TAP with glucose. **(A)** 1D ROE analysis of the reaction of TAP with glucose, pH 7. Saturation of the anomeric signal at 4.45 ppm gives a pronounced TOCSY effect at 3.85 ppm and ROEs at 3.5-3.6 ppm, consistent with a *β-C*-pyranoside, as seen in the 1D ROE analysis of TAP-Glc6P (a *β-C*-pyranoside, Fig. S22C). **(B)** 1D ROE analysis of the reaction of TAP with glucose, pH 1. Saturation of the anomeric signal at 4.45 ppm gives a similar result as found in the reaction of TAP with glucose at pH 7.

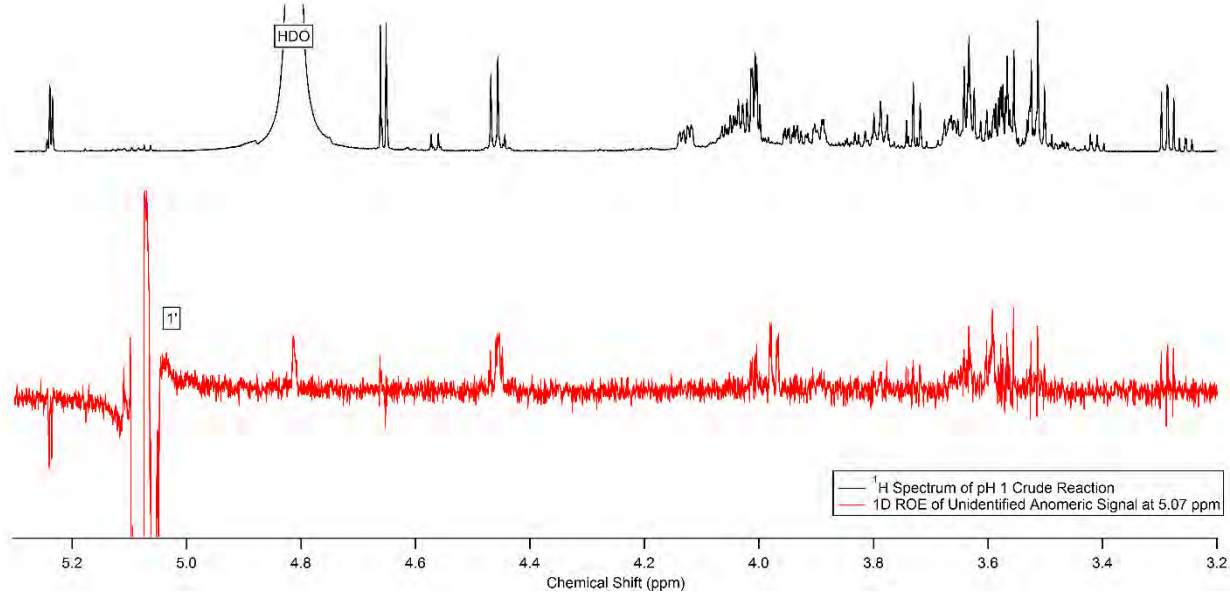
A.


HDO
1'
3', 5'?
^{1}H Spectrum of pH 7 Crude Reaction
1D ROE of Unidentified Anomeric Signal at 5.10 ppm
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
Chemical Shift (ppm)

B.


HDO
1'
3', 5'?
^{1}H Spectrum of pH 7 Crude Reaction (collected at 50°C)
1D ROE of Unidentified Anomeric Signal at 5.13 ppm (collected at 50°C)
5.8
5.6
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
Chemical Shift (ppm)

C.



D.

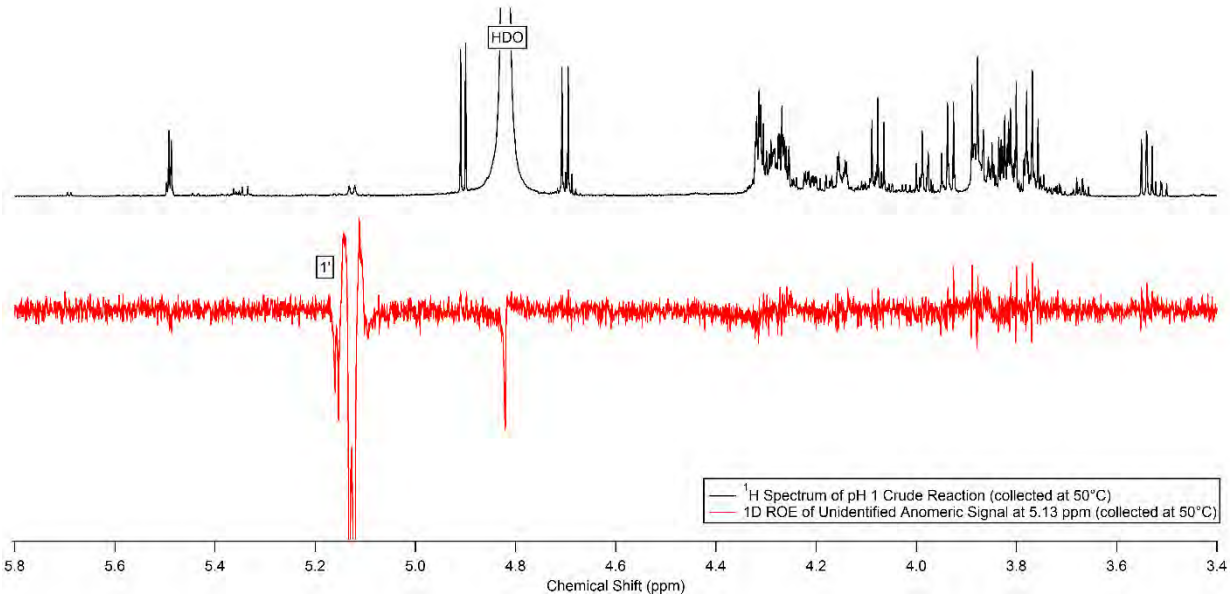


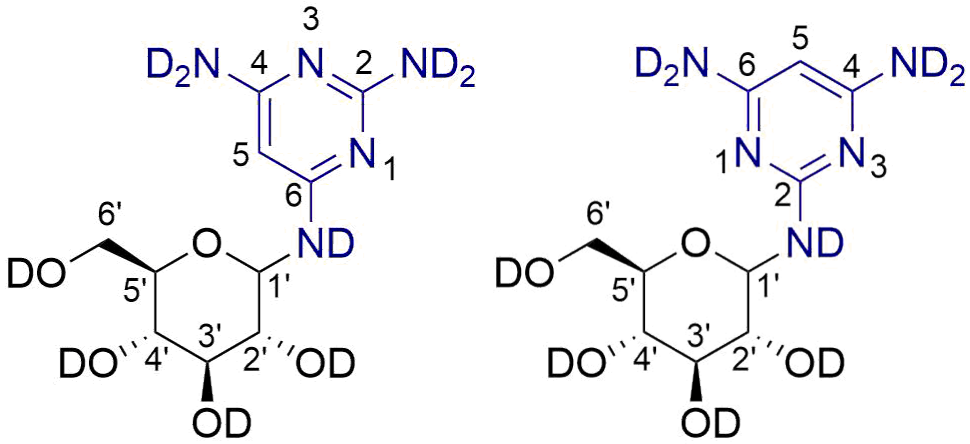
Fig. S19. 1D ROE NMR analysis of the crude reactions of TAP with glucose-6-phosphate. **(A)** 1D ROE analysis of the reaction of TAP with Glc6P, pH 7, performed at 25°C. Saturation of the anomeric signal at 5.10 ppm gives ROEs at 3.8-3.9 ppm, consistent with a *β*-*N*-pyranoside, as seen in the 1D ROE analysis of TAP-Glc (a mixture of *β*-*N*-pyranosides, Fig. S20C). **(B)** 1D ROE analysis of the reaction of TAP with Glc6P, pH 1, performed at 50°C in order to sharpen the broad anomeric signal ordinary found at 4.9 ppm at 25°C. Saturation of this anomeric signal, now at 5.13 ppm, gives a similar result as found in (A), but absolute structural certainty cannot be ascertained. **(C)** 1D ROE analysis of the reaction of TAP with Glc6P, pH 1, performed at 25°C. Due to the low abundance of this species, saturation of the anomeric signal at 5.10 does not

produce interpretable ROEs. However, the chemical shift of this anomeric signal is consistent with a *β*-*N*-pyranoside, as seen for glucose in Fig. S20A. **(D)** 1D ROE analysis of the reaction of TAP with Glc6P, pH 1, performed at 50°C in order to sharpen the broad anomeric signal ordinary found at 4.9 ppm at 25°C. Saturation of this anomeric signal, now at 5.13 ppm, gives a similar result as found in (C). Although no discernable ROEs are present, the chemical shift and broadening of this signal upon heating are consistent with a *β*-*N*-pyranoside, as seen for glucose in Fig. S20A.

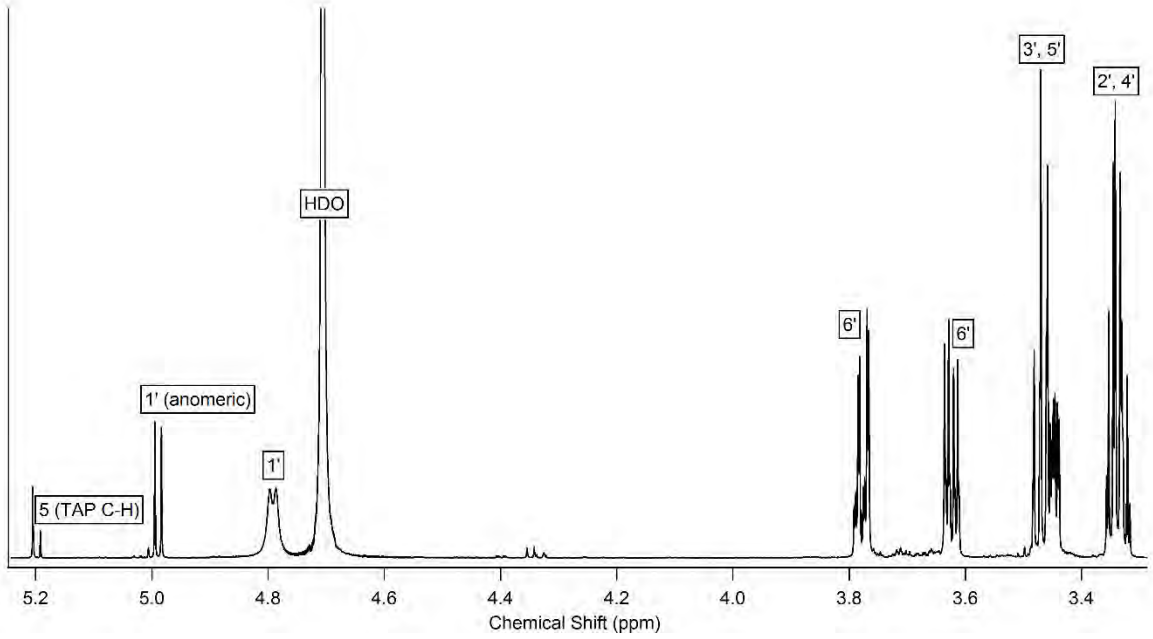
II. Analysis of Purified TAP Glycosides

Figure S20. NMR analysis of purified TAP glucosides (TAP-glucose)

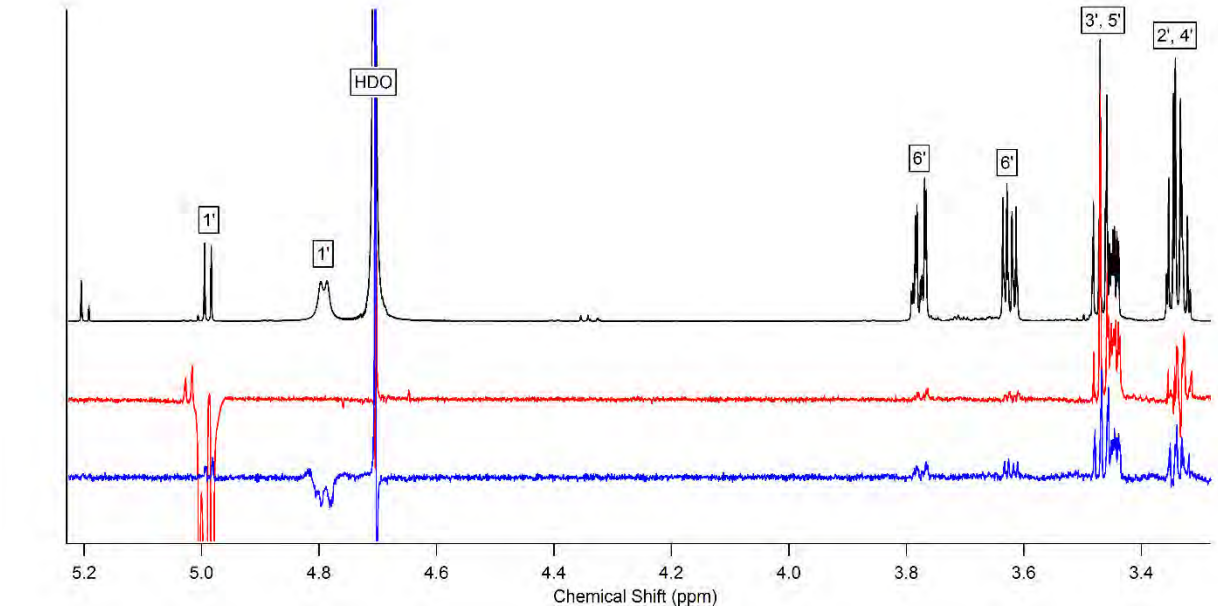
A.



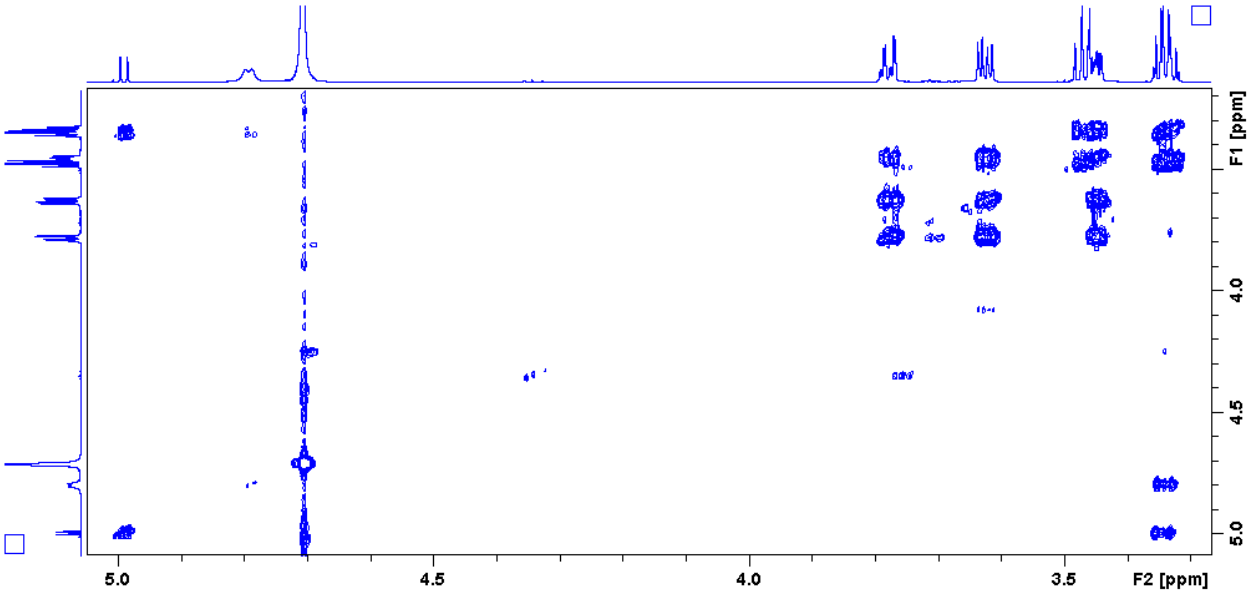
B.



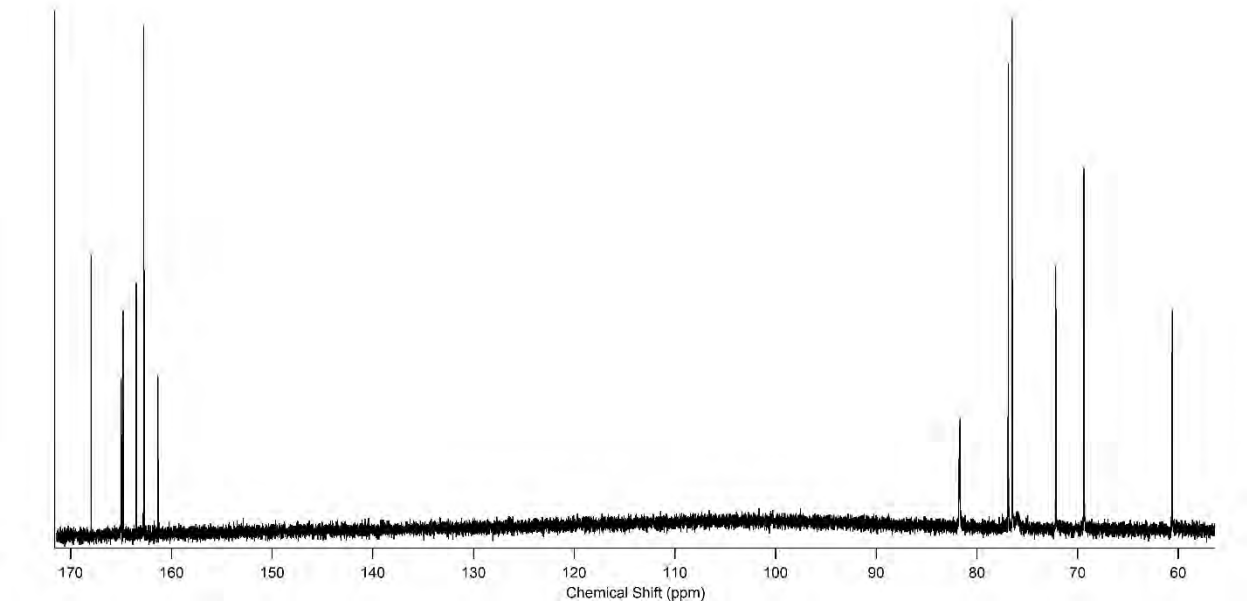
C.


HDO
1'
1'
6'
6'
3', 5'
2', 4'
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
Chemical Shift (ppm)

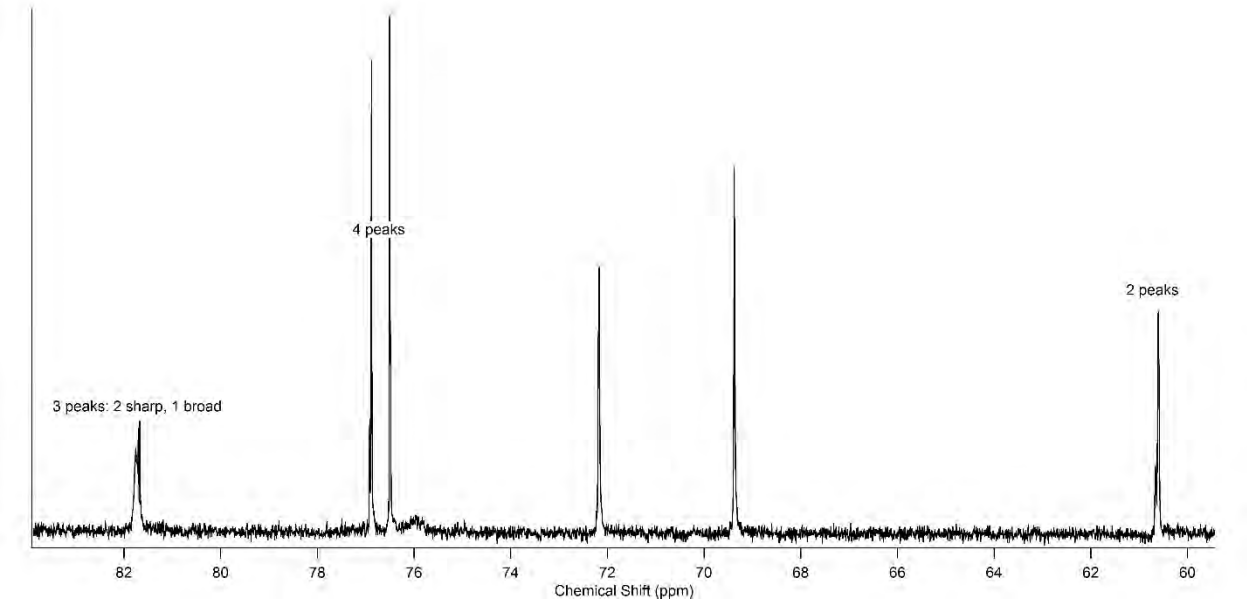
D.


F1 [ppm]
4.0
4.5
5.0
5.0
4.5
4.0
3.5
F2 [ppm]

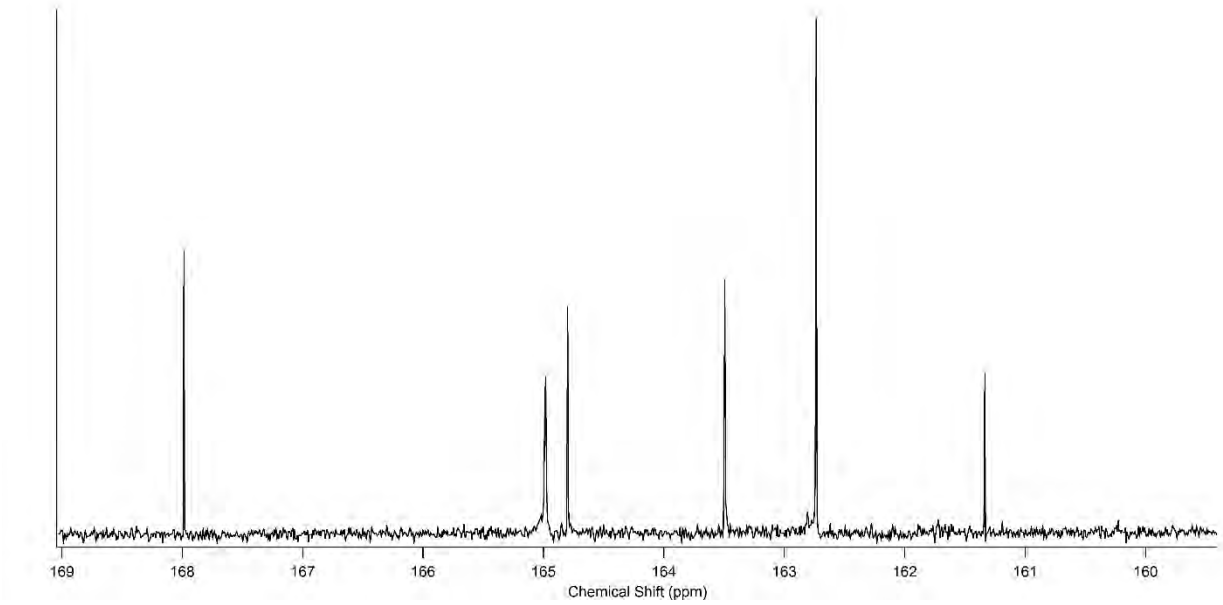
E. i.


170
160
150
140
130
120
110
100
90
80
70
60
Chemical Shift (ppm)

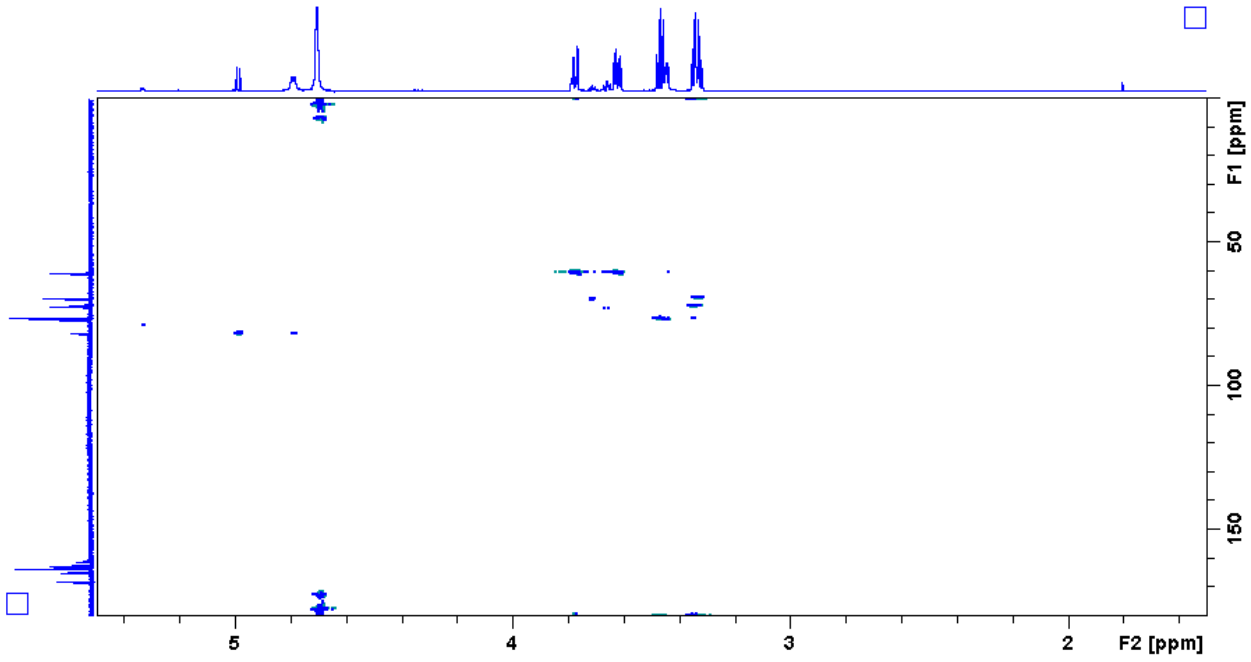
ii.


