JOC Notes

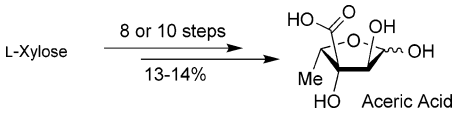
Synthesis of the Branched-Chain Sugar Aceric Acid: A Unique Component of the Pectic Polysaccharide Rhamnogalacturonan-II†

Nigel A. Jones,‡ Sergey A. Nepogodiev,*,‡ Colin J. MacDonald,‡ David L. Hughes,‡ and Robert A. Field*,‡,§

School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich, NR4 7TJ, U.K., and Department of Biological Chemistry, John Innes Centre, Norwich Research Park, Norwich, NR4 7UH, U.K.

s.nepogodiev@uea.ac.uk; r.a.field@uea.ac.uk

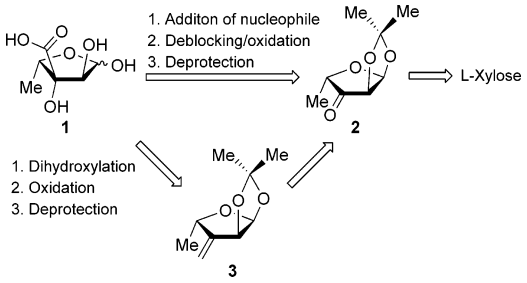
Received May 20, 2005



Described herein is the synthesis of 3-*C*-carboxy-5-deoxy-L-xylose (aceric acid), a rare branched-chain sugar found in the complex pectic polysaccharide rhamnogalacturonan-II. The key synthetic step in the construction of aceric acid was the stereoselective addition of 2-trimethylsilyl thiazole to 5-deoxy-1,2-*O*-isopropylidene-α-L-*erythro*-pentofuran-3-ulose (**2**), which was prepared from L-xylose. The thiazole group was efficiently converted into the required carboxyl group via conventional transformations. Aceric acid was also synthesized by dihydroxylation of a 3-*C*-methylene derivative of **2** followed by oxidation of the resulting hydroxylmethyl group. The C-2 epimer of aceric acid was also synthesized using thiazole addition chemistry, starting from L-arabinose.

The complex pectic polysaccharide rhamnogalacturonan-II (RG-II), which is found in the cell walls of all higher plants, plays a key role in plant growth, development, and resistance to disease.[1] The backbone of RG-II contains at least eight α-(1 → 4)-linked D-Gal*p*A residues to which are attached four structurally distinct oligosaccharide side chains (designated A−D).[2] It has been shown that relatively minor changes in the structure of the side chains can cause severe plant growth defects.[3] These defects have been attributed to changes in primary cell wall architecture, which is highly dependent on the ability of RG-II to exist in a dimeric form that is covalently cross-linked through a borate diester.[3] Only one particular monosaccharide residue of each RG-II molecule, namely β-D-apiofuranose of side chain A, is thought to be involved in borate diester formation.[4] Side chain B of RG-II also contains an apiofuranose residue which is, however, not thought to be involved in borate cross-linking.

SCHEME 1. Retrosynthetic Analysis for Synthesis of Aceric Acid



Fragments of chains A and B will be useful for investigating the borate complexation phenomenon; synthesis of several fragments of side chain A has recently been reported.[5] Construction of side chain B requires access to a unique branched-chain acidic monosaccharide 3-*C*-carboxy-5-deoxy-L-xylose, referred to as aceric acid (Ace*f* **1**, Scheme 1).[6] This branched sugar has only been isolated[6] in analytical quantities from RG-II itself and has not previously been prepared by chemical synthesis. We describe herein approaches to the synthesis of aceric acid and its C-2 epimer from readily available monosaccharide precursors.

The construction of aceric acid was investigated using two different pathways, according to the retrosynthetic analysis outlined in Scheme 1. In both cases L-xylose, which has the correct stereochemistry at C-2 and C-4, can serve as a convenient starting monosaccharide. L-Xylose was trapped in the furanose form as a 1,2-*O*-isopropylidene derivative[7] and then transformed into 5-deoxy-3-ulose derivative **2**, as reported[8] for the D-enantiomer. One commonly used approach in the synthesis of branched-chain sugars with a branching ar-

† Dedicated to Professor Steve Ley on the occasion of his 60th birthday.

‡ University of East Anglia.

§ John Innes Centre.

