Imaging Tools for Probing Glycosaminoglycans In Vivo in Developing Zebrafish

by

Brendan James Beahm

A dissertation submitted in partial satisfaction of the requirements for the degree
Doctor of Philosophy
In
Chemistry
in the
Graduate Division
of the
University of California, Berkeley

Committee in charge:
Professor Carolyn R. Bertozzi, Chair
Professor Matthew B. Francis
Professor Michelle C. Chang
Professor Craig T. Miller

Spring 2014
Imaging Tools for Probing Glycosaminoglycans In Vivo in Developing Zebrafish

© 2014

by Brendan James Beahm
Abstract

Imaging Tools for Probing Glycosaminoglycans In Vivo in Developing Zebrafish

by

Brendan James Beahm

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Carolyn R. Bertozzi, Chair

Glycosaminoglycans are linear polysaccharides that decorate the surfaces of all animal cells. The heterogeneity of the polysaccharide, especially the pattern of sulfation on the distal end of the chain, endows these glycans with diverse biological functions including essential roles in signaling during animal development. Thus, glycosaminoglycans represent an extremely important and interesting class of glycans to probe using metabolic labeling in concert with bioorthogonal chemistry.

This thesis describes the extension of the metabolic labeling strategy to glycosaminoglycans (GAGs). Prior to the work described herein, GAGs could not be metabolically labeled because xylose, the monosaccharide specific to these glycans, lacks a salvage or biosynthetic pathway capable of converting xylose derivatives, or other simple sugars, to the ultimate sugar donor UDP-xylose. In order to label GAGs, the UDP-xylose derivative itself must be delivered to cells. However, the charged nature of nucleotide sugars prevents passive diffusion through hydrophobic membranes, and eukaryotic cells lack a plasma membrane nucleotide sugar transporter to provide facilitated or active transport. Prior research in our lab, however, had demonstrated that unnatural nucleotide sugars could be delivered to cells by microinjection into the yolk sacs of developing zebrafish embryos. In addition to providing a tractable solution to deliver nucleotide sugars, zebrafish are a well-established vertebrate development model organism with excellent properties for molecular imaging.

It was hypothesized that GAGs could be metabolically labeled by microinjecting unnatural variants of UDP-xylose into zebrafish embryos. To this end, three analogs of UDP-xylose were synthesized, microinjected into zebrafish embryos, and analyzed for incorporation onto zebrafish embryo cell surfaces by reaction with fluorescent cyclooctyne probes. Two of the analogs, UDP-2-XylAz and UDP-3-XylAz, did not provide azide-dependent labeling. