Organic & Biomolecular Chemistry

Dynamic Article Links

Cite this: *Org. Biomol. Chem.*, 2011, **9**, 8371

www.rsc.org/obc

PAPER

Exploring the synthetic potency of the first furanothioglycoligase through original remote activation†

Mélanie Almendros,[a,b] Dantcho Danalev,[a,b] Marc François-Heude,[a,b] Pascal Loyer,[c] Laurent Legentil,[a,b] Caroline Nugier-Chauvin,[a,b] Richard Daniellou*[a,b] and Vincent Ferrières*[a,b]

Received 22nd July 2011, Accepted 20th September 2011
DOI: 10.1039/c1ob06227a

Thioglycosidic bonds are of utmost importance in biomolecules as their incorporation led to more stable glycomimetics with potential drug activities. Until now only chemical methods were available for their incorporation into glycofuranosyl conjugates. Herein, we wish to describe the use of the first furanothioglycoligase for the preparation of a great variety of thioaryl derivatives with moderate to excellent yields. Of great interest, a stable 1-thioimidoyl arabinofuranose, classically used in chemical glycosylation, was able to efficiently act as a donor through an original enzymatic remote activation mechanism. Study of the chemical structure as well as the nucleophilicity of the thiol allowed us to optimize this biocatalyzed process. As a consequence, this mutated enzyme constitutes an original, mild and eco-friendly method of thioligation.

Introduction

Carbohydrates play an important part in a vast array of biological processes and glycomimetics are currently becoming a powerful class of novel therapeutics.[1] Amongst them, thioglycosides, in which a sulfur atom has replaced the glycosidic oxygen atom, are tolerated by most biological systems and are less sensitive to acid/base or enzyme-mediated hydrolysis. Such compounds have already been demonstrated to be valuable tools as good chemical donors for synthetic purpose,[2,3] as stable intermediates in X-ray crystallographic analysis of proteins[4] and, of particular interest, as competitive inhibitors of a wide range of glycosidases involved in numerous diseases.[5] Their methods of preparation have recently been reviewed.[6] One promising strategy relies on the use of eco-friendly and versatile enzymes. Thioglycoligases,[7] mainly developed by the team of Withers, represent mutant retaining glycosidases with low hydrolytic activities that are able to efficiently perform the synthesis of *S*-linked oligosaccharides and thioglycoproteins.[8–12] Since they are lacking the catalytic acid/base amino acid residue (Fig. 1), such enzymes generally required particular glycosyl donors owning good leaving groups like dinitrophenyl derivatives, for example, and deoxythio sugars as better nucleophilic acceptors. However, up to now, only examples related to pyranosyl hydrolases have been described in the literature, although the corresponding mutated furanosidases have been easily obtained.[13,14] In that case the scientific key seems to rely on the chemical synthesis of both furanosyl substrates. Still, glycofuranosyl-conjugates are widespread in nature and, notably, are present in numerous pathogens responsible for global threats such as tuberculosis, leishmaniosis, leprae, etc.[15,16] In addition, they are completely absent from mammals, except in the case of the D-ribose and 2-deoxy-D-ribose in nucleic acids, and thus can constitute original molecular structures for the development of drugs or immunostimulating additives in vaccines.[17–20]

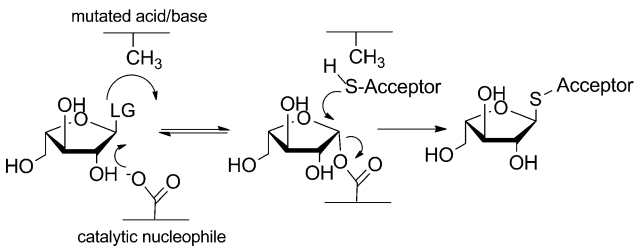


Fig. 1 General mechanism of the furanothioglycoligase.

As a consequence, we were interested in demonstrating the thioglycoligase potency of the E173A mutant of the α-L-arabinofuranosidase Ara*f*51 from *Clostridium thermocellum*[14] for the preparation of thiodisaccharides (Fig. 1). First, the prerequisite was the chemical synthesis of both a stable furanosyl donor owning a better leaving group and various thiol acceptors. Secondly, enzymatic reactions were performed and helped us in gaining insights into the substrate specificity and the mechanism of the biocatalyzed process. Finally, application of this enzymatic procedure allowed us to synthesize various *S*-arabinofuranosides, thus revealing an efficient and general method of ligation of sugars onto thiophenol derivatives.

[a]*Ecole Nationale Supérieure de Chimie de Rennes, CNRS, UMR 6226, Avenue du Général Leclerc, CS 50837, 35708, Rennes Cedex 7, France. E-mail: richard.daniellou@ensc-rennes.fr; Fax: +33-2-23-23-80-46; Tel: +33-2-23-23-80-96*
[b]*Université Européenne de Bretagne, France*
[c]*Inserm UMR S-991, Liver Metabolisms and Cancer, Université de Rennes 1, 35033, Rennes, France*

† Electronic supplementary information (ESI) available: Spectra of compounds **7**, **10**, **11**, **12**, **13**, **14** and **15**. See DOI: 10.1039/c1ob06227a

Results and discussion

The arabinofuranosidase Araf51 is a thermostable glycosidase able to catalyze the release of L-arabinose from a range of plant polysaccharides. Its substrate specificity as well as its structural determination by X-ray analysis have already been published.[14] Recently, this native enzyme has also been demonstrated to be very efficient in promoting auto-condensation reactions, thus leading to a great variety of oligofuranosides.[18] The E173A mutant of this biocatalyst was overexpressed in our laboratory using common procedures previously described[14] and easily purified by heating the mixture for one minute at 70 °C to precipitate out all other proteins. This procedure afforded us tens of milligrams of protein. Then, we focused our attention on the chemical synthesis of a convenient glycosyl donor as well as numerous acceptors.

Chemical synthesis of donors and acceptors

p-Nitrophenyl glycosides like **1** are well known as synthetic substrates of glycosidases and, as a consequence, most of them are commercially available (Fig. 2). On the contrary, thioglycoligases require glycosyl donors with good leaving groups that do not need acid catalysis, for example, dinitrophenyl groups, to allow formation of the glycosyl-enzyme intermediate (Fig. 1). Unfortunately, in the case of furanosides, the preparation of dinitrophenyl- or chloronitrophenyl L-arabinofuranosides **2** involves tedious reactions and leads to unstable water-sensitive compounds.[13,21] Moreover, such chemical procedures were hardly reproducible and allowed us to produce these compounds only in very low yields. In order to overcome this problem, unprotected 1-thioimidoyl arabinofuranoses like **3** were thought to represent good donors since i) they proved to be very stable in water (no degradation was observed by NMR in 2 days), and ii) upon remote activation, their aglycone constitutes good leaving groups which enhanced their reactivity in the glycosylation process.[22–24]

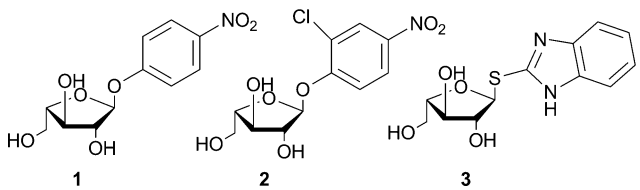


Fig. 2 Substrates of the wild-type arabinofuranosidase **1** and the thioglycoligase mutant **2**–**3**.