(1) (a) Höfte, H. *Science* **2001**, *294*, 795−797. (b) Ridley, B. L.; O'Neill, M. A.; Mohnen, D. A. *Phytochemistry* **2001**, *57*, 929−967. (c) O'Neill, M. A.; Ishii, T.; Albersheim, P.; Darvill, A. G. *Annu. Rev. Plant Biol.* **2004**, *55*, 109−139.

(2) (a) Darvill, A. G.; McNeil, M.; Albersheim, P. *Plant Physiol.* **1978**, *62*, 418−422. (b) Glushka, J. N.; Terrell, M.; York, W. S.; O'Neill, M. A.; Gucwa, A.; Darvill, A. G.; Albersheim, P.; Prestegard, J. H. *Carbohydr. Res.* **2003**, *338*, 341−352. (c) Rodríguez-Carvajal, M. A.; Hervé du Penhoat, C.; Mazeau, K.; Doco, T.; Pérez, S. *Carbohydr. Res.* **2003**, *338*, 651−671. (d) Hervé du Penhoat, C.; Gey, C.; Pellerin, P.; Pérez, S. *J. Biomol. NMR* **1999**, *14*, 253−271.

(3) O'Neill, M. A.; Eberhand, S.; Albersheim, P.; Darvill, A. G. *Science* **2001**, *294*, 846−849.

(4) Ishii, T.; Matsunaga, T.; Pellerin, P.; O'Neill, M. A.; Darvill, A.; Albersheim, P. *J. Biol. Chem.* **1999**, *274*, 13098−13104.

(5) (a) Chauvin, A.-L.; Nepogodiev, S. A.; Field, R. A. *Carbohydr. Res.* **2004**, *339*, 21−27. (b) Buffet, M. A. J.; Rich, J. R.; McGavin, R. S.; Reimer, K. B. *Carbohydr. Res.* **2004**, *339*, 2507−2513. (c) Chauvin, A.-L.; Nepogodiev, S. A.; Field, R. A. *J. Org. Chem.* **2005**, *70*, 960−966.

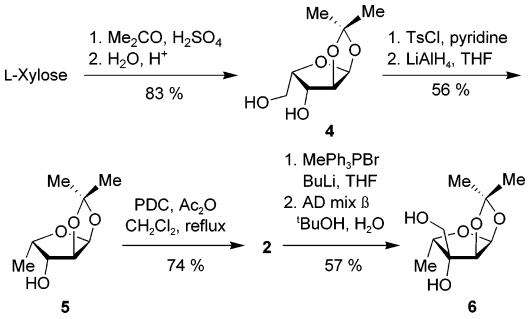
(6) Spellman, M. W.; McNeil, M.; Darvill, A. G.; Albersheim, P. *Carbohydr. Res.* **1983**, *122*, 115−129.

(7) Moravcová, J.; Capková, J.; Stanek, J. *Carbohydr. Res.* **1994**, *263*, 61−66.

(8) Tronchet, J. M. J.; Graf, R. *Helv. Chim. Acta* **1972**, *55*, 613−622.

10.1021/jo051012b CCC: $30.25
Published on Web 09/09/2005

SCHEME 2

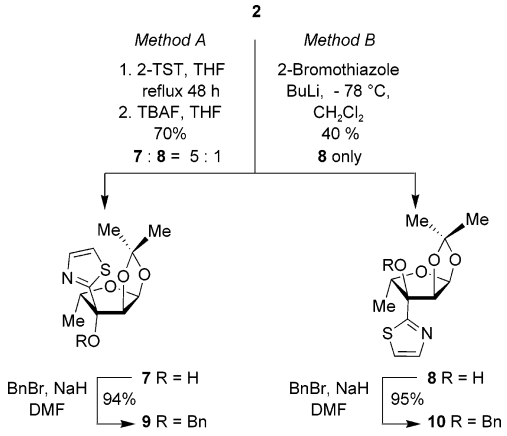


rangement similar to that found in aceric acid involves addition of a nucleophile to an ulose.[9] However, potentially useful Grignard[10] and organolithium[11] reagents are not likely to lead to the correct C-3 stereochemistry in this example. The bicylo[3.3.0]octane ring system present in 3-ulose **2** is expected to favor *exo*-addition of the nucleophile to the keto group on steric grounds. An example of highly stereoselective *endo*-addition to a bicyclic 3-ketofuranose was reported[12] in a reaction utilizing 2-(trimethylsilyl)thiazole (2-TST) as a nucleophile. Since the 2-TST group is a formyl group equivalent,[13] it is also a convenient precursor of the carboxylic group, which can be easily generated by standard oxidation of the formyl group.