4 peaks
2 peaks
3 peaks: 2 sharp, 1 broad
82
80
78
76
74
72
70
68
66
64
62
60
Chemical Shift (ppm)

iii.


169
168
167
166
165
164
163
162
161
160
Chemical Shift (ppm)

F.


F1 [ppm]
50
100
150
5
4
3
2
F2 [ppm]

G.

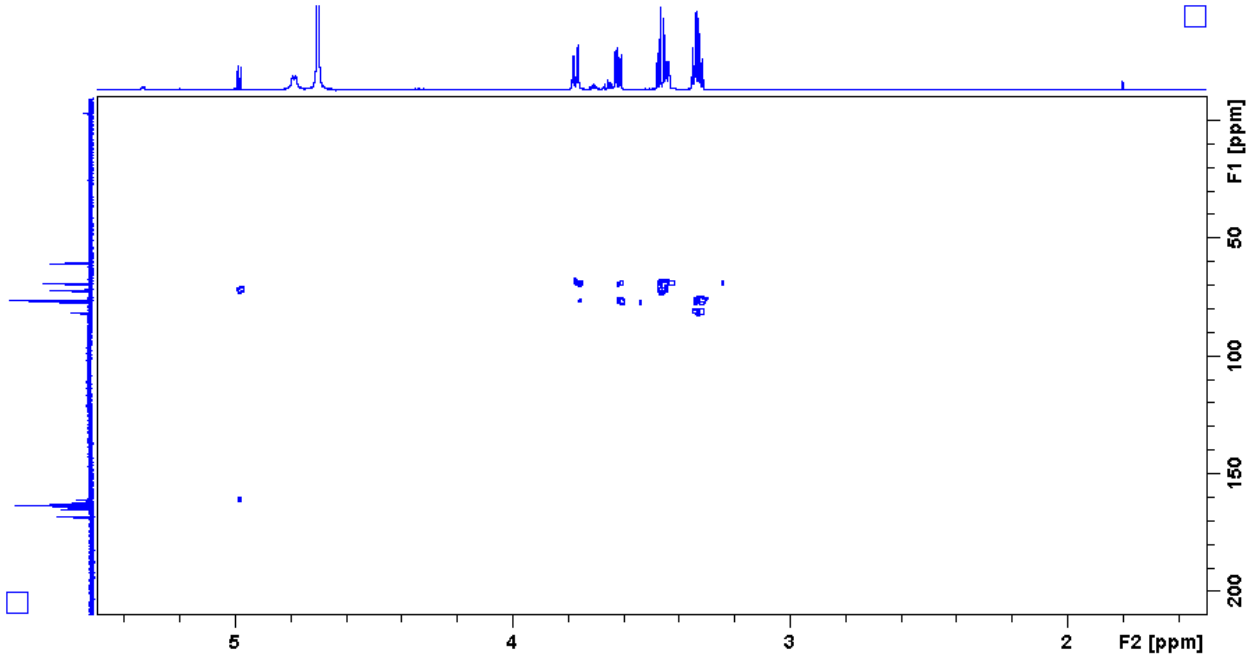
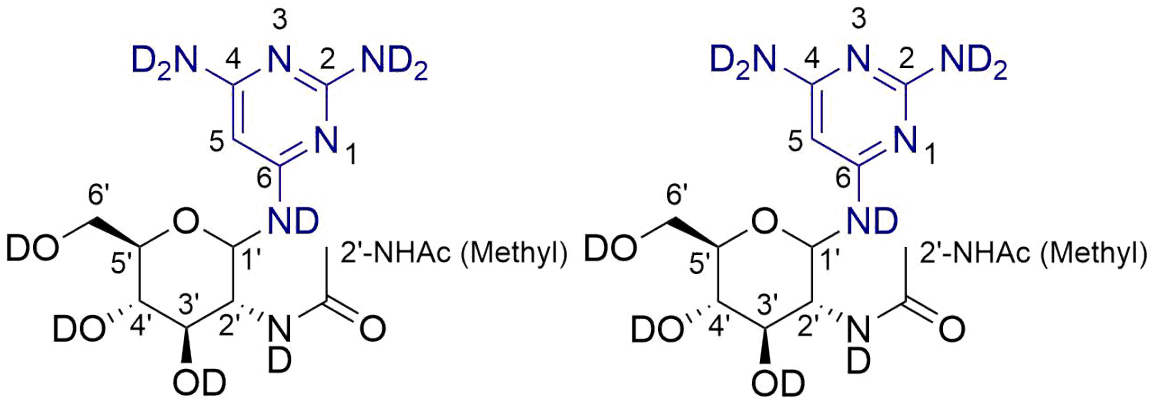


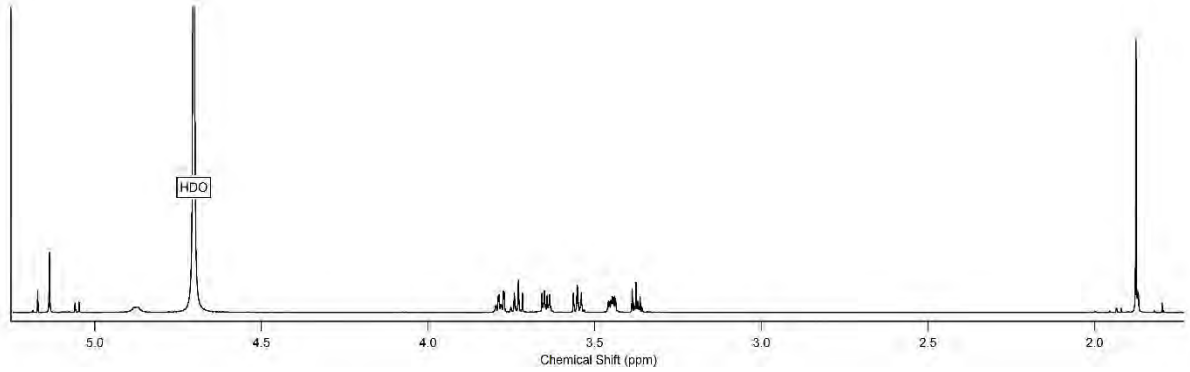
Fig. S20. NMR analysis of purified TAP glucosides (TAP-glucose). **(A)** Proposed structures of the isolated TAP-glucose species. **(B)** ^{1}H NMR spectrum. All peak assignments made through COSY (see below). **(C)** 1D ROE analysis of anomeric signals of TAP-glucose species. Saturation of either anomeric signal (protons designated as 1') gives a ROE to protons 3' and 5', with TOCSY to proton 2'. **(D)** COSY of TAP-glucose species. **(E)** ^{13}C NMR spectrum of TAP-glucose species. **(i)** Full ^{13}C NMR spectrum. **(ii)** Detail of upfield region of ^{13}C spectrum. **(iii)** Detail of downfield region of ^{13}C spectrum. **(F)** HSQC spectrum of TAP-glucose species. **(G)** HMBC spectrum of TAP-glucose species.

Figure S21. NMR analysis of purified TAP *N*-acetylglucosamine glycosides (TAP-GlcNAc)

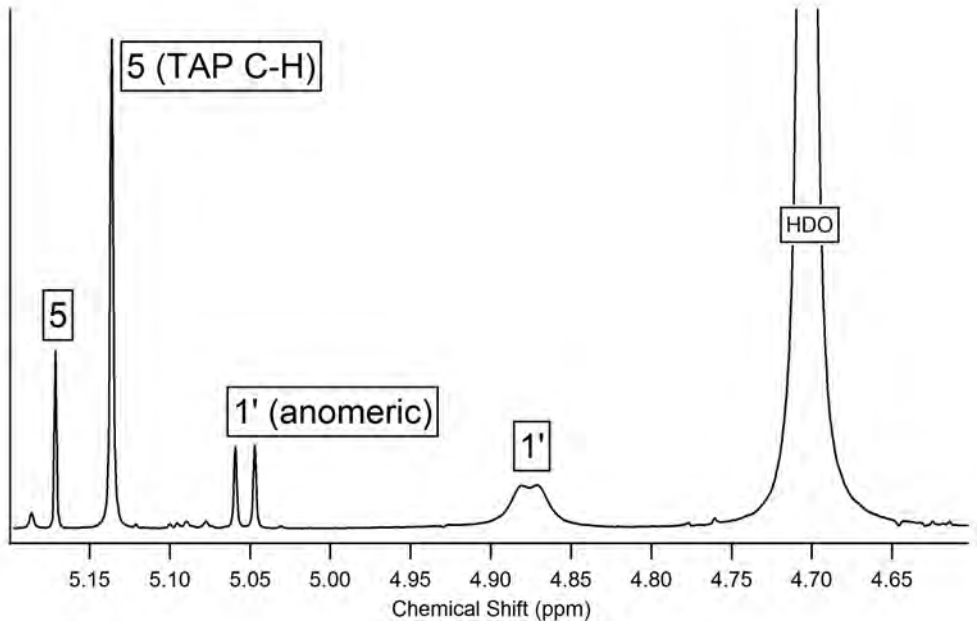
A.



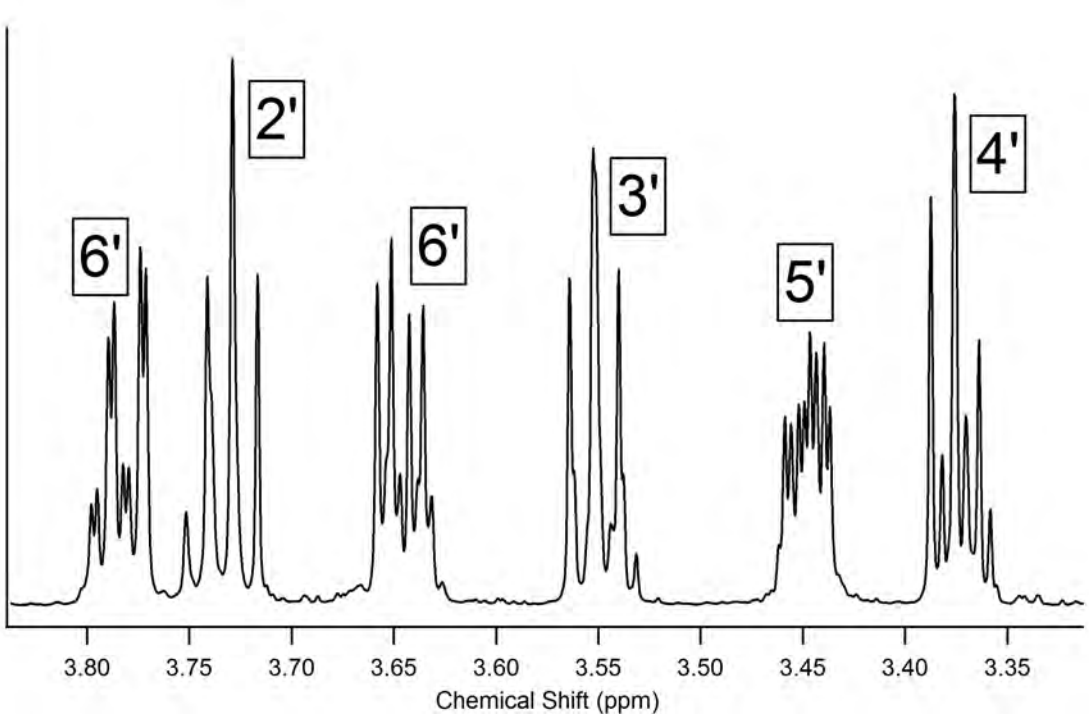
B. i.