On this basis, we thus envisioned that **3** should be well-recognized by Araf51 and might act as an efficient glycosyl donor for enzymatic reactions, eventually through remote activation of the aglycone thanks to an acidic residue close to the active site. The biocatalyzed hydrolysis of **3** in the presence of the wild-type Araf51 was followed by ^{1}H NMR spectroscopy in D_2O (data not shown). The spectra rapidly exhibited the appearance of signals corresponding to those of the anomeric protons of L-arabinose, thus confirming the substrate potency of the 1-thioimidoyl arabinofuranose. This was unambiguously confirmed using a competitive assay in the presence of **1** (1 mM) at pH = 7.0, a concentration equal to 0.15 mM of **3** being necessary to prevent 50% of the hydrolysis of the former. Docking experiments were further performed in order to assess the role of the aglycone. Interestingly, **3** displays a similar conformation as the arabinofuranosyl moiety of the arabinotrioside present in the –1 subsite in the crystal structure previously reported (Fig. 3).[14] As expected, the aglycone undergoes hydrophobic stacking with Trp-178 in the +1 subsite. More interestingly, this thioimidoyl leaving group is axial to the nucleophile Glu-292 that is poised for in-line nucleophilic attack (2.9 Å). Similarly, the hydrogen of the OH-2 of the arabinofuranosyl moiety interacts with the nucleophile O2 atom of Glu-292 (1.8 Å). In addition, the NH group of the benzimidazoyl aglycone also interacts with the other nucleophile O1 atom of the same glutamate residue (2.2 Å). Consequently, enzymatic remote activation of the leaving group seems to occur directly with Glu-292 present in the active site, *i.e.* the nucleophile residue commonly responsible for the formation of the glycosyl–enzyme intermediate. To our knowledge, such activation constitutes the first report of an enzymatic remote activation of a stable donor in glycosidases. More importantly, as all the members of this class of enzymes working with retention of configuration share a common mechanism,‡ such results pave the way for the further development of thioimidoyl furanoses and pyranoses, classically used in chemical glycosylation, in biocatalyzed reactions.

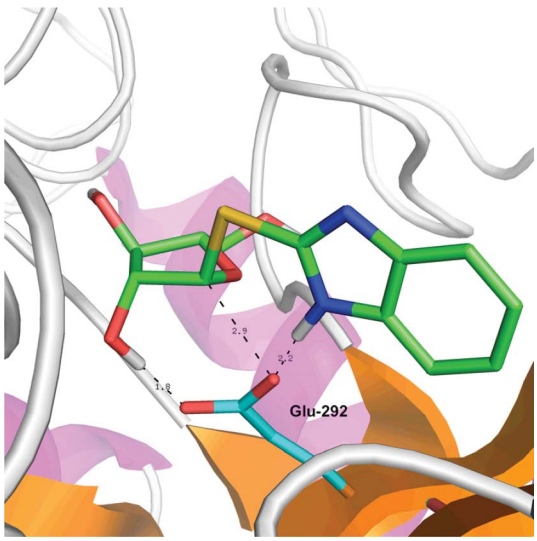
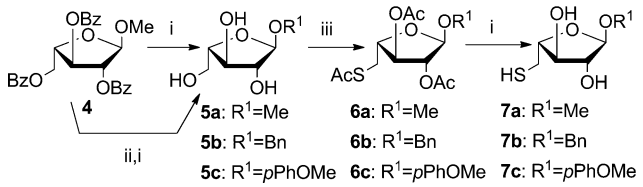


Fig. 3 Docking of **3** in the active site of the acid/base mutant of Araf51.

With this original potential enzymatic donor in hand, we turned our attention to the chemical synthesis of the 5-deoxy-5-thio-α-L-arabinofuranosides **7** (Scheme 1). Two kinds of derivatives were considered: the methyl **7a** for its ease of preparation and the benzyl **7b** or the *p*-methoxyphenyl **7c** since such groups are known to be good mimics of sugar entities and they can be easily visualized using standard analytical methods. Starting from readily accessible protected methyl arabinofuranoside **4**,[25] a simple Zemplen deprotection gave us **5a** in quantitative yield. Alternatively, the sugar **4** was also activated by acetolysis and the desired aromatic groups were introduced using common glycosylation reactions promoted by a borontrifluoride–etherate complex and the corresponding alcohols. Zemplen deprotection afforded us **5b–c** with 63% and 71% overall yield, respectively. Then, Mitsunobu conditions were used in order to substitute the

‡ See the Carbohydrate Active Enzyme Database at www.cazy.org

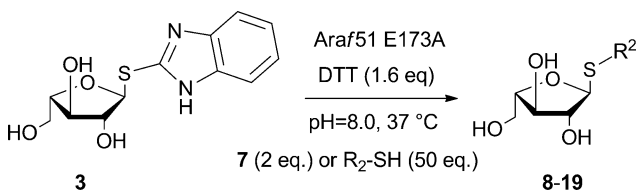


Scheme 1 Chemical synthesis of 5-deoxy-5-thio-α-L-arabinofuranosides: i) MeoNa (0.01 M), MeOH, rt: **5a** (100%), **7a** (100%), **7b** (94%), **7c** (98%); ii) a) Ac_2O, cat. H_2SO_4, CH_2Cl_2, rt, b) BF_3–OEt_2 (3.1 eq.), Et_3N (0.08 eq.), alcohol (1.2 eq.), CH_2Cl_2, rt, 24 h: **5b** (71%), **5c** (63%); iii) a) thioacetic acid (1.8 eq.), PPh_3 (1.7 eq.), DIAD (1.7 eq.), THF, rt, overnight, b) Ac_2O/pyridine (1/1), rt, 3 h: **6a** (69%), **6b** (74%), **6c** (74%).

primary alcohol with a thioacetyl function. After concentration and filtration on a silica pad, the resulting crude compounds were directly peracetylated to facilitate their purification. Compounds **6a–c**, obtained with 69 to 74% yields, were stored at –20 °C and finally deprotected using sodium methoxide to give **7** just before enzymatic reaction in order to prevent the formation of their corresponding disulfide. The absence of the latter was easily verified by checking the chemical shift of the ^{13}C NMR signal corresponding to C-5, around 27 ppm for the thiol and 43.5 ppm for its disulfide counterpart.

Biocatalyzed-thioligation

The thioimidoyl donor **3** was then allowed to react with the various 5-deoxy-5-thio-arabinofuranosides **7a–c** (2 eq.) in the presence of the thioglycoligase (Scheme 2) and dithiothreitol (1.6 eq., DTT). Experiments were performed in 50 mM phosphate buffer at pH = 8.0 and 37 °C. Reactions were monitored by thin layer chromatography until complete consumption of the donor (24 h). The resulting compounds were purified by flash column chromatography and subsequently analyzed by NMR spectroscopy.



Scheme 2 Biocatalyzed synthesis of thiofuranosides **8–19**.

Unexpectedly, in all three cases, we were not able to detect the formation of the desired thioglycoside **8**. In fact, **3** was totally converted into its hydrolysis product (L-arabinopyranose) and the thio-sugars **7** were partially oxidized as disulfides. These observations led us to worry about i) the recognition of the carbohydrates **7** by the active-site of the mutated enzyme and ii) more importantly, their reactivity as nucleophiles towards the ligation reaction. Since **7** represent slightly modified substrates compared to the natural ones, we envisioned the second point to be the major problem. Taking into account results obtained in the native chemical ligation reaction of peptides,[26] our hypothesis was that aryl thiols constitute better nucleophiles than carbohydrates **7** and might act as good acceptors in this biocatalyzed ligation reaction.