On the other hand, asymmetric dihydroxylation[14] of alkene **3**, which can be prepared from 3-ulose **2** (Scheme 1) via a Wittig reaction, presents an alternative route to aceric acid. In this approach, dihydroxylation should proceed from the more accessible *exo*-face of the bicyclic system, affording a branched-chain derivative with the correct orientation of the 3-*C*-hydroxymethyl group relative to the furanose ring. Selective oxidation of the primary OH group may then provide the required 3-*C*-carboxy-derivative. Due to its simplicity, this approach was investigated first.

1,2-*O*-Isopropylidene-L-xylofuranose (**4**), which is readily accessible in 83% yield via a two-step procedure[7] starting from L-xylose, was converted into 5-deoxy sugar **5** in 56% yield over two steps involving tosylation of the primary OH group and reduction with $LiAlH_4$ (Scheme 2). Compound **5** was oxidized with pyridinium dichromate to afford the unstable 3-ulose **2** in a 74% yield, which was converted into the 3-*C*-methylene derivative **3**, a low boiling liquid, in 70% yield by reaction with $H_2C{=}PPh_3$. Dihydroxylation of alkene **3** using Sharpless conditions[14] afforded the same single stereoisomer **6**, regardless of whether AD mix α or AD mix *β* was used. One can anticipate that the preferred face for dihydroxylation of olefin **3** is *exo* with respect to the bicyclic ring system, and therefore the reaction product **6** was tentatively assigned as having the L-*xylo*-configuration. However, 2D NOESY NMR spectroscopy data for **6**, analyzed in the absence of data for stereoisomeric compounds, were inconclusive, leaving this assignment ambiguous. Furthermore, a literature precedent[15] shows that dihydroxylation of a similar 3-*C*-alkene derivative of a 1,2-*O*-isopropylidine derivative of a pentofuranose results in a diol having the newly formed hydroxymethyl group *anti* with respect to the isopropylidene group, indicative of dihydroxylation of the more hindered *endo*-face of the alkene. Due to the potential ambiguity in the structure of **6**, an alternative approach to installing the chiral quaternary center was considered.

SCHEME 3



To investigate a thiazole addition approach to aceric acid, 3-ulose **2** was treated with 2-TST in THF, under thermodynamic control, followed by desilylation with TBAF. This reaction sequence led to a 5:1 mixture of *endo*- and *exo*-adducts **7** and **8**, respectively, in an overall yield of 70% (Scheme 3, Method A). In contrast, 2-thiazolyl lithium addition to compound **2** gave only the adduct **8** arising from sterically controlled *exo*-addition to the trioxabicylo[3.3.0]octane ring system in an unoptimized 40% yield (Scheme 3, Method B). Both isomeric adducts **7** and **8** were benzylated to afford 3-*O*-benzyl derivatives **9** and **10**, respectively. The stereochemistry of thiazole addition to 3-ulose **2** was established by X-ray structural analysis of compounds **7** and **10** (Figure 1), which revealed that **10** has the required C-3 configuration for aceric acid synthesis.

The stereoselectivity of the reaction of 3-ketofuranose **2** with 2-thiazolyl lithium can be attributed to kinetic control of the reaction and steric hindrance to addition to the *endo*-face of the ketone. The mechanism of the reaction of **2** with 2-TST is more complex: as noted for a similar system,[12] it is likely to involve several steps, including equilibration of thiazolium anion intermediates, which allows stereodifferentiation between the *endo* and *exo*-faces of the bicyclic molecule.

(9) For reviews on the synthesis of branched-chain sugars, see: (a) Yoshimura, J. *Adv. Carbohydr. Chem. Biochem.* **1984**, *42*, 69−134. (b) Chapleur, Y.; Chrétien, F. In *Preparative Carbohydrate Chemistry*; Hanessian, S., Ed.; Marcel Dekker: New York, 1997; pp 207−262. (c) Sviridov, A. F.; Shmyrina, A. Y.; Chizhov, O. S.; Kochetkov, N. K. *Bioorg. Khim.* **1982**, *8*, 293−325.

(10) Dyer, J. R.; McGonigal, W. E.; Rice, K. C. *J. Am. Chem. Soc.* **1965**, *87*, 654−655.