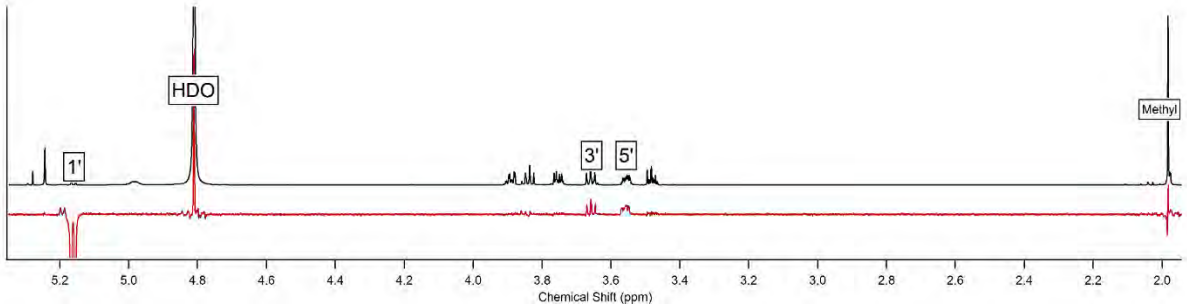
ii.



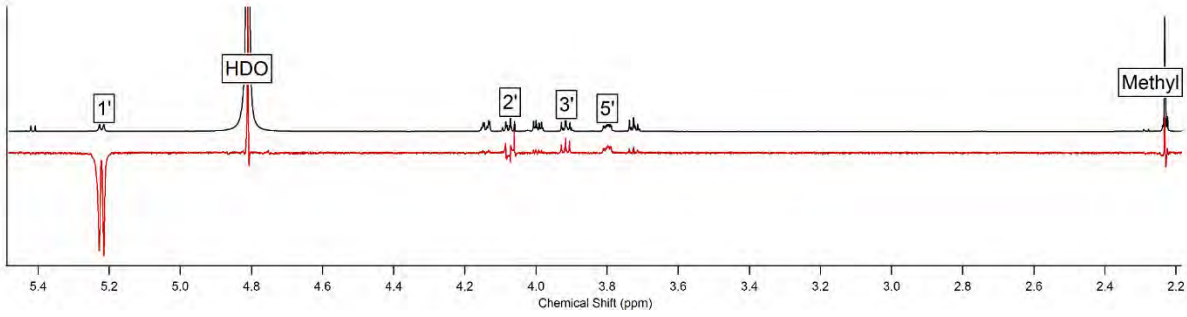
iii.


6'
2'
6'
3'
5'
4'
3.80
3.75
3.70
3.65
3.60
3.55
3.50
3.45
3.40
3.35
Chemical Shift (ppm)

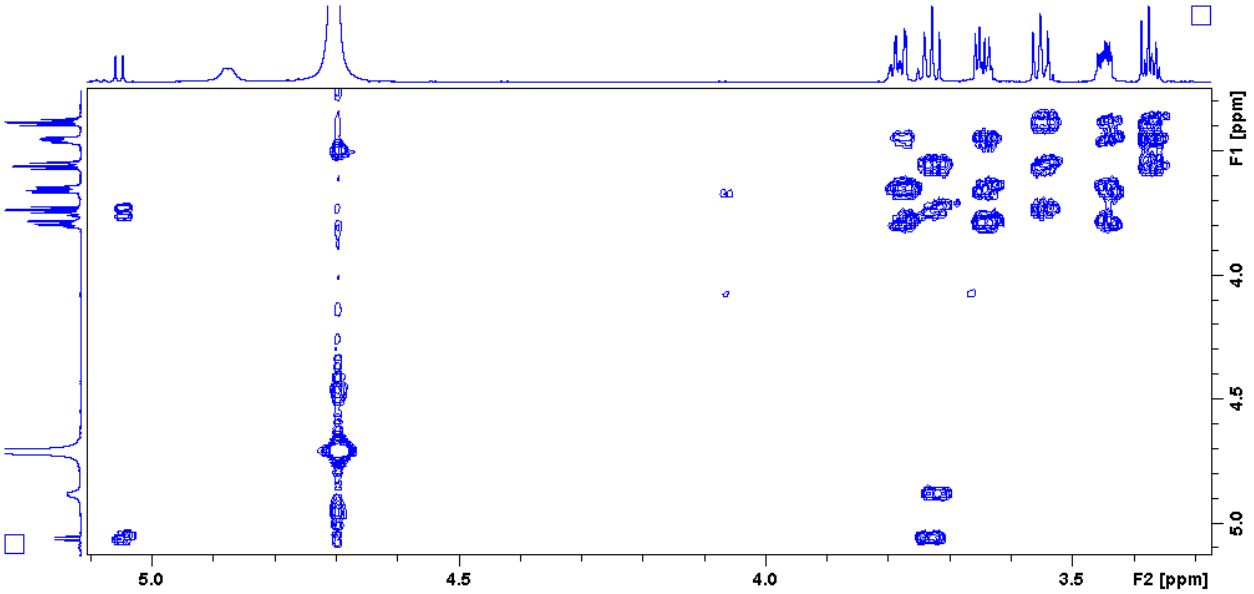
C. i.


HDO
1'
3'
5'
Methyl
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
Chemical Shift (ppm)

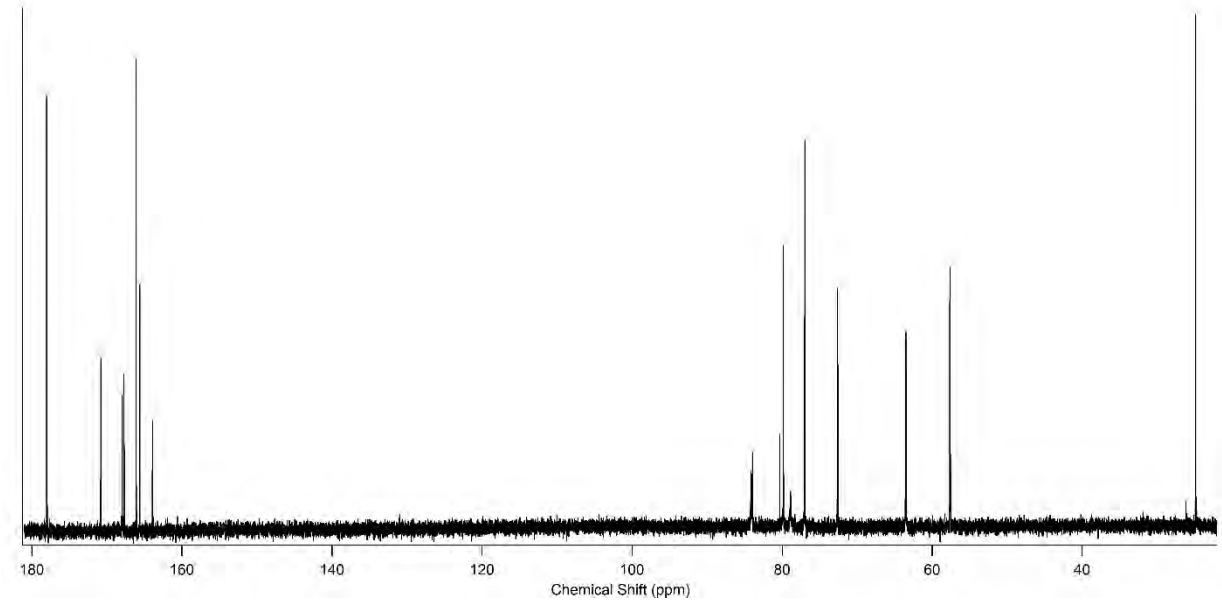
ii.


HDO
1'
2'
3'
5'
Methyl
5.4
5.2
5.0
4.8
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
Chemical Shift (ppm)

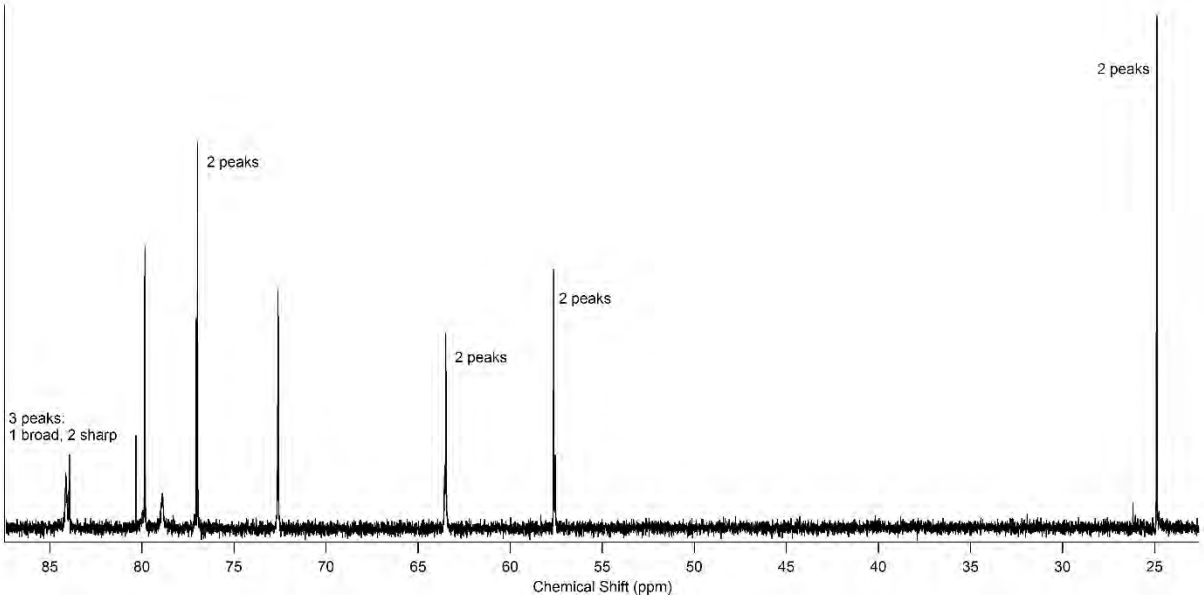
D.


F1 [ppm]
4.0
4.5
5.0
5.0
4.5
4.0
3.5
F2 [ppm]

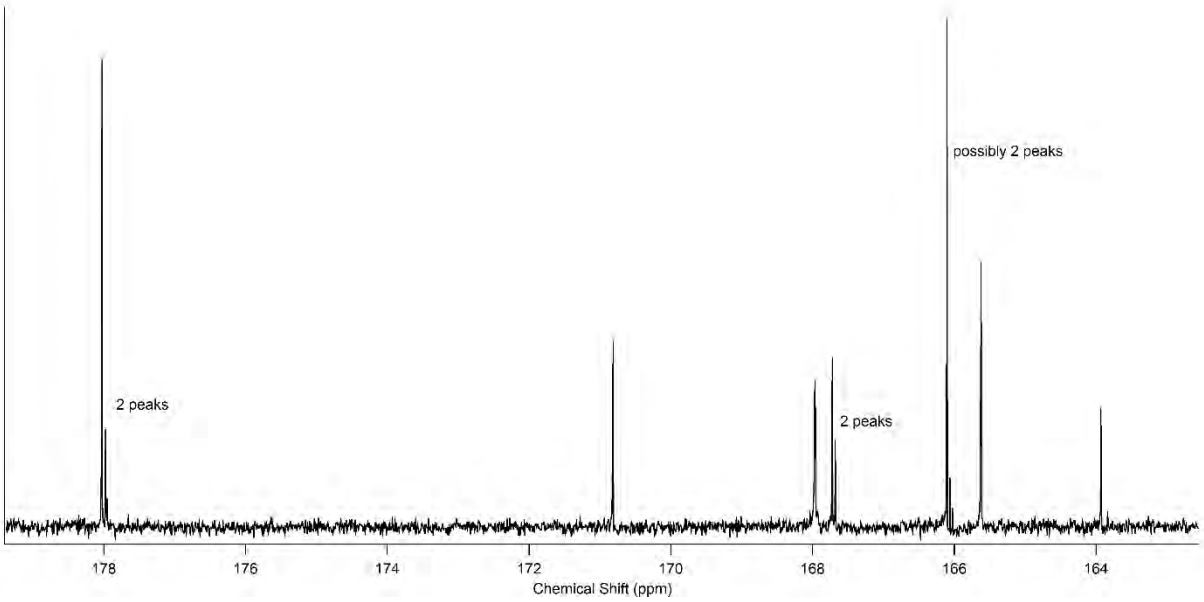
E. i.