Therefore, a set of commercially available thiols§ (50 eq.), both alkyl and aryl, were chosen based on the pK_a of the thiol moiety (Table 1). As envisioned, octane thiol, owning the highest pK_a, was totally unreactive in our conditions (entry 2). Interestingly, all aryl thiols were able to act as substrates and led to compounds **11–18** with moderate to excellent yields (50–98%, entries 4–12). The only exception was the benzyl mercaptan (entry 3), which, unsurprisingly with respect to its pK_a of 9.43, led to only 5% yield of the thioglycoside **10**. However, the ligation reaction seems to be general for aryl thiols of pK_a between 8.75 to 5.83. In our case, the best result was obtained with the thiophenol as emphasized by the 98% yield for derivative **15** (entry 8). Substituents like chlorine or a methyl group in the *ortho*, *meta* and *para* position did not prevent the function of the biocatalyst (entries 5–6 and 10–11 for examples). In addition, the enzyme was able to tolerate many other chemical functions (OH, NH_2, CO_2H) and was shown to be fully selective for the creation of the thioglycosidic linkage (entries 4, 6 and 12). It should also be noted that DTT (pK_a of 9.2 and 10.1) doesn't seem to participate in the reaction since we never observed its incorporation in the product. Consequently the E173A mutant of the Ara*f*51 constitutes an original and powerful method to create a new thioglycosidic bond using mild and eco-compatible conditions.

STD-NMR

In addition, no inhibition was observed when performing a competitive assay with both benzyl and phenyl mercaptan. Compound **10**, even with a ratio up to 1 : 10 (donor **3** : **10**), was also not able to inhibit the enzymatic reaction thus ruling out a possible product inhibition. Therefore, one question was still present in this study: is the benzyl mercaptan simply less reactive than the thiophenol or does the Ara*f*51 E173A recognize it less? Saturation Transfer Difference (STD) NMR experiments were performed in order to solve this problem.[27] This technique allowed us to determine rapidly the binding epitopes of the thioglycosides **10** and **15** (Fig. 4). Normally, an STD value of 100% is arbitrarily attributed to the highest peak. Here, as it corresponded to the peak of the 5 protons of the aromatic, a value of 500% was given. As depicted in Fig. 4, the *S*-phenyl arabinofuranoside **15** revealed to be tightly bound to the enzyme. Moreover, the aglycon seemed to be as well recognized by the protein as the sugar moiety, emphasizing once again the importance of the hydrophobic stacking with Trp-178 in the +1 subsite. Similar values were obtained for the *S*-benzyl arabinofuranoside **10**. This simple observation led us to the conclusion that the recognition does not represent a major problem but more surely that the benzyl mercaptan is less reactive than the thiophenol during this biocatalyzed procedure.

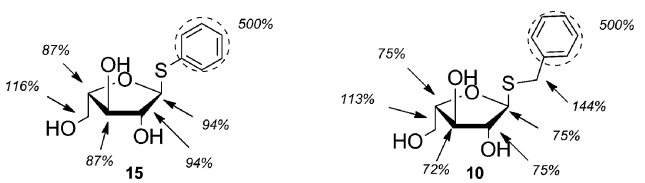
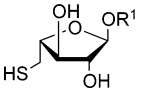
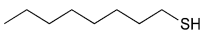
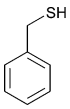
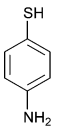
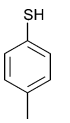
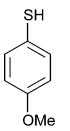
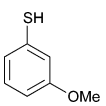
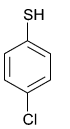
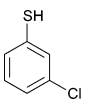
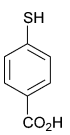


Fig. 4 *S*-Bn and *S*-ph arabinofuranosides **10** and **15**. STD effects observed by NMR are indicated in italics for each proton.

§ First reactions were performed with 2 eq. but yields were limited to 30%.

Table 1 Thiols tested in the thioligase reaction

Entry	R_2-SH	pK_a	Product (yield%)
1		10.60[a]	**8**(—)
2		10.00	**9**(—)
3		9.43	**10**(5)
4		8.75	**11**(60)
5		6.81	**12**(50)
6		6.79	**13**(52)
7		6.75	**14**(81)
8		6.61	**15**(98)
9		6.36	**16**(67)
10		6.11	**17**(77)
11		5.83	**18**(73)
12		5.83	**19**(84)

[a] Estimated value by comparison with methyl and ethyl mercaptan.

Conclusions

Thioglycosidic bonds are of utmost importance in biomolecules as their incorporation leads to more stable glycomimetics with potential drug activities. In addition to well-known chemical procedures, we have developed an original method using the first furanothioglycoligase. As a prerequisite, both donors and acceptors were prepared using standard chemical reactions with good overall yields. Interestingly, thioimidoyl glycosides like **3** were shown to represent a promising new class of stable substrate that can undergo remote activation inside glycosidases. As they are easily available, their further use even in the pyranose form will with undoubtedly help us in increasing the synthetic potency of new thioglycoligases. Moreover, even if unreactive in our case, three different 5-deoxy-5-thio furanosides are now available for such studies. Even more importantly, insights were also gained on the detailed mechanism of this particular mutated enzyme. In total accordance with what was previously described, Araf51 constitutes a tremendous synthetic tool due to its great versatility toward the acceptor recognition subsite *i.e.* furanosides, pyranosides, aryl, thioaryl, etc. In addition, the nucleophilicity of the thiols seems to represent the major parameter that influences the yield of the biocatalyzed reaction. Experiments are currently in progress to extend this enzymatic thioligation to more challenging biological objects.

Experimental

General procedures

Thin layer chromatography (TLC) analyses were conducted on E. Merck 60 F_{254} Silica Gel non activated plates and compounds were visualized by UV-absorption at 254 nm, by dipping in a 5% solution of orcinol in 5% solution of H_2SO_4 in EtOH followed by heating. Preparative chromatography purifications were performed on Geduran Si 60 (40–63 µm) Silica Gel. Optical rotations were measured on a Perkin–Elmer 341 polarimeter. NMR spectra were recorded with a Brüker ARX 400 spectrometer at 400 MHz for ^{1}H and 100 MHz for ^{13}C. Chemical shifts are given in δ-units (ppm) measured from the solvent signal. Coupling constants J were calculated in hertz (Hz). Abbreviations were used to present signal multiplicity: s (singlet), d (doublet), t (triplet), m (multiplet), dd (doublet doublet). Mass spectra were measured with a MS/MS ZabSpec TOF Micromass using *m*-nitrobenzylic alcohol as a matrix and accelerated caesium ions for ionization.

Docking experiment

Rotatable bonds in the ligand were assigned with Autodock Tools – an accessory program that allows the user to interact with Autodock. Docking simulations were performed with AutoDock 4.2 using a Lamarckian genetic algorithm. The standard docking procedure was used for a rigid protein and a flexible ligand whose torsion angles were identified (for 100 independent runs per ligand). A grid of 60, 60 and 60 points in x, y, and z directions was built, centered on the nucleophile residue Glu-292. A grid spacing of 0.375 Å and a distance-dependent function of the dielectric constant were used for the calculation of the energetic map. The default settings were used for all parameters. At the end of docking, the ligand with the most favorable free energy of binding was selected as the resultant complex structure. All calculations were carried out on Apple based machines using Mac Os X.6. The

resultant structure file was analyzed using Pymol 1.3 visualization programs.