(11) (a) Dahlman, O.; Garegg, P. J.; Mayer, H.; Schramek, S. *Acta Chem. Scand.* **1986**, *40*, 15−20. (b) Sharma, G. V. M.; Reddy, V. G.; Krishna, P. R.; Sankar, A. R.; Kunwar, A. C. *Tetrahedron* **2002**, *58*, 3801−3812. (c) Dondoni, A.; Marra, A.; Scherrmann, M. C. *Tetrahedron Lett.* **1993**, *34*, 7323−7326.

(12) Carcano, M.; Vasella, A. *Helv. Chim. Acta* **1998**, *81*, 889−901.

(13) (a) Dondoni, A. *Pure Appl. Chem.* **1990**, *62*, 643−652. (b) Dondoni, A.; Marra, A. *Chem. Rev.* **2004**, *104*, 2557−2599.

(14) Sharpless, K. B.; Amberg, W.; Bennani, Y. L.; Crispino, G. A.; Hartung, J.; Jeong, K.; Kwong, H.; Morikawa, K.; Wang, Z.; Xu, D.; Zhang, X. *J. Org. Chem.* **1992**, *57*, 2768−2771.

(15) Houseknecht, J. B.; Lowary, T. L. *J. Org. Chem.* **2001**, *67*, 4150−4164.

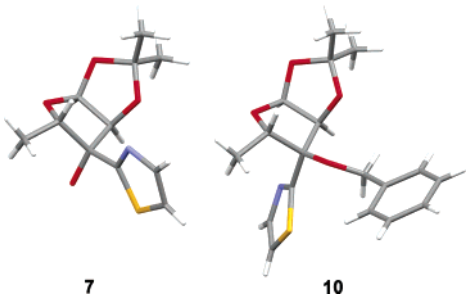


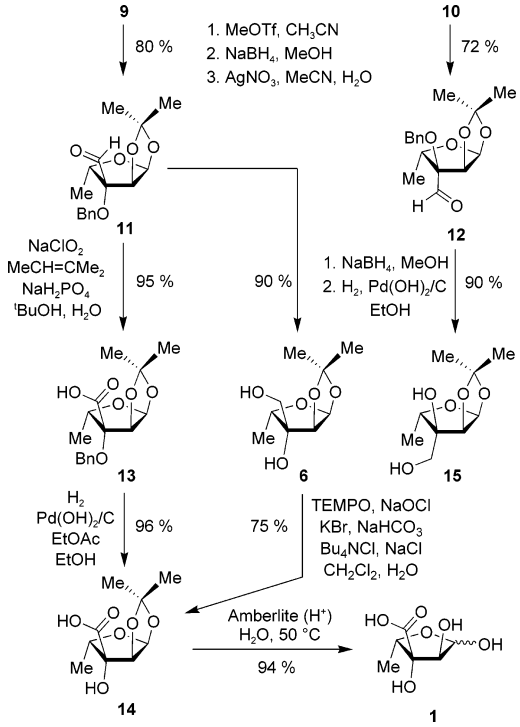
FIGURE 1. X-ray crystal structures of *endo*-thiazolyl derivative **7** and 3-*O*-benzylated *exo*-thiazolyl derivative **10**.

Conversion of the thiazolyl group of compounds **9** and **10** into the formyl group was achieved using essentially a one-pot, three-step literature method[16] involving *N*-methylation ($MeOSO_2CF_3$), reduction ($NaBH_4$), and metal-promoted hydrolysis ($AgNO_3$ in aqueous MeCN) (Scheme 4). Aldehydes **11** and **12** were prepared this way in 80 and 72% yield, respectively. Oxidation of aldehyde **11** to carboxylic acid **13**, followed by removal of protecting groups by successive catalytic hydrogenolysis (compound **14**) and acid hydrolysis, afforded target aceric acid **1** in an overall 86% yield starting from **11**. The structure of synthetic aceric acid **1** was confirmed by ^1H and ^{13}C NMR data, which were in good agreement with those reported for natural aceric acid isolated from primary cell walls of sycamore (*Acer pseudoplantanus*).[6]