180
160
140
120
100
80
60
40
Chemical Shift (ppm)

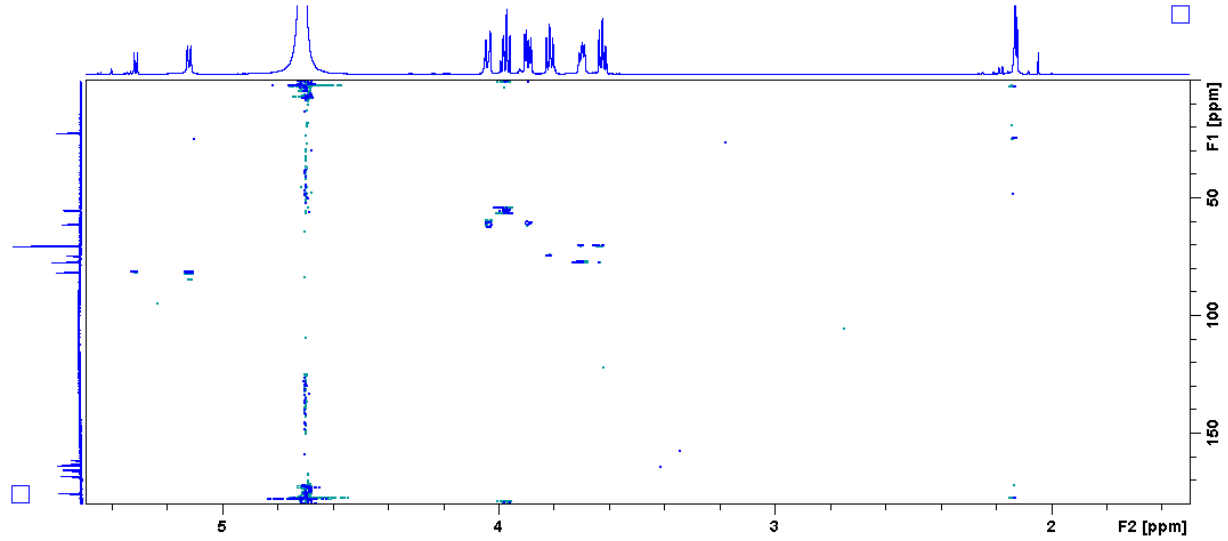
ii.


2 peaks
2 peaks
2 peaks
2 peaks
3 peaks:
1 broad, 2 sharp
85
80
75
70
65
60
55
50
45
40
35
30
25
Chemical Shift (ppm)

iii.


possibly 2 peaks
2 peaks
2 peaks
178
176
174
172
170
168
166
164
Chemical Shift (ppm)

F.



G.

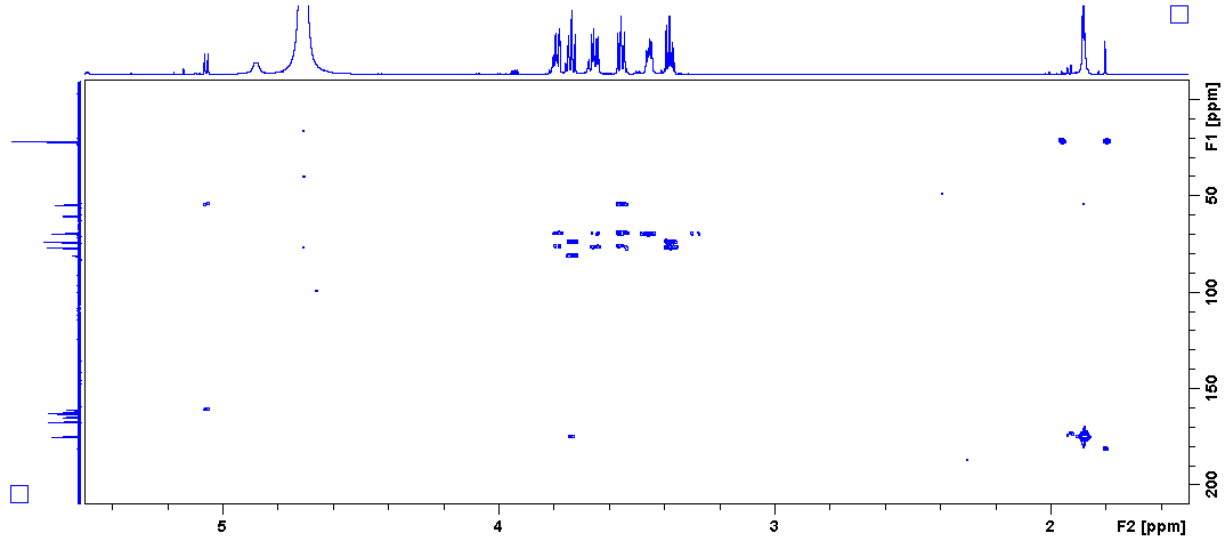
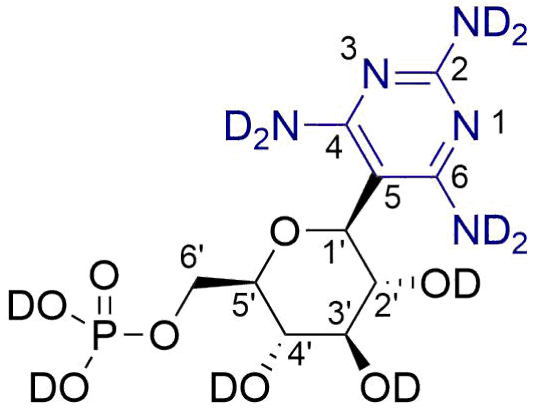


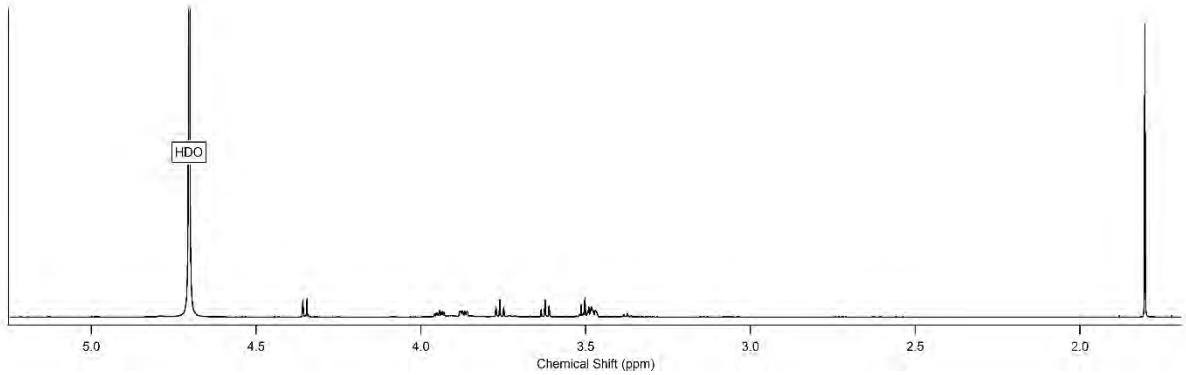
Fig. S21. NMR analysis of purified TAP *N*-acetylglucosamine glycosides (TAP-GlcNAc). **(A)** Proposed structures of the isolated TAP-GlcNAc species. **(B)** ^{1}H NMR spectrum of TAP-GlcNAc species. All peak assignments made through COSY (see below). **(i)** Full ^{1}H NMR spectrum. **(ii)** Detail of anomeric region of ^{1}H NMR spectrum. **(iii)** Detail of sugar peak region of ^{1}H NMR spectrum. **(C)** 1D ROE analysis of anomeric signals of TAP-GlcNAc species. **(i)** 1D ROE analysis of downfield anomeric signal, measured at 25°C. **(ii)** 1D ROE analysis of upfield anomeric signal, measured at 50°C. **(D)** COSY specrum of TAP-GlcNAc products. **(E)** ^{13}C NMR spectrum of TAP-GlcNAc species. **(i)** Full ^{13}C NMR spectrum. **(ii)** Detail of upfield region of ^{13}C spectrum. **(iii)** Detail of downfield region of ^{13}C spectrum. **(F)** HSQC spectrum of TAP-GlcNAc species. **(G)** HMBC spectrum of TAP-GlcNAc species.

Figure S22. NMR analysis of purified TAP glucose-6-phosphate glycoside (TAP-Glc6P)

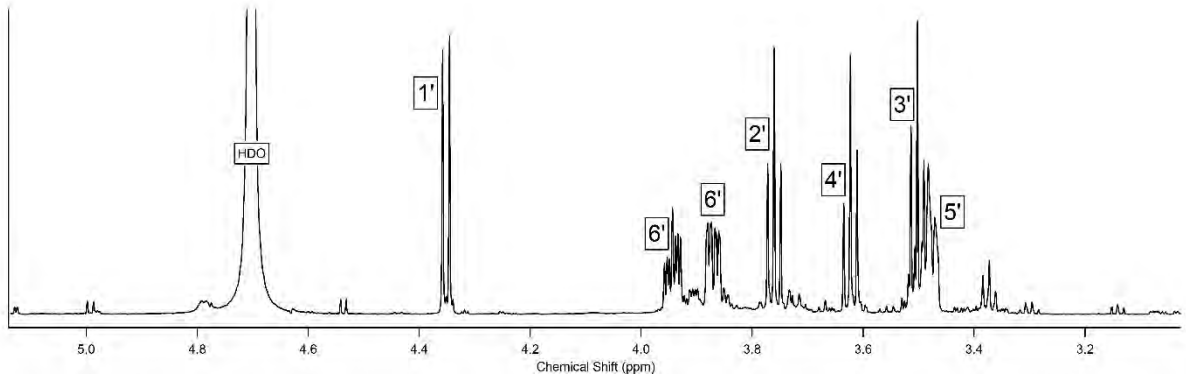
A.



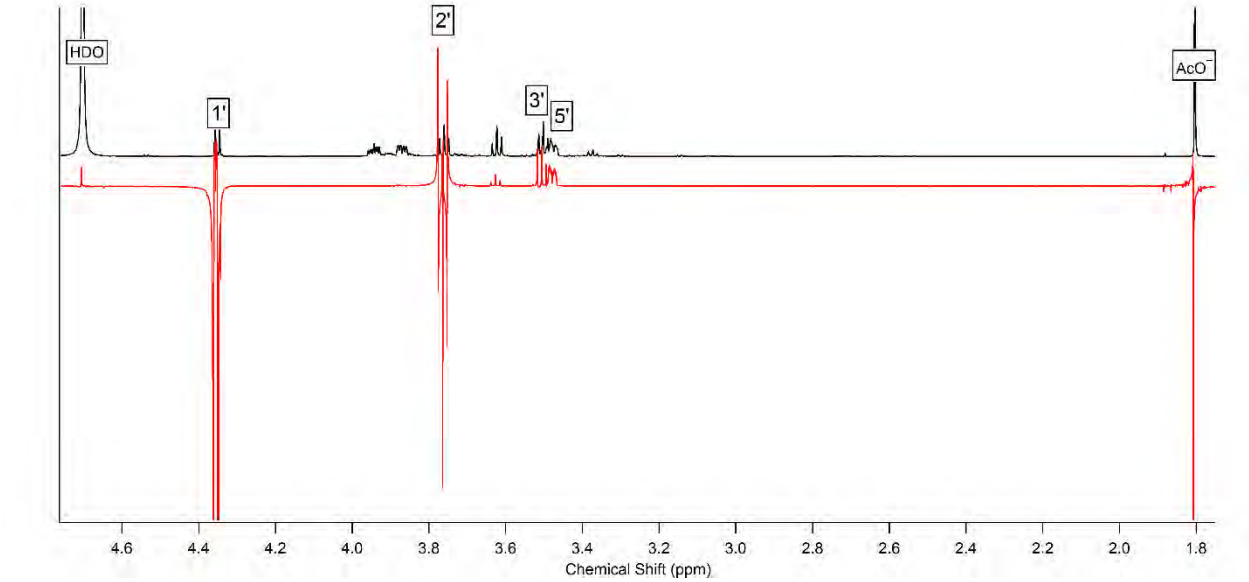
B. i.