General procedure for the Zemplen deprotection

The peracetylated carbohydrate (0.15 mmol) was dissolved in MeOH (3 mL) and 0.1 M MeONa in MeOH (2.7 eq) was added. The mixture was stirred at room temperature for 15 min. After neutralization with amberlite IR-120 (H^+ form), the resin was filtered off and the MeOH removed under pressure. The residue was finally purified by column chromatography on silica gel (9 : 1 CH_2Cl_2/MeOH) to afford **5a–c** or **7a–c**.

Benzyl α-L-arabinofuranoside (5b)

To a solution of 1-*O*-acetyl-2,3,5-tri-*O*-benzoyl-α-L-arabinofuranoside (4.0 g, 7.82 mmol) in anhydrous dichloromethane (40 mL) were successively added $BF_3{\cdot}OEt_2$ (2.9 mL, 24.5 mmol) and Et_3N (0.54 mL, 0.62 mmol) at 0 °C. After stirring for 10 min, benzyl alcohol (9.28 mmol) was added. The reaction mixture was stirred for 24 h at room temperature, then diluted with water (80 mL) and extracted with dichloromethane (150 mL). The organic layer was washed with 1 N aqueous HCl (3 × 40 mL) followed by a saturated solution of aqueous $NaHCO_3$ (3 × 40 mL). After drying over $MgSO_4$ and concentration, the residue was deprotected using the general Zemplen procedure to afford **5b** as a colourless oil (1.3 g, 71%): $[\alpha]_D^{20}$ −12.3 (*c* 1.00, MeOH); R_f 0.33 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.38–7.25 (m, 5H, H arom), 4.94 (d, 1H, $J_{1,2}$ = 1.5 Hz, H-1), 4.75 (d, 1H, *J* = 11.9 Hz, CH_2 benzyl), 4.53 (d, 1H, CH_2 benzyl), 4.03 (dd, 1H, $J_{2,3}$ = 3.8 Hz, H-2), 3.99 (ddd, 1H, $J_{3,4}$ = 6.4 Hz, $J_{4,5a}$ = 3.1 Hz, $J_{4,5b}$ = 5.5 Hz, H-4), 3.86 (dd, 1H, H-3), 3.78 (dd, 1H, $J_{5a,5b}$ = 12.0 Hz, H-5a), 3.66 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 139.3 (C-1′), 129.3 (C-2′, C-6′), 129.0 (C-4′), 128.6 (C-3′, C-5′), 108.5 (C-1), 85.5 (C-4), 83.7 (C-2), 78.9 (C-3), 70.0 (*C*H_2 benzyl), 63.0 (C-5); HRESIMS: Calcd for $C_{12}H_{16}NaO_5$ [M + Na]$^+$: *m*/*z* 263.0667; found: *m*/*z* 263.0666.

4′-Methoxyphenyl α-L-arabinofuranoside (5c)

Compound **5c** was synthesized as previously described for **5b** using 4-methoxyphenol (9.28 mmol). White solid (1.3 g, 63%): $[\alpha]_D^{20}$ −131.4 (*c* 1.00, MeOH); R_f 0.33 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 6.97 (d, 2H, *J* = 9.2 Hz, H-2′, H-6′), 6.80 (d, 2H, H-3′, H-5′), 5.38 (d, 1H, $J_{1,2}$ = 2.0 Hz, H-1), 4.19 (dd, 1H, $J_{2,3}$ = 4.2 Hz, H-2), 4.04 (ddd, 1H, $J_{3,4}$ = 6.5 Hz, $J_{4,5a}$ = 3.1 Hz, $J_{4,5b}$ = 5.1 Hz, H-4), 3.94 (dd, 1H, H-3), 3.76 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, H-5a), 3.73 (s, 3H, OCH_3), 3.64 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 156.4 (C-4′), 152.4 (C-1′), 119.2 (C-2′, C-6′), 115.4 (C-3′, C-5′), 108.7 (C-1), 85.9 (C-4), 83.7 (C-2), 78.2 (C-3), 62.8 (C-5), 56.0 (O*C*H_3); HRESIMS: Calcd for $C_{12}H_{16}NaO_6$ [M + Na]$^+$: *m*/*z* 279.0844; found: *m*/*z* 279.0848.

General procedure for the Mitsunobu reaction and the following peracetylation

Thioacetic acid was purified prior to its use by low-temperature distillation under nitrogen. To a cooled solution of thioacetic acid (1.8 eq.) and **5** (1 eq.) in dry THF (20 mL) was added a cooled solution of triphenylphosphine (1.7 eq.) and diisopropyl azodicarboxylate (1.7 eq.) in dry THF (20 mL). The reaction was allowed to warm up to room temperature and stirred overnight under nitrogen. After concentration *in vacuo*, the residue was suspended in EtOAc (50 mL) before filtration through a silica pad. The resulting filtrate was concentrated and dissolved in dry pyridine (5 mL). Acetic anhydride (5 mL) was added and the mixture was stirred at room temperature for 3 h. After concentration *in vacuo* and co-distillation with toluene (3 × 15 mL), the resulting oil was diluted with EtOAc (100 mL) and successively washed with 5% aqueous HCl (3 × 20 mL), saturated solution of aqueous $NaHCO_3$ (3 × 20 mL) and water (3 × 20 mL). After drying over $MgSO_4$ and concentration *in vacuo*, the residue was purified by column chromatography on silica gel (95 : 5 toluene/EtOac) to afford **6**.

Methyl 2,3-di-*O*-acetyl-5-deoxy-5-thioacetyl-α-L-arabinofuranoside (6a)

Colourless oil (0.78 g, 69%): $[\alpha]_D^{20}$ −65.5 (*c* 0.95, $CHCl_3$); R_f 0.33 (9 : 1 toluene/EtOac);^{1}H NMR ($CDCl_3$): 5.01 (dd, 1H, $J_{2,3}$ = 2.0 Hz, H-2), 4.89 (dd, 1H, $J_{3,4}$ = 5.7 Hz, H-3), 4.87 (d, 1H, H-1), 4.17 (ddd, 1H, $J_{4,5a}$ = 4.9 Hz, $J_{4,5b}$ = 6.9 Hz, H-4), 3.77 (s, 3H, OCH_3), 3.38 (dd, 1H, $J_{5a,5b}$ = 13.9 Hz, $J_{4,5a}$ = 2.6 Hz, H-5a), 3.18 (dd, 1H, H-5b), 2.36 (s, 3H, SCOCH_3), 2.11 (s, 3H, OCOCH_3), 2.10 (s, 3H, OCOCH_3); ^{13}C NMR ($CDCl_3$): δ 194.8 (S*C*OCH$_3$), 170.2 (O*C*OCH$_3$), 169.9 (O*C*OCH$_3$), 106.7 (C-1), 81.8 (C-2), 80.5 (C-4), 79.3 (C-3), 55.1 (O*C*H$_3$), 31.3 (C-5), 30.6 (SCO*C*H$_3$), 20.9 (2x OCO*C*H$_3$); HRESIMS: Calcd for $C_{12}H_{18}NaO_7S$ [M + Na]$^+$: *m*/*z* 329.0670; found: *m*/*z* 329.0668.