Certainty about the stereochemistry of aldehydes **11** and **12** (arising from crystal structures of **7** and **10**, respectively) made it possible to confirm the C-3 configuration of compound **6** obtained by the dihydroxylation approach (Scheme 2). To make this comparison, aldehydes **11** and **12** were converted into 3-*C*-hydroxymethyl derivatives **6** and **15**, respectively, via $NaBH_4$ reduction followed by removal of 3-*O*-benzyl group by catalytic hydrogenolysis. Diol **6** synthesized from the *endo*-thiazolyl derivative **9** proved to be identical to that obtained earlier from the dihydroxylation route. Completion of aceric acid synthesis route required selective oxidation of the primary hydroxyl group which was successfully carried out using a TEMPO-based procedure,[17] but purification of carboxylic acid **14** obtained by this route proved difficult. Hence, although requiring fewer steps, the dihydroxylation route to aceric acid proved less practical than the 2-trimethylsilylthiazole addition route.

The aceric acid residue present in RG-II (Scheme 1) has a synthetically challenging 1,2-*cis*-configuration. Due to the particular difficulty in creating a 1,2-*cis*-glycofuranosidic bond,[18] an indirect approach,[19] involving initial formation of 1,2-*trans*-glycosides followed by inversion of the C-2 configuration of the glycosyl residue, was considered. In the current context, this requires the availability of the C-2 epimer of aceric acid. The 3-keto-

SCHEME 4



furanose derivative **16**, prepared by known methods[10] from L-arabinofuranose, served as a convenient starting material for introduction of the C-3 branch point (Scheme 5). Application of 2-thiazolyllithium as a nucleophile afforded a single isomer, **17**, with the required C-3 stereochemistry in 70% yield. The same *exo*-addition product, **17**, was identified as the major component of a 2:1 mixture of stereoisomers obtained in a combined 48% yield in the reaction of **16** with 2-TST.[20] After protecting the 3-OH group in **17** via benzylation, 3-*C*-thiazolyl derivative **18** was transformed into aldehyde **19** in 80% overall yield, using the three-step procedure described for the synthesis of L-*xylo*-isomer **11**. Sodium chlorite oxidation of **19** afforded 3-*C*-carboxy derivative **20**, which was deprotected using catalytic hydrogenolysis (compound **21**) and subsequent acid hydrolysis leading to the target C-2 epimer of aceric acid, **22**, in 74% overall yield from **19**. Similar to the case of aceric acid **1**, NMR spectra of **22** in D_2O showed two sets of signals that indicated the presence of a mixture of anomers in which the isomer with 1,2-*trans*-configuration (α-anomer) dominates (α-**22**/β-**22** = 2.6:1). Assignment of the anomeric configuration of α-**22** and β-**22** was based on anomeric ^{13}C chemical shifts, which showed a difference of 4.2 ppm, consistent with literature data.[21]

The structure of aldehyde **19** was confirmed by conversion into known 1,2-*O*-isopropylidene derivative **23**,[22]

(16) (a) Altman, L. J.; Richheim, S. L. *Tetrahedron Lett.* **1971**, *12*, 4709−4711. (b) Dondoni, A.; Fogagnolo, M.; Medici, A.; Pedrini, P. *Tetrahedron Lett.* **1985**, *26*, 5477−5480.

(17) de Nooy, A. E. J.; Besemer, A. C.; van Bekkum, H. *Synthesis* **1996**, 1153−1174.

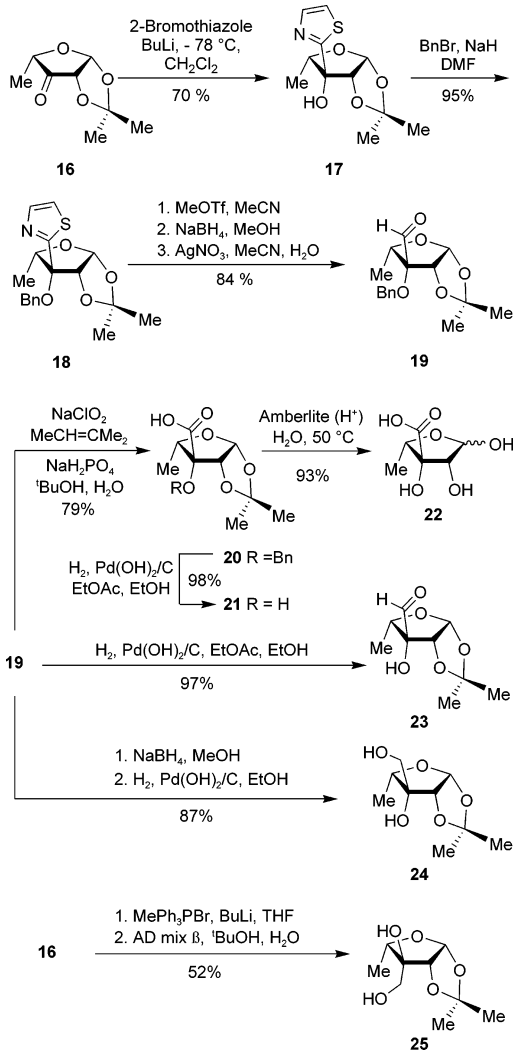
(18) Gadikota, R. R.; Callam, C. S.; Wagner, T.; Del Fraino, B.; Lowary, T. L. *J. Am. Chem. Soc.* **2003**, *125*, 4155−4165.