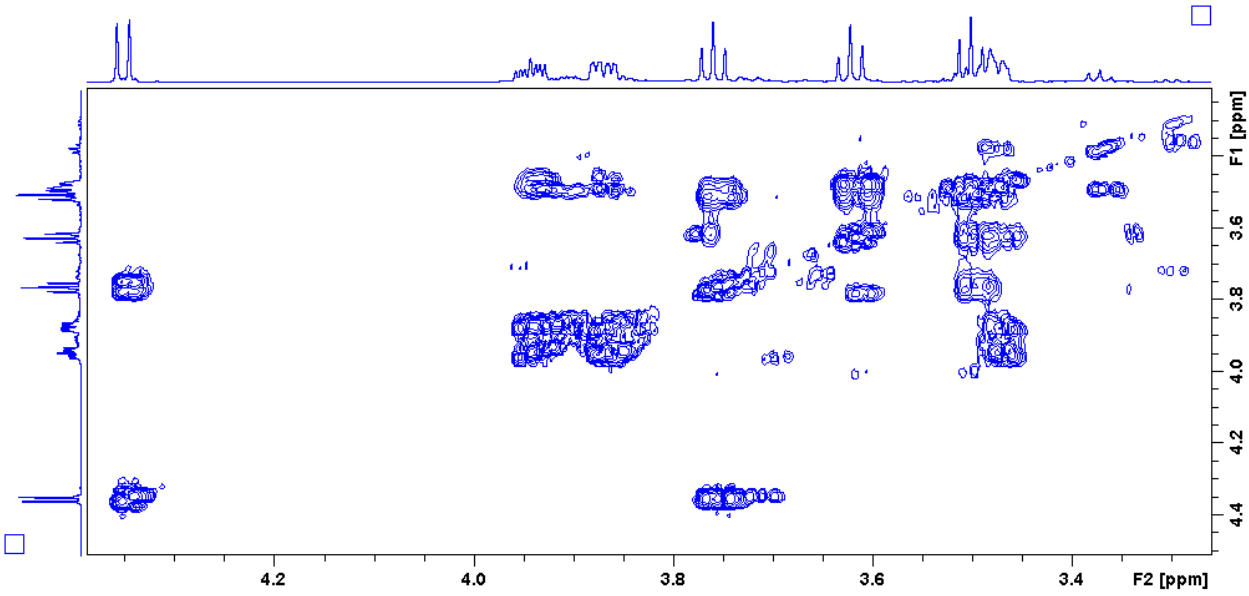
ii.



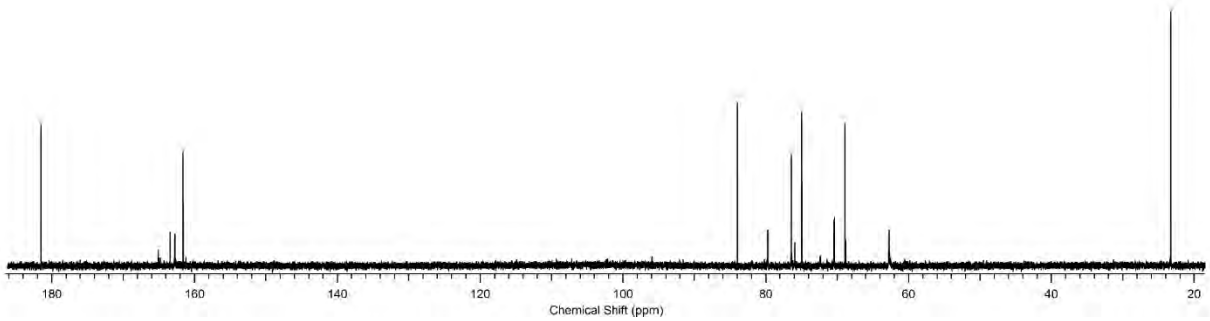
C.


HDO
1'
2'
3'
5'
AcO⁻
4.6
4.4
4.2
4.0
3.8
3.6
3.4
3.2
3.0
2.8
2.6
2.4
2.2
2.0
1.8
Chemical Shift (ppm)

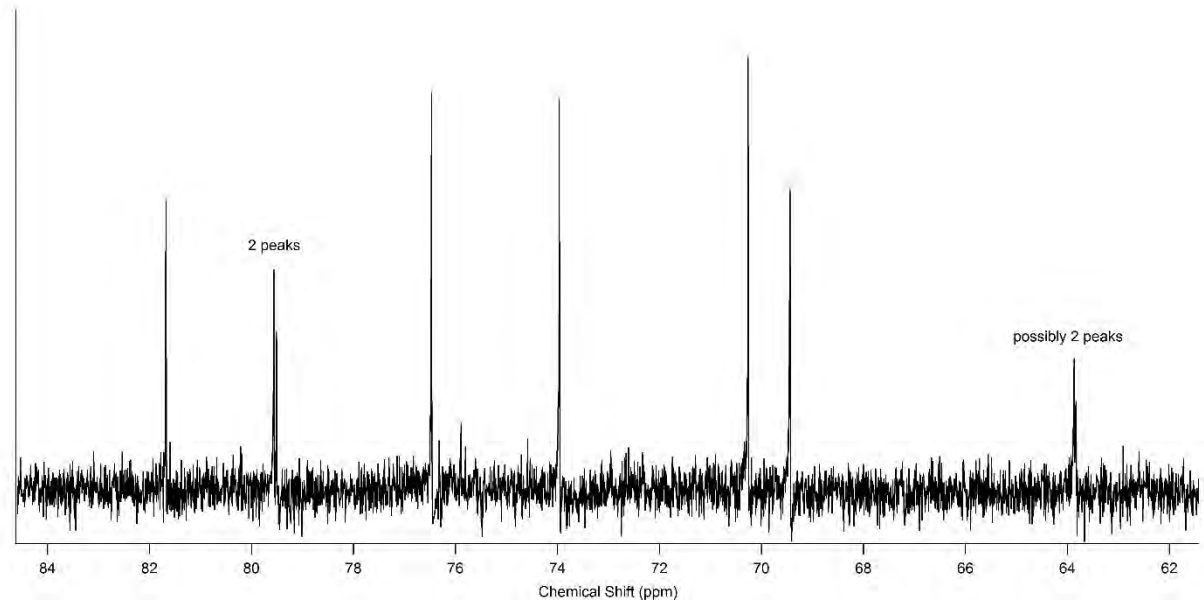
D.


F1 [ppm]
3.6
3.8
4.0
4.2
4.4
4.2
4.0
3.8
3.6
3.4
F2 [ppm]

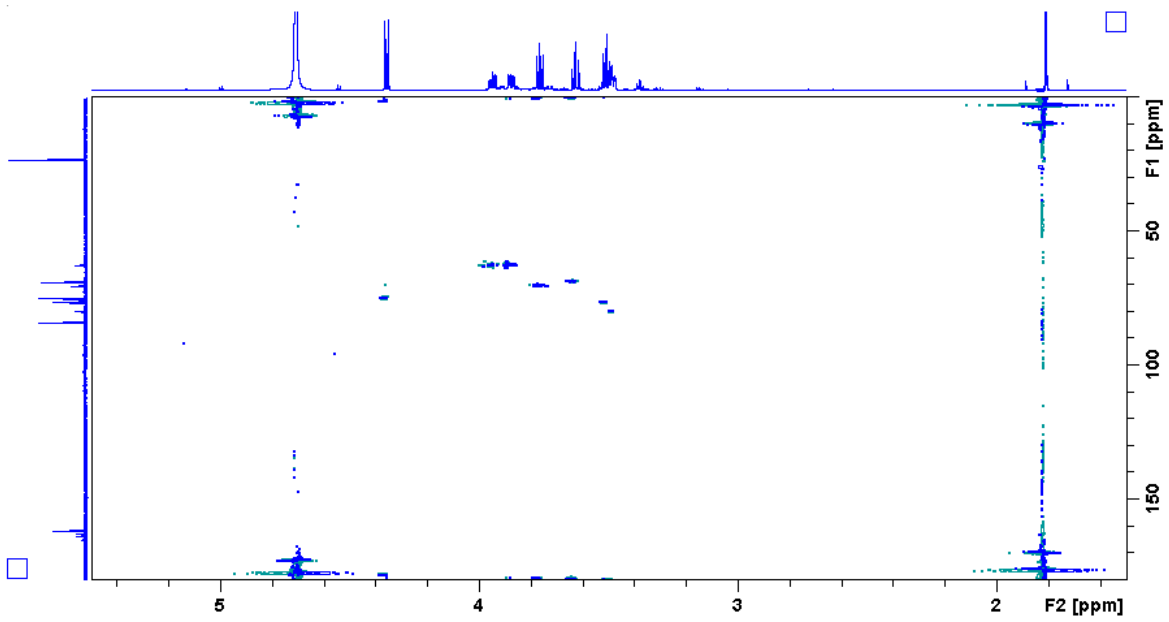
E. i.


180
160
140
120
100
80
60
40
20
Chemical Shift (ppm)

ii.


2 peaks
possibly 2 peaks
84
82
80
78
76
74
72
70
68
66
64
62
Chemical Shift (ppm)

F.



G.

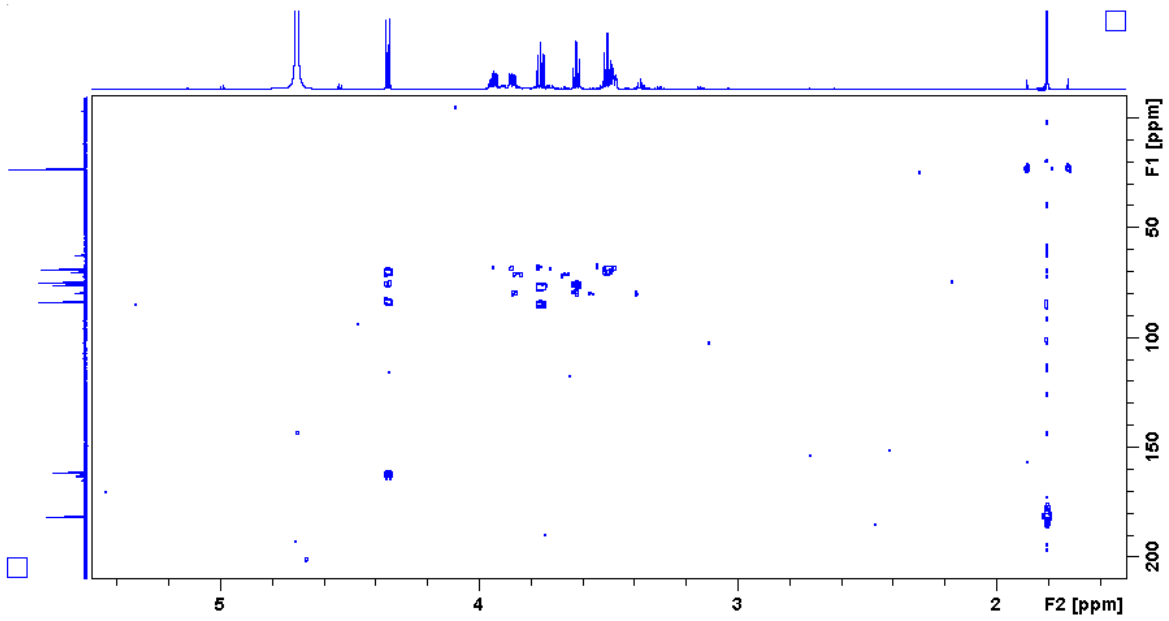


Fig. S22. NMR analysis of purified TAP glucose-6-phosphate glycoside (TAP-Glc6P). **(A)** Proposed structure of the isolated TAP-Glc6P species. **(B)** ^{1}H NMR spectrum of TAP-Glc6P. All peak assignments made through COSY (see below). **(i)** Full ^{1}H NMR spectrum. The large peak at 1.8 ppm represents acetate from the ammonium acetate buffer used in the purification of TAP-Glc6P. **(ii)** Sugar peak region of ^{1}H NMR. **(C)** 1D ROE analysis of anomeric signals of TAP-Glc6P. Saturation of the anomeric proton (labeled 1') gives ROEs for protons 3' and 5', with TOCSY to proton 2'. **(D)** COSY specrum of TAP-Glc6P. **(E)** ^{13}C NMR spectrum of TAP-Glc6P. **(i)** Full ^{13}C NMR spectrum. **(ii)** Detail of upfield region of ^{13}C spectrum. **(F)** HSQC spectrum of TAP-Glc6P species. **(G)** HMBC spectrum of TAP-Glc6P species.