Benzyl 2,3-di-*O*-acetyl-5-deoxy-5-thioacetyl-α-L-arabinofuranoside (6b)

Colourless oil (1.1 g, 74%): $[\alpha]_D^{20}$ −55.3 (*c* 0.88, $CHCl_3$); R_f 0.45 (95 : 5 toluene/EtOac);^{1}H NMR ($CDCl_3$): 7.28–7.18 (m, 5H, H arom), 5.04 (dd, 1H, $J_{2,3}$ = 2.0 Hz, H-2), 4.98 (d, 1H, H-1), 4.84 (dd, 1H, $J_{3,4}$ = 5.6 Hz, H-3), 4.67 (d, 1H, *J* = 12.1 Hz, CH_2 benzyl), 4.46 (d, 1H, CH_2 benzyl), 4.14 (ddd, 1H, $J_{4,5a}$ = 4.9 Hz, $J_{4,5b}$ = 6.9 Hz, H-4), 3.31 (dd, 1H, $J_{5a,5b}$ = 13.8 Hz, H-5a), 3.10 (dd, 1H, H-5b), 2.28 (s, 3H, SCOCH_3), 2.02 (s, 3H, OCOCH_3), 2.01 (s, 3H, OCOCH_3); ^{13}C NMR ($CDCl_3$): δ 195.9 (S*C*OCH$_3$), 170.2 (O*C*OCH$_3$), 169.8 (O*C*OCH$_3$), 137.3 (C-1′), 128.5 (C-2′, C-6′), 128.4 (C-4′), 127.9 (C-3′, C-5′), 104.8 (C-1), 81.9 (C-2), 80.7 (C-4), 79.3 (C-3), 69.0 (*C*H_2 benzyl), 31.3 (C-5), 30.6 (SCO*C*H$_3$), 20.9 (2x OCO*C*H$_3$); HRESIMS: Calcd for $C_{18}H_{22}NaO_7S$ [M + Na]$^+$: *m*/*z* 405.0983; found: *m*/*z* 405.0983.

4′-Methoxyphenyl 2,3-di-*O*-acetyl-5-deoxy-5-thioacetyl-α-L-arabinofuranoside (6c)

Colourless oil (1.1 g, 74%): $[\alpha]_D^{20}$ −94.1 (*c* 1.00, $CHCl_3$); R_f 0.44 (95 : 5 toluene/EtOac);^{1}H NMR ($CDCl_3$): 6.97 (d, 2H, *J* = 9.2 Hz, H-2′, H-6′), 6.82 (d, 2H, H-3′, H-5′), 5.51 (d, 1H, H-1), 5.30 (dd, 1H, $J_{2,3}$ = 2.2 Hz, H-2), 5.00 (dd, 1H, $J_{3,4}$ = 5.5 Hz, H-3), 4.37 (ddd, 1H, $J_{4,5a}$ = 5.2 Hz, $J_{4,5b}$ = 6.5 Hz, H-4), 3.77 (s, 3H, OCH_3), 3.38 (dd, 1H, $J_{5a,5b}$ = 14.1 Hz, H-5a), 3.22 (dd, 1H, H-5b), 2.36 (s, 3H, SCOCH_3), 2.14 (s, 6H, 2xOCOCH_3); ^{13}C NMR ($CDCl_3$): δ 194.8 (S*C*OCH$_3$), 170.2 (O*C*OCH$_3$), 169.8 (O*C*OCH$_3$), 155.4 (C-4′), 150.2 (C-1′), 118.5 (C-2′, C-6′), 114.7 (C-3′, C-5′), 105.1 (C-1), 81.9 (C-2), 81.3 (C-4), 78.8 (C-3), 55.8 (O*C*H3), 31.1

(C-5), 30.6 (SCO*C*H$_3$), 20.9 (2x OCO*C*H$_3$); HRESIMS: Calcd for $C_{18}H_{22}NaO_8S$ [M + Na]$^+$: *m/z* 421.0933; found: *m/z* 421.0929.

Methyl 5-deoxy-5-thio-α-L-arabinofuranoside (7a)

Colourless oil (27 mg, 100%): $[\alpha]_D^{20}$ −3.1 (*c* 0.24, MeOH); R_f 0.70 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): 4.75 (d, 1H, $J_{1,2}$ = 1.8 Hz, H-1), 3.95 (dd, 1H, $J_{2,3}$ = 4.0 Hz, H-2), 3.94 (ddd, 1H, $J_{3,4} = J_{4,5b}$ = 6.4 Hz, $J_{4,5a}$ = 4.6 Hz, H-4), 3.85 (dd, 1H, H-3), 3.37 (s, 3H, OC*H*$_3$), 2.80 (dd, 1H, $J_{5a,5b}$ = 13.9 Hz, H-5a), 2.70 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 106.7 (C-1), 85.2 (C-4), 83.7 (C-2), 80.7 (C-3), 55.3 (O*C*H$_3$), 27.4 (C-5).

Benzyl 5-deoxy-5-thio-α-L-arabinofuranoside (7b)

Colourless oil (36 mg, 94%): $[\alpha]_D^{20}$ −7.2 (*c* 0.07, MeOH); R_f 0.76 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): 7.39–7.32 (m, 5H, H arom), 4.95 (d, 1H, $J_{1,2}$ = 1.8 Hz, H-1), 4.75 (d, 1H, *J* = 11.9 Hz, C*H*$_2$ benzyl), 4.53 (d, 1H, C*H*$_2$ benzyl), 4.06 (dd, 1H, $J_{2,3}$ = 4.0 Hz, H-2), 4.01 (ddd, 1H, $J_{3,4}$ = 6.6 Hz, $J_{4,5b}$ = 6.4 Hz, $J_{4,5a}$ = 4.6 Hz, H-4), 3.89 (dd, 1H, H-3), 2.84 (dd, 1H, $J_{5a,5b}$ = 13.9 Hz, H-5a), 2.73 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 139.2 (C-1′), 129.3, 129.0, 128.6 (C-2′, C-3′, C-4′, C-5′, C-6′), 108.4 (C-1), 85.1 (C-4), 83.9 (C-2), 81.0 (C-3), 70.2 (*C*H$_2$ benzyl), 27.4 (C-5).

4′-Methoxyphenyl 5-deoxy-5-thio-α-L-arabinofuranoside (7c)

Colourless oil (40 mg, 98%): $[\alpha]_D^{20}$ −71.7 (*c* 0.06, MeOH); R_f 0.77 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): 6.88 (d, 2H, *J* = 9.2 Hz, H-2′, H-6′), 6.73 (d, 2H, H-3′, H-5′), 5.27 (d, 1H, $J_{1,2}$ = 2.0 Hz, H-1), 4.12 (dd, 1H, $J_{2,3}$ = 4.4 Hz, H-2), 3.98 (ddd, 1H, $J_{3,4}$ = 6.6 Hz, $J_{4,5b}$ = 6.2 Hz, $J_{4,5a}$ = 4.8 Hz, H-4), 3.87 (dd, 1H, H-3), 3.64 (s, 3H, OC*H*$_3$), 2.74 (dd, 1H, $J_{5a,5b}$ = 13.9 Hz, H-5a), 2.63 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 156.5 (C-1′), 152.3 (C-4′), 119.4 (C-2′, C-6′), 115.5 (C-3′, C-5′), 108.7 (C-1), 85.5 (C-4), 83.9 (C-2), 80.3 (C-3), 56.0 (O*C*H3), 27.4 (C-5).