(19) (a) David, S.; Malleron, A.; Dini, C. *Carbohydr. Res.* **1989**, *188*, 193−200. (b) Weiler, S.; Schmidt, R. R. *Tetrahedron Lett.* **1998**, *39*, 2299−2302.

(20) The presence of the *endo*-isomer of **17** as the minor component in the mixture of reaction products was shown by NMR spectroscopy and mass spectrometry (see Supporting Information). 2-TST addition to ketone **16** appears to give a "mismatched" reaction, with the *exo*-addition product formed in modest excess. In contrast, the corresponding addition to diastereomeric ketone **2** appears to be "matched" with respect to generation of the desired *endo*-adduct required for aceric acid synthesis.

(21) (a) Angyal, S. J.; Bethell, G. S. *Aust. J. Chem.* **1976**, *29*, 1249−1265. (b) Bock, K.; Pedersen, C. *Adv. Carbohydr. Chem. Biochem.* **1983**, *41*, 27−66.

SCHEME 5



which represents a protected form of the well-known branched-chain sugar streptose,[10] a key component of the antibiotic streptomycin. $NaBH_4$ reduction and subsequent catalytic hydrogenolysis of aldehyde **19** gave diol **24**, a diastereomer of diols **6** and **15**. The fourth diastereomer of this series of branched-chain sugar diols, namely **25**, was synthesized starting from 3-ketofuranose **16**. The latter was first converted into its 3-*C*-methylene derivative and then dihydroxylated, using the same conditions as applied for transformation of 3-ketofuranose **2**, giving diol **25**.

In addition to crystallographic data for 3-*C*-thiazolyl derivatives **7** and **10**, assignment of the C-3 configuration in the synthetic branched-chain sugars was accomplished using selective 1D NOE experiments. Investigation of a systematic series of diastereomeric 1,2-*O*-isopropylidene derivatives **6**, **15**, **24**, and **25** revealed NOEs indicative of the relative orientation of the hydroxymethyl group and the C-5 methyl group on the furanose ring. A 1,2-*trans*-orientation of these groups was established in compounds **6** and **24** and a 1,2-*cis*-orientation in compounds **15** and **25**.[23]

In summary, we successfully synthesized aceric acid **1** from a readily available 1,2-*O*-isopropylidene derivative of L-xylofuranose **2** using two different approaches. The first approach was based on efficient, diastereoselective 2-trimethylsilylthiazole addition to 3-ulose **2** under thermodynamic control, followed by simple unmasking of the 3-*C*-formyl group and its oxidation to the 3-*C*-carboxy group. In the second approach, dihydroxylation of a 3-*C*-methylene derivative of L-xylofuranose **3** was applied, followed by TEMPO−NaOCl oxidation of the primary OH group. The overall yield starting from L-xylose was 13% (over 10 steps) and 14% (over eight steps) in the first and the second routes, respectively. In addition, the 2-epimer of aceric acid, **22**, was also synthesized starting from L-arabinofuranose via derivative **16** (13 steps, 6% yield).