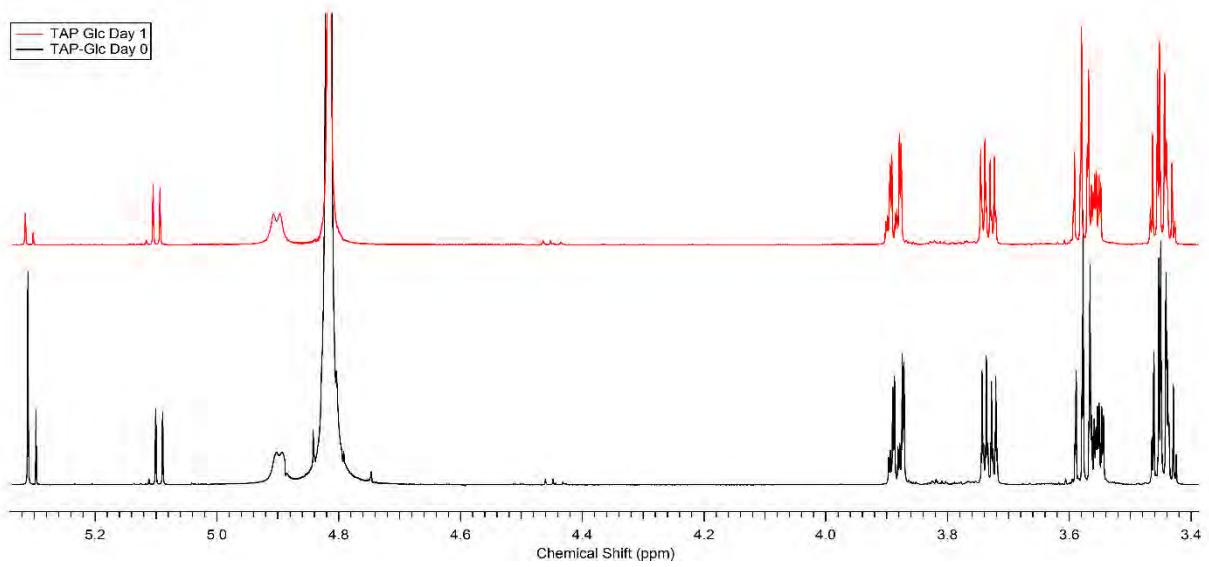


Fig. S23. Evolution of ^{1}H NMR spectra of purified TAP-glucose over 1 day at 5°C. When a solution of TAP-glucose glycosides is recorded immediately after dissolution in D_2O, two sharp singlets can be observed at 5.3 ppm. After incubating at 5°C for 24 hours, the peak intensities have greatly diminished, suggesting slow exchange with the solvent.

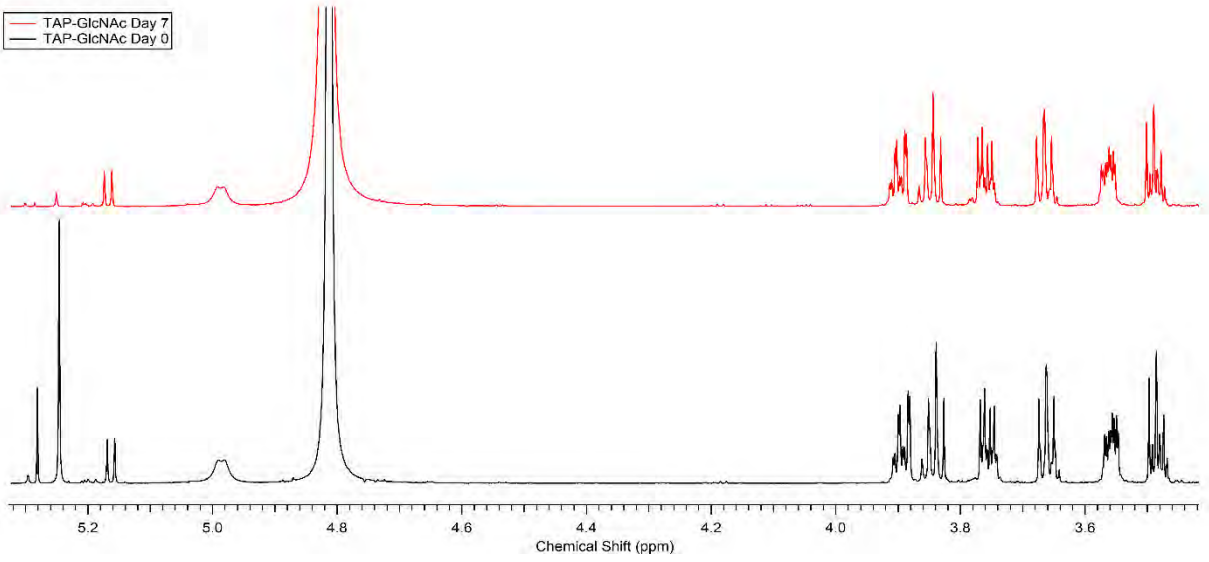


Fig. S24. Evolution of ^{1}H NMR spectra of purified TAP-GlcNAC over 7 days at 5°C. When a solution of TAP-GlcNAc glycosides is recorded immediately after dissolution in D_2O, two sharp singlets can be observed at 5.24-5.3 ppm. After incubating at 5°C for 7 days, the peak intensities have greatly diminished, suggesting slow exchange with the solvent.

III. Discussion of the Mechanism of TAP Glycosylation

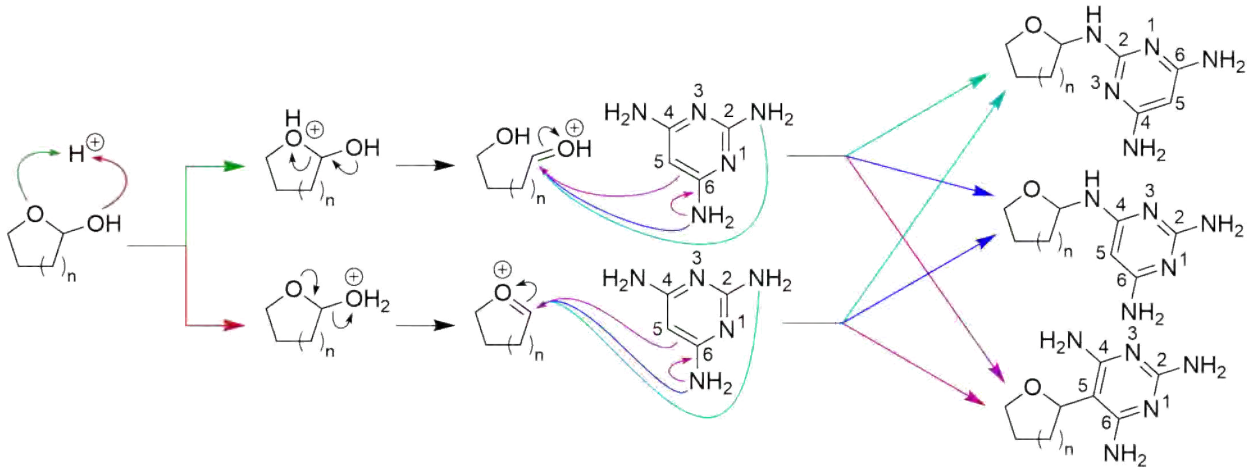


Fig. S25. Possible mechanisms of TAP glycosylation. Electrophilic activation of a sugar proceeds under acidic conditions either by an endocyclic C-O bond cleavage path (green) or an exocyclic C-O bond cleavage (red). Attack on the resulting oxocarbenium ion by TAP results in substitution at the exocyclic amine of TAP at position 2 (cyan), the exocyclic amine of TAP at position 4/6 (blue), or electrophilic aromatic substitution at position 5 of TAP (violet).

The possible mechanisms of TAP glycosylation are shown in Fig. S25. Glycosylation in water is expected to be acid-catalyzed, and may proceed either by protonation of the endocyclic oxygen atom, followed by endocyclic C-O bond cleavage to give a protonated aldehyde electrophile (as an acyclic-form sugar, green path), or by protonation of the exocyclic anomeric hydroxyl group, followed by exocyclic C-O bond cleavage to give an oxocarbenium electrophile (as a cyclic-form sugar, red path). Previous studies on aqueous glycosylation of heteroaromatic compounds[1] supported the endocyclic (green) mechanism that proceeds by attack of the nucleophilic heterocycle on the protonated aldehyde of an acyclic-form sugar. However, for pyranosides (glucose presents mainly as *α*- and *β*-pyranoses in water), solvolysis generally follows the exocyclic (red) oxocarbenium mechanism, suggesting that the reverse reaction, glycosylation, also follows this mechanism.[2] The dominant mechanism in the glycosylation of TAP by glucose and its derivatives is unclear. The acyclic path is reasonable for the formation of *β*-*N*-pyranosides, first proceeding through Schiff base formation with an exocyclic amine of TAP, followed by attack of the 5' hydroxyl group on the protonated Schiff base to form the glycosylamine. However, for *β*-*C*-pyranoside formation through an acyclic sugar, a Knoevenagel condensation-type intermediate must be formed which would have severe 1,3-allylic strain. The formation of *β*-*C*-pyranosides is not ameliorated through the cyclic path, which, through consideration of the most stable half-chair conformation of the oxocarbenium species, is expected to give *α*-pyranosides. It may be that *β*-*C*-pyranosides are formed over *α*-*C*-

pyranosides due to the steric bulk of a TAP nucleophile in electrophilic aromatic substitution at C5. Although attack on the *β* face of the oxocarbenium species would initially give a twist-boat conformation, this is perhaps sterically more permissible than attack on the *α* face, which would produce a chair conformation with severe 1,3-diaxial interactions.

The apparent enhancement in yield of the *β*-*C*-pyranoside of Glc6P over *β*-*N*-pyranosides may be due to stabilization of the transition state in the electrophilic aromatic substitution step of the cyclic path. It is possible that the phosphate of Glc6P acts to dock an incoming TAP nucleophile in the proper conformation for electrophilic aromatic substitution by the exocyclic cleavage path (red) in a manner not possible for glucose or GlcNAc.

Although anchimeric assistance by the acetamido group of GlcNAc is expected to enhance *β*-substitution, the formation of this more stabilized cation may not provide sufficient electrophilicity to incite electrophilic aromatic substitution of TAP, thus preventing the formation of the *β*-*C*-pyranoside.

IV. Comparison of Estimated Glycoside Yields as Determined by ^{1}H NMR and LC-MS/UV.

Table S1 TAP-glycoside yields as estimated by ^{1}H NMR and LC-MS/UV

Sugar	pH	Glycoside yields est. by ^{1}H NMR[a]	Monoglycoside yields est. by LC-MS/UV[b]
Hexoses			
Glucose	1	31%	40
	7	44%	40
GlcNAc	1	28%	22
	7	5%	17
Glc6P	1	27%	49
	7	22%	43
Fructose[b]	1	2%	2
	7	12%	27
Galactose	1	35%	37
	7	48%	56
Galacturonic Acid	1	40%	52
	7	53%	49
Glucosamine	1	10%	5
	7	15%	0
Glucuronic Acid	1	42%	40
	7	24%	44
Mannose	1	20%	46
	7	13%	43
Pentoses			
Arabinose	1	61%	57
	7	55%	59
Lyxose	1	20%	44
	7	40%	52
Ribose	1	28%	46
	7	31%	50
R5P	1	11%	45
	7	8%	17
Ribulose[b]	1	3%	17
	7	31%	46
Xylose	1	27%	43
	7	34%	57
Tetroses			
Erythrose	1	<1%	X[c]
	7	17%	
Threose	1	3%	
	7	2%	

[a]Based on integrated intensity of new resonances (not ureacted sugar) in the anomeric region of ^{1}H NMR spectra relative to a internal concentration standard.
[b]Based on integrated intensity (at 272 nm) of all peaks with the expected m/z value for TAP with a single conjugated sugar relative the total integrated intensity of peaks with a 272 nm absorbance.
[c]Estimated yields for the tetroses, erythrose and threose, are not reported because HPLC peaks containing monoglycosides overlap with peaks of unidentified products and 4,6-diamino-2-hydroxypyrimidine (DAHP, a hydrolysis product of TAP).