Expression and purification of Ara*f*51 E173A

Ara*f*51 E173A was produced in *Escherichia coli* strain BL21 DE3 cells harbouring the appropriate recombinant plasmid,[14] cultured in LB (Luria–Bertani) broth containing 50 mg L^{-1} kanamycin at 37 °C. Cells were grown to mid-exponential phase [A_{600} (absorbance) of 0.7], at which point isopropyl β-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM. The cultures were then incubated for a further 16 h at 37 °C. The cells were harvested by centrifugation at 4 °C for 30 min and recombinant proteins were purified from cell-free extracts by sonication followed by 15 min heat denaturation step at 70 °C. A final centrifugation at 4 °C for 30 min was performed.

General procedure for the enzymatic ligation

2′-Benzimidazolyl 1-thio-α-L-arabinofuranoside **3** (10.0 mg, 35.4 μmol) and dithiothreitol (8.9 mg, 1.6 eq.) were dissolved in tris(hydroxymethyl)aminomethane/HCl buffer (50 mM, pH 8, 10 mL). To the mixture were added successively at 37 °C the enzyme Ara*f*51 E173A (0.22 mg mL^{-1}, 45 nmol, 5 mL) freshly warmed at 37 °C and the thiol/thiophenol derivative (50 eq.). The reaction was monitored by TLC until disappearance of the donor (generally 3 h). Then the mixture was concentrated under reduced pressure and the residue was purified by column chromatography (9 : 1 CH_2Cl_2/MeOH) on silica gel to afford the desired alkyl/aryl 1-thio-α-L-arabinofuranoside.

Benzyl 1-thio-α-L-arabinofuranoside (10)

Colourless oil (0.5 mg, 5%): $[\alpha]_D^{20}$ −35.0 (*c* 0.06, MeOH); R_f 0.3 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.37–7.20 (m, 5H, C_6H_5), 4.93 (d, 1H, $J_{1,2}$ = 3.8 Hz, H-1), 4.00 (ddd, 1H, $J_{3,4}$ = 7.2 Hz, $J_{4,5b}$ = 5.4 Hz, $J_{4,5a}$ = 3.1 Hz, H-4), 3.91 (d, 1H, *J* = 12.8 Hz, C*H*$_2$ benzyl), 3.90 (dd, 1H, H-2), 3.86 (dd, 1H, $J_{2,3}$ = 7.2 Hz, $J_{3,4}$ = 4.4 Hz, H-3), 3.81 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, H-5a), 3.80 (d, 1H, C*H*$_2$ benzyl), 3.68 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 136.4 (C-1′), 132.6 (C-2′, C-6′), 129.9 (C-3′, C-5′), 128.2 (C-4′), 88.0 (C-1), 84.2 (C-4), 84.0 (C-2), 77.1 (C-3), 61.1 (C-5), 34.3 (*C*H$_2$ benzyl); HRESIMS: Calcd for $C_{12}H_{16}NaO_4S$ [M + Na]$^+$: *m/z* 279.0667; found: *m/z* 279.0666.

4′-Aminophenyl 1-thio-α-L-arabinofuranoside (11)

Colourless oil (5.5 mg, 60%): $[\alpha]_D^{20}$ −113.7 (*c* 0.21, MeOH); R_f 0.3 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.29 (d, 2H, *J* = 8.5 Hz, H-2′, H-6′), 6.65 (d, 2H, H-3′, H-5′), 4.99 (d, 1H, $J_{1,2}$ = 5.0 Hz, H-1), 3.92 (dd, 1H, $J_{2,3}$ = 5.1 Hz, H-2), 3.92–3.85 (m, 2H, H-3, H-4), 3.75 (dd, 1H, $J_{5a,5b}$ = 12.0 Hz, $J_{4,5a}$ = 2.5 Hz, H-5a), 3.61 (dd, 1H, $J_{4,5b}$ = 4.5 Hz, H-5b); ^{13}C NMR (CD_3OD): δ 149.7 (C-4′), 136.7 (C-2′, C-6′), 121.1 (C-1′), 116.5 (C-3′, C-5′), 93.8 (C-1), 84.0 (C-4), 82.6 (C-2), 77.5 (C-3), 62.5 (C-5); HRESIMS: Calcd for $C_{11}H_{15}NaNO_4S$ [M + Na]$^+$: *m/z* 280.06195; found: *m/z* 280.0619.

4′-Methylphenyl 1-thio-α-L-arabinofuranoside (12)

Colourless oil (4.5 mg, 50%): $[\alpha]_D^{20}$ −115.8 (*c* 0.24, MeOH); R_f 0.5 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.41 (d, 2H, *J* 8.0 Hz, H-2′, H-6′), 7.14 (d, 2H, H-3′, H-5′), 5.20 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 3.97 (dd, 1H, $J_{2,3}$ = 5.3 Hz, H-2), 3.96–3.92 (m, 1H, H-4), 3.93 (dd, 1H, $J_{3,4}$ = 7,5 Hz, H-3), 3.77 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5a}$ = 2.6 Hz, H-5a), 3.64 (dd, 1H, $J_{4,5b}$ = 4.6 Hz, H-5b), 2.32 (s, 3H, CH_3); ^{13}C NMR (CD_3OD): δ 140.7 (C-4′), 138.7 (C-1′), 133.4 (C-2′, C-6′), 130.6 (C-3′, C-5′), 93.3 (C-1), 84.2 (C-4), 83.2 (C-2), 77.6 (C-3), 62.4 (C-5), 21.1 (CH_3); HRESIMS: Calcd for $C_{12}H_{16}NaO_4S$ [M + Na]$^+$: *m/z* 279.0667; found: *m/z* 279.0669.

2′-Hydroxyphenyl 1-thio-α-L-arabinofuranoside (13)

Colourless oil (5.0 mg, 52%): $[\alpha]_D^{20}$ +122.9 (*c* 0.07, MeOH); R_f 0.4 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.49 (dd, 1H, $J_{5',6'}$ = 7.7 Hz, $J_{4',6'}$ = 1.8 Hz, H-6′), 7.18 (ddd, 1H, $J_{3',4'}$ = 8.2 Hz, $J_{4',5'}$ = 7.2 Hz, H-4′), 6.83 (dd, 1H, $J_{3',5'}$ = 1.1 Hz, H-3′), 6.83 (ddd, 1H, H-5′), 5.42 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 4.19 (dd, 1H, $J_{2,3}$ = 4.2 Hz, H-2), 4.06 (dd, 1H, H-3), 3.84 (ddd, 1H, $J_{3,4}$ = 8.4 Hz, $J_{4,5b}$ = 4.8 Hz, $J_{4,5a}$ = 3.8 Hz, H-4), 3.75 (dd, 1H, $J_{5a,5b}$ = 11.7 Hz, H-5a), 3.70 (dd, 1H, H-5b); ^{13}C NMR (CD_3OD): δ 158.8 (C-2′), 136.3 (C-6′), 131.0 (C-4′), 121.3 (C-5′), 120.6 (C-1′), 116.6 (C-3′), 93.0 (C-1), 87.1 (C-4), 79.5 (C-2), 78.0 (C-3), 63.2 (C-5); HRESIMS: Calcd for $C_{11}H_{14}NaO_5S$ [M + Na]$^+$: *m/z* 281.04597; found: *m/z* 281.0457.