Experimental Section

5-Deoxy-1,2-*O*-isopropylidene-3-*C*-(thiazol-2-yl)-α-L-xylofuranose (7) and 5-Deoxy-1,2-*O*-isopropylidene-3-*C*-(thiazol-2-yl)-α-L-ribofuranose (8). 2-(Trimethylsilyl)thiazole (1.46 mL, 9.13 mmol) was added to a stirred solution of compound **2** (0.79 g, 4.57 mmol) in THF (10 mL), and the resulting mixture was refluxed for 36 h. On cooling and concentration *in vacuo*, the residue was dissolved in THF (10 mL) and treated with a 1 M solution of TBAF in THF (4.6 mL, 4.6 mmol). After being stirred for 30 min at room temperature, the solution was concentrated *in vacuo* to afford a mixture of isomeric thiazoles **7** and **8** in a 5:1 ratio (from integration of anomeric signals in ^{1}H NMR spectra). Individual products were isolated by silica gel chromatography (hexane/EtOAc 8:1 → 4:1) to give compound **7** as a crystalline solid (0.68 g, 58%) and compound **8** as an oil (0.14 g, 12%). Compound **7**: R_f = 0.20 (hexane/EtOAc 1:1); mp 86−88 °C (EtOAc/hexane); $[\alpha]_D$ −83 (*c* 1.0, $CHCl_3$); IR (thin film): ν_{max} 3100 (OH) cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$): δ 1.16 (d, $J_{4,5}$ = 6.4 Hz, 3 H, H-5), 1.36 (s, 3 H, Me), 1.64 (d, 3 H, Me), 4.46 (q, $J_{4,5}$ = 6.4 Hz, 1 H, H-4), 4.51 (d, $J_{1,2}$ = 3.7 Hz, 1 H, H-2), 6.07 (d, $J_{1,2}$ = 3.7 Hz, 1 H, H-1), 7.45 (d, $J_{A,B}$ = 3.2 Hz, 1 H, thiazole), 7.80 (d, $J_{A,B}$ = 3.2 Hz, 1 H, thiazole); ^{13}C NMR (100 MHz, $CDCl_3$): δ 11.6 (C-5), 26.8 (Me), 27.0 (Me), 81.2 (C-4), 82.8 (C-3), 86.2 (C-2), 105.3 (C-1), 113.0 (*C*Me$_2$), 121.9, 141.2, 166.5 (thiazole); HR ESI-MS: found *m/z* 258.0795 [M + H]$^+$, calcd for $C_{11}H_{16}NO_4S$ 258.0795. Compound **8**: R_f = 0.30 (hexane/EtOAc 1:1); $[\alpha]_D$ −88 (*c* 1.0, $CHCl_3$); IR (thin film) ν_{max} 3109 (OH) cm^{-1}; ^{1}H NMR (400 MHz, $CDCl_3$): δ 0.99 (d, $J_{4,5}$ = 6.4 Hz, 3 H, H-5), 1.42 (s, 3 H, Me), 1.66 (s, 3 H, Me), 3.40 (broad s, 1 H, OH), 4.27 (q, $J_{4,5}$ = 6.4 Hz, 1 H, H-4), 4.65 (d, $J_{1,2}$ = 3.9 Hz, 1 H, H-2), 6.16 (d, $J_{1,2}$ = 3.9 Hz, 1 H, H-1), 7.36 (d, $J_{A,B}$ = 3.3 Hz, 1 H, thiazole), 7.82 (d, $J_{A,B}$ = 3.3 Hz, 1 H, thiazole); ^{13}C NMR (100 MHz, $CDCl_3$): δ 13.6 (C-5), 26.8 (Me), 26.9 (Me), 79.3, 82.7, 83.3 (C-2, C-3, C4), 105.0 (C-1), 113.2 (*C*Me$_2$), 119.3, 143.1, 169.9 (thiazole); HR ESI-MS: found *m/z* 258.0792 [M + H]$^+$, calcd for $C_{11}H_{16}NO_4S$ 258.0795.

Acknowledgment. This research was supported by the U.K. Engineering and Physical Sciences Research Council and the Weston Foundation. We are indebted to the EPSRC Mass Spectrometry Service Centre, Swansea for invaluable support.

Supporting Information Available: Experimental procedures, characterization data and NMR spectra for compounds **1**, **2**, **6**−**15**, **17**−**25**, X-ray data for compounds **7** and **10**, 1D NOE spectra for compounds **6**, **15**, **24**, and **25**. This material is available free of charge via the Internet at http://pubs.acs.org.

JO051012B

(22) Paulsen, H.; Sinnwell, V.; Stadler, P. *Chem. Ber.* **1972**, *105*, 1978−1987.

(23) NOEs for compounds **6**, **15**, **24**, and **25** were in the 2−3% range. Details of the observed NOEs, which were obtained for these compounds upon selective irradiation of the signals corresponding to H-1, H-2, H-3, H-4, and H-3′, are summarized in Supporting Information.