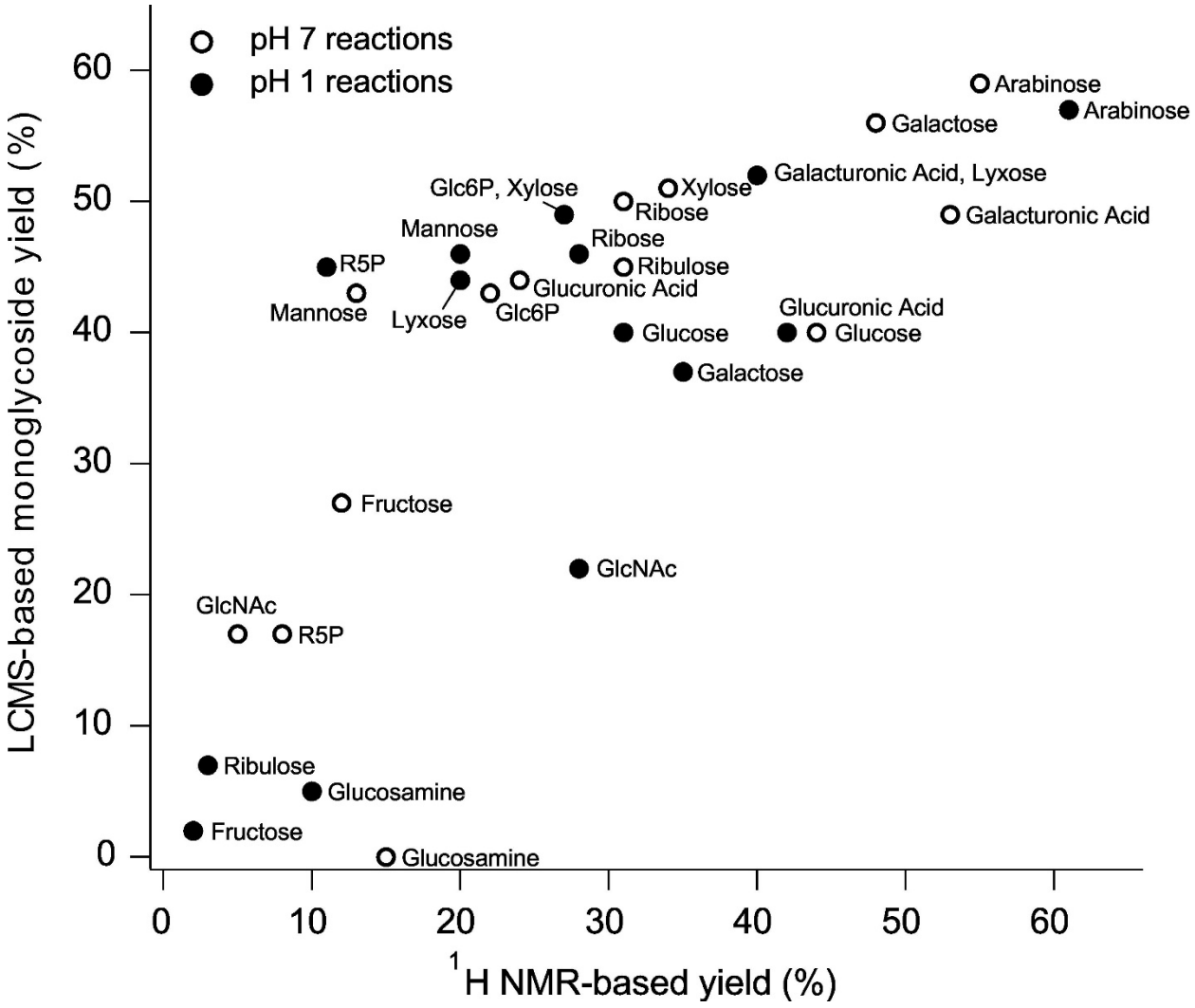


Fig. S26. Plot of estimated TAP-glycoside yields based on integrated absorption at 272 nm of peaks observed in LC-MS/UV chromatograms with m/z values corresponding to single sugar conjugates of TAP versus estimated TAP-glycoside yields based on integration of new (not unreacted sugar) resonances in the anomeric regions of ^{1}H NMR spectra. LC-MS/UV-based yields are relative to total integrated intensity of peaks with absorption at 272 nm (i.e., total recovered TAP and TAP conjugates). ^{1}H NMR-based yields are determined by comparison of integrated intensities to the integrated intensities of the resonances of an internal standard of known concentration.

V. AFM Analysis of TAP Glycoside Assemblies with Cyanuric Acid and CyCo6

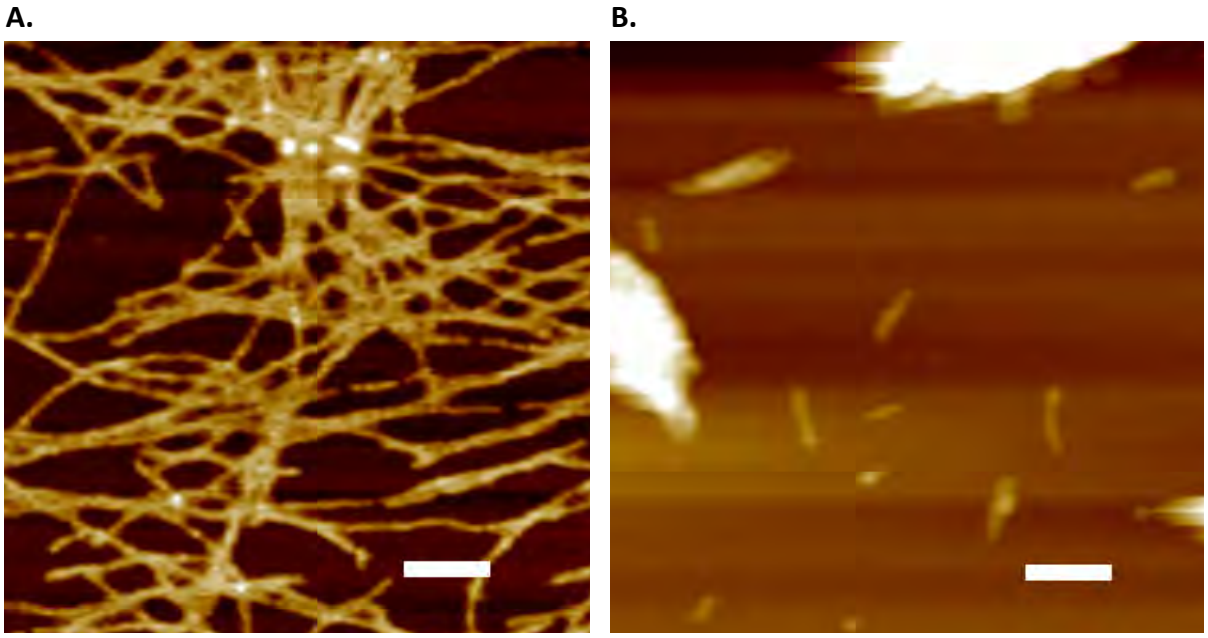


Fig. S27. AFM images of structures formed by TAP-Glc6P reaction products and cyanuric acid (CA). **(A)** Linear assemblies formed upon the mixing of crude TAP-Glc6P products of the (pH 1) reaction with CA (50 mM in total TAP species, with 50 mM CA, pH 7, and approximately 0.3 M NaCl from initial reaction pH adjustment with HCl and subsequent solution pH adjustment with NaOH). Gelation and precipitation were observed prior to deposition. Scale bar indicates 100 nm. **(B)** Representative image showing variety of irregular structures formed when purified TAP-Glc6P is mixed with CA (50 mM total TAP species, 50 mM CA, pH 7, approximately 0.1 M NH_4OAc from purification of TAP-Glc6P). Precipitation observed prior to deposition. Scale bar indicates 100 nm.

A. **B.**

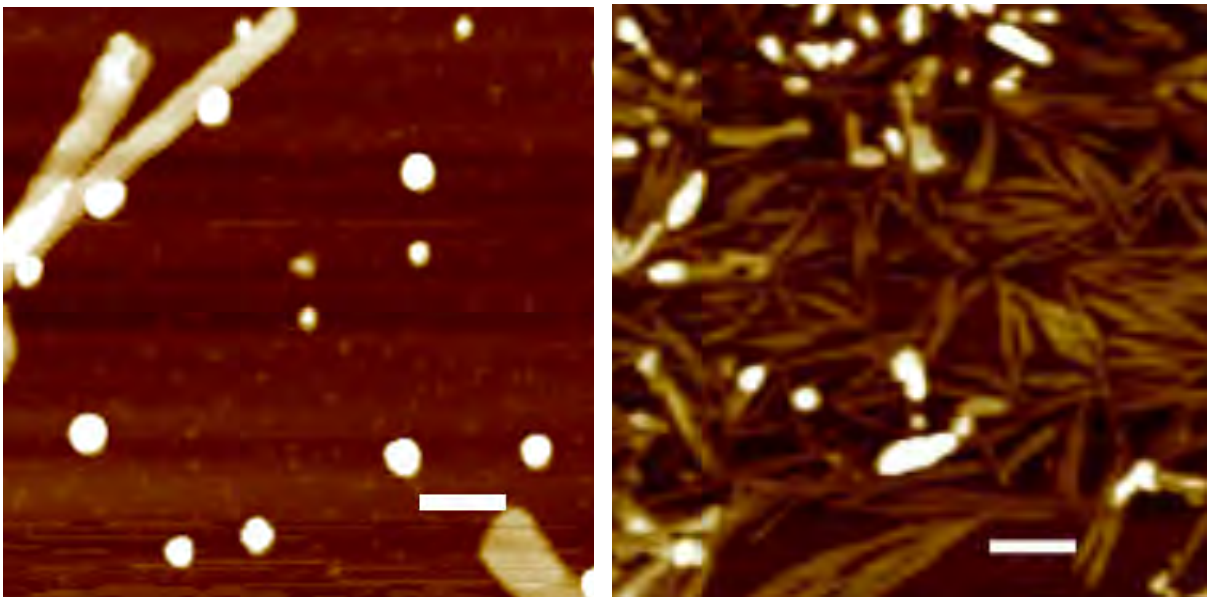


Fig. S28. AFM images of structures formed by products of the TAP-glucose reaction when mixed with CyCo6 (CA conjugated with hexaonic acid). **(A)** Structures formed upon the mixing of the crude products of the TAP-glucose (pH 1) reaction with CyCo6 (50 mM in total TAP species, 50 mM CyCo6, pH 7, and approximately 0.3 M NaCl from initial reaction pH adjustment with HCl and subsequent solution adjustment with NaOH). Precipitation observed prior to deposition. **(B)** Structures observed when purified TAP-glucosides are mixed with CyCo6 (50 mM total TAP glycosides, 50 mM CyCo6, pH 7, and approximately 0.1 M NaCl from pH adjustment with NaOH and HCl). No precipitation or gelation observed prior to deposition. Scale bar is 100 nm in both images.

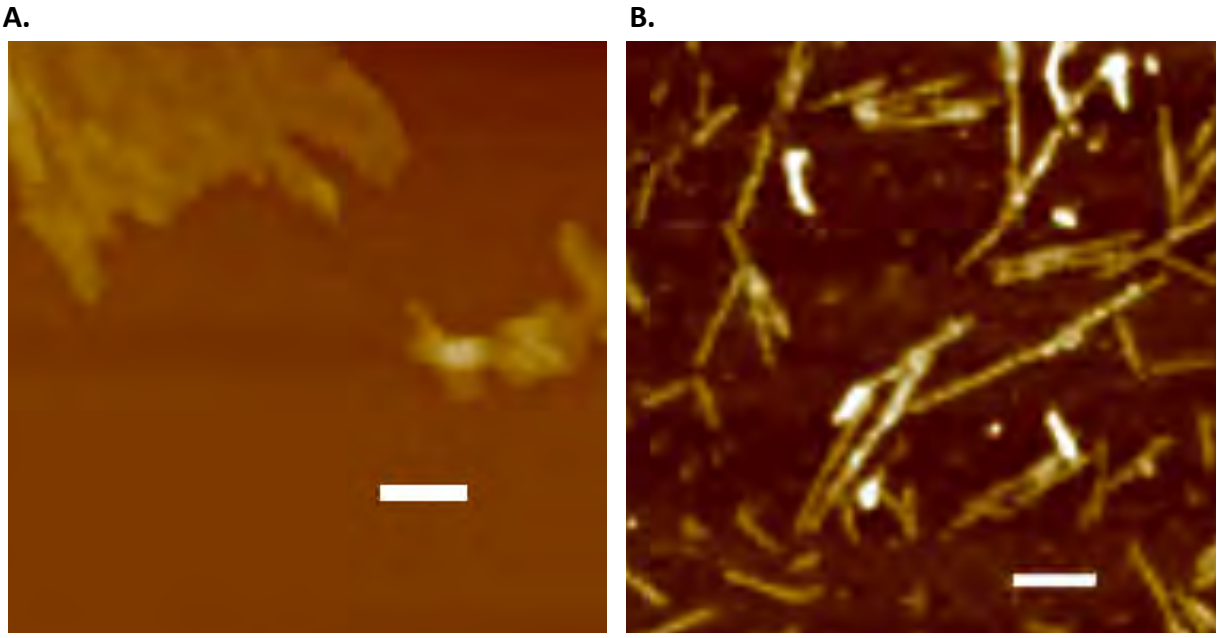


Fig. S29. AFM images of structures formed when products of the TAP-GlcNAc reaction were mixed with CyCo6. **(A)** Structures formed upon the mixing of the crude TAP-GlcNAc reaction (pH 1) products with CyCo6 (50 mM in total TAP species, 50 mM CyCo6, pH 7, and approximately 0.3 M NaCl from initial reaction pH adjustment with HCl and subsequent solution adjustment with NaOH). Precipitation was observed prior to deposition. **(B)** Structures formed when purified TAP-GlcNAc glycosides were mixed with CyCo6 (50 mM in total TAP glycosides, 50 mM CyCo6, pH 7, and approximately 0.1 M NaCl from pH adjustment with NaOH and HCl). No precipitation or gelation observed prior to deposition. Scale bar is 100 nm in both images.

VI. Supplementary References

1. J. P. Dworkin and S. L. Miller, *Carbohydr. Res.*, 2000, **39**, 359-365.

2. N. A. Horenstein, *Ad. Phys. Org. Chem.*, 2006, **41**, 275-314.