4′-Methoxyphenyl 1-thio-α-L-arabinofuranoside (14)

Colourless oil (7.8 mg, 81%): $[\alpha]_D^{20}$ −137.5 (*c* 0.24, MeOH); R_f 0.5 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.44 (d, 2H, J = 8.8 Hz, H-2′, H-6′), 6.85 (d, 2H, H-3′, H-5′), 5.03 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 3.90 (t, 1H, $J_{2,3}$ = 4.6 Hz, H-2), 3.90–3.86 (m, 2H, H-3, H-4), 3.75 (s, 3H, OCH_3), 3.73 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5a}$ = 2.2 Hz, H-5a), 3.59 (dd, 1H, $J_{4,5b}$ = 4.6 Hz, H-5b); ^{13}C NMR (CD_3OD): δ 161.3 (C-4′), 136.2 (C-2′, C-6′), 125.8 (C-1′), 115.4 (C-3′, C-5′), 93.6 (C-1), 84.1 (C-4), 82.9 (C-2), 77.5 (C-3), 62.4 (C-5), 55.8 (OCH_3); HRESIMS: Calcd for $C_{12}H_{16}NaO_5S$ $[M + Na]^+$: *m*/*z* 295.0616; found: *m*/*z* 295.0612.

Phenyl 1-thio-α-L-arabinofuranoside (15)

Colourless oil (8.6 mg, 98%): $[\alpha]_D^{20}$ −145.2 (*c* 0.62, MeOH); R_f 0.5 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.52–7.49 (m, 2H, H-2′, H-6′), 7.31–7.21 (m, 3H, H-3′, H-4′, H-5′), 5.26 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 3.98 (dd, 1H, $J_{2,3}$ = 4.8 Hz, H-2), 3.97–3.93 (m, 1H, H-4), 3.91 (dd, 1H, $J_{3,4}$ = 7.2 Hz, H-3), 3.77 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5a}$ = 2.6 Hz, H-5a), 3.64 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5b}$ = 4.6 Hz, H-5b); ^{13}C NMR (CD_3OD): δ 136.4 (C-1′), 132.6 (C-2′, C-6′), 129.9 (C-3′, C-5′), 128.2 (C-4′), 93.1 (C-1), 84.3 (C-4), 83.4 (C-2), 77.6 (C-3), 62.4 (C-5); HRESIMS: Calcd for $C_{11}H_{14}NaO_4S$ $[M + Na]^+$: *m*/*z* 265.0510; found: *m*/*z* 265.0503.

3′-Methoxyphenyl 1-thio-α-L-arabinofuranoside (16)

Colourless oil (6.5 mg, 67%): $[\alpha]_D^{20}$ −185.2 (*c* 0.25, MeOH); R_f 0.5 (9 : 1 CH_2Cl_2–MeOH);^{1}H NMR (CD_3OD): δ 7.21 (t, 1H, $J_{4',5'}$ = 8.4 Hz, H-5′), 7.10–7.08 (m, 2H, H-2′, H-6′), 6.83 (dd, 1H, $J_{4',6'}$ = 1.6 Hz, H-4′), 5.30 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 4.00 (dd, 1H, $J_{2,3}$ = 4.8 Hz, H-2), 4.00–3.95 (m, 1H, H-4), 3.93 (dd, 1H, $J_{3,4}$ = 7.3 Hz, H-3), 3.79 (dd, 1H, $J_{5a,5b}$ = 12.2 Hz, $J_{4,5a}$ = 2.8 Hz, H-5a), 3.79 (s, 3H, OCH_3), 3.67 (dd, 1H, $J_{4,5b}$ 4.6 Hz, H-5_b); ^{13}C NMR (CD_3OD): δ 161.3 (C-3′), 137.6 (C-1′), 130.7 (C-5′), 124.6 (C-6′), 117.7 (C-2′), 113.9 (C-4′), 93.0 (C-1), 84.4 (C-4), 83.4 (C-2), 77.7 (C-3), 62.4 (C-5), 55.7 (OCH_3); HRESIMS: Calcd for $C_{12}H_{16}NaO_5S$ $[M + Na]^+$: *m*/*z* 295.0616; found: *m*/*z* 295.0619.

4′-Chlorophenyl 1-thio-α-L-arabinofuranoside (17)

Colourless oil (7.5 mg, 77%): $[\alpha]_D^{20}$ −107.6 (*c* 0.46, MeOH); R_f 0.4 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.51 (d, 2H, J = 8.7 Hz, H-2′, H-6′), 7.32 (d, 2H, H-3′, H-5′), 5.27 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 3.99 (dd, 1H, $J_{2,3}$ = 4.8 Hz, H-2), 3.99–3.95 (m, 1H, H-4), 3.93 (dd, 1H, $J_{3,4}$ = 7.3 Hz, H-3), 3.78 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5a}$ = 2.6 Hz, H-5a), 3.66 (dd, 1H, $J_{4,5b}$ = 4.8 Hz, H-5b); ^{13}C NMR (CD_3OD): δ 135.3 (C-4′), 134.3 (C-1′), 134.0 (C-2′, C-6′), 130.0 (C-3′, C-5′), 93.1 (C-1), 84.5 (C-4), 83.4 (C-2), 77.7 (C-3), 62.4 (C-5); HRESIMS: Calcd for $C_{11}H_{13}ClNaO_4S$ $[M + Na]^+$: *m*/*z* 299.0121; found: *m*/*z* 299.0112.

3′-Chlorophenyl 1-thio-α-L-arabinofuranoside (18)

Colourless oil (7.1 mg, 73%): $[\alpha]_D^{20}$ −206.5 (*c* 0.40, MeOH); R_f 0.4 (9 : 1 CH_2Cl_2/MeOH);^{1}H NMR (CD_3OD): δ 7.54 (dd, 1H, $J_{2',6'}$ = 1.8 Hz, $J_{2',4'}$ = 1.3 Hz, H-2′), 7.44 (ddd, 1H, $J_{5',6'}$ = 7.2 Hz, $J_{4',6'}$ = 1.5 Hz, H-6′), 7.29 (dd, 1H, $J_{4',5'}$ = 7.9 Hz, H-5′), 7.26 (ddd, 1H, H-4′), 5.34 (d, 1H, $J_{1,2}$ 4.4 Hz, H-1), 4.01 (dd, 1H, $J_{2,3}$ = 4.6 Hz, H-2), 4.00–3.97 (m, 1H, H-4), 3.95 (dd, 1H, $J_{3,4}$ = 7.2 Hz, H-3), 3.80 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5a}$ = 2.6 Hz, H-5a), 3.68 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5b}$ = 4.6 Hz, H-5b); ^{13}C NMR (CD_3OD): δ 139.1 (C-3′), 135.5 (C-1′), 131.3 (C-5′), 131.2 (C-2′), 130.2 (C-6′), 128.0 (C-4′), 92.9 (C-1), 84.6 (C-4), 83.5 (C-2), 77.6 (C-3), 62.3 (C-5); HRESIMS: Calcd for $C_{11}H_{13}ClNaO_4S$ $[M + Na]^+$: *m*/*z* 299.0121; found: *m*/*z* 299.0123.

4′-Carboxyphenyl 1-thio-α-L-arabinofuranoside (19)

Colourless oil (8.5 mg, 84%): $[\alpha]_D^{20}$ −142.1 (*c* 0.61, MeOH); R_f 0.35 (15 : 1 AcOEt/AcOH);^{1}H NMR (CD_3OD): δ 7.81 (d, 2H, J = 8.4 Hz, H-2′, H-6′), 7.42 (d, 2H, H-3′, H-5′), 5.30 (d, 1H, $J_{1,2}$ = 4.6 Hz, H-1), 3.93 (dd, 1H, $J_{2,3}$ = 4.6 Hz, H-2), 3.91–3.84 (m, 2H, H-3, H-4), 3.69 (dd, 1H, $J_{5a,5b}$ = 12.1 Hz, $J_{4,5a}$ = 2.6 Hz, H-5a), 3.61 (dd, 1H, $J_{4,5b}$ = 4.6 Hz, H-5b); ^{13}C NMR (CD_3OD): δ 131.0 (C-2′, C-6′), 130.4 (C-3′, C-5′), 92.4 (C-1), 84.6 (C-4), 83.6 (C-2), 77.7 (C-3), 62.4 (C-5); HRESIMS: Calcd for $C_{12}H_{14}NaO_6S$ $[M + Na]^+$: *m*/*z* 286.0511; found: *m*/*z* 285.0514.

STD-NMR

Samples were prepared in 0.5 mL of D_2O and contained ~30 μmol of **10** or **15**. STD-NMR spectra were recorded on a Bruker Avance III 400 MHz spectrometer. Processing of all data was performed on a PC with Bruker Topspin v2.0 software. After the determination of the optimal conditions, *i.e.* temperature, delay between pulse (d_{20}) and molecular ratio (protein/ligand), STD-NMR experiments were performed at 283 K as followed. The protein (1 : 100 ratio) was saturated on resonance at 0.7 ppm and off resonance at 40 ppm with a cascade of 40 selective gaussian-shaped pulses of 50 ms duration with a 100 μs delay between each pulse. The total duration of the saturation time was set to 2s. A total of 256 scans/STD-NMR experiment was acquired. A WATERGATE sequence was used to suppress residual HOD signal. A spin lock filter with strength of 5 KHz and duration of 10 ms was also applied to suppress the protein background. A similar experiment with no enzyme was used as reference in order to verify the absence of STD effect in these experimental conditions. Intensities of all STD effects were calculated though integrals over the respective signals in ^{1}H NMR reference spectra. The largest STD effect in each spectrum was set to 100% and relative intensities were determined, as common for non-refined STD effects. Hence, sufficient comparisons of relative STD effects between sugars were possible, but absolute binding intensities could not be determined.

Acknowledgements

We are grateful to the Agence Nationale de la Recherche (ANR JCJC06_140075) for financial supports and for a grant to M. A., and to the Agence des Universités Francophones (AUF) for a grant to D. D. We also thank Prof. G. J. Davies for providing the wild type and the E173A mutant of Ara*f*51, Prof. S. J. Withers for helpful discussions, and Jean-Paul Guégan (ENSCR) for recording NMR spectra.

Notes and references

1 B. Ernst and J. L. Magnani, *Nat. Rev. Drug Discovery*, 2009, **8**, 661.
2 X. M. Zhu and R. R. Schmidt, *Angew. Chem., Int. Ed.*, 2009, **48**, 1900.
3 J. D. C. Codee, R. Litjens, L. J. van den Bos, H. S. Overkleeft and G. A. van der Marel, *Chem. Soc. Rev.*, 2005, **34**, 769.

4 H. Driguez, *ChemBioChem*, 2001, **2**, 311.
5 D. J. Wardrop and S. L. Waidyarachchi, *Nat. Prod. Rep.*, 2010, **27**, 1431.
6 K. Pachamuthu and R. R. Schmidt, *Chem. Rev.*, 2006, **106**, 160.
7 M. Jahn, J. Marles, R. A. J. Warren and S. G. Withers, *Angew. Chem., Int. Ed.*, 2003, **42**, 352.
8 Y. W. Kim, H. M. Chen, J. H. Kim and S. G. Withers, *FEBS Lett.*, 2006, **580**, 4377.
9 Y. W. Kim, A. L. Lovering, H. M. Chen, T. Kantner, L. P. McIntosh, N. C. J. Strynadka and S. G. Withers, *J. Am. Chem. Soc.*, 2006, **128**, 2202.
10 J. Mullegger, H. M. Chen, R. A. J. Warren and S. G. Withers, *Angew. Chem., Int. Ed.*, 2006, **45**, 2585.
11 Y. W. Kim, H. M. Chen, J. H. Kim, J. Mullegger, D. Mahuran and S. G. Withers, *ChemBioChem*, 2007, **8**, 1495.
12 Z. Armstrong, S. Reitinger, T. Kantner and S. G. Withers, *ChemBioChem*, 2010, **11**, 533.
13 D. Shallom, V. Belakhov, D. Solomon, G. Shoham, T. Baasov and Y. Shoham, *J. Biol. Chem.*, 2002, **277**, 43667.
14 E. J. Taylor, N. L. Smith, J. P. Turkenburg, S. D'Souza, H. J. Gilbert and G. J. Davies, *Biochem. J.*, 2006, **395**, 31.
15 P. Peltier, R. Euzen, R. Daniellou, C. Nugier-Chauvin and V. Ferrières, *Carbohydr. Res.*, 2008, **343**, 1897.
16 M. R. Richards and T. L. Lowary, *ChemBioChem*, 2009, **10**, 1920.
17 I. Chlubnova, B. Sylla, C. Nugier-Chauvin, R. Daniellou, L. Legentil, B. Kralova and V. Ferrières, *Nat. Prod. Rep.*, 2011, **28**, 937.
18 I. Chlubnova, D. Filipp, V. Spiwok, H. Dvorakova, R. Daniellou, C. Nugier-Chauvin, B. Kralova and V. Ferrières, *Org. Biomol. Chem.*, 2010, **8**, 2092.
19 P. Peltier, M. Belanova, P. Dianiskova, R. Zhou, R. Zheng, J. Pearcey, M. Joe, P. Brennan, C. Nugier-Chauvin, V. Ferrières, T. Lowary, R. Daniellou and K. Mikusova, *Chem. Biol.*, 2010, **17**, 1356.
20 R. Dureau, F. Robert-Gangneux, J. Gangneux, C. Nugier-Chauvin, L. Legentil, R. Daniellou and V. Ferrières, *Carbohydr. Res.*, 2010, **345**, 1299.
21 M. A. Kelly, M. L. Sinnott and D. Widdows, *Carbohydr. Res.*, 1988, **181**, 262.
22 V. Ferrières, S. Blanchard, D. Fischer and D. Plusquellec, *Bioorg. Med. Chem. Lett.*, 2002, **12**, 3515.
23 R. Euzen, J. P. Guegan, V. Ferrières and D. Plusquellec, *J. Org. Chem.*, 2007, **72**, 5743.
24 A. V. Demchenko, N. N. Malysheva and C. De Meo, *Org. Lett.*, 2003, **5**, 455.
25 K. Mizutani, R. Kasai, M. Nakamura, O. Tanaka and H. Matsuura, *Carbohydr. Res.*, 1989, **185**, 27.
26 E. C. B. Johnson and S. B. H. Kent, *J. Am. Chem. Soc.*, 2006, **128**, 6640.
27 M. Mayer and B. Meyer, *Angew. Chem., Int. Ed.*, 1999, **38**, 1784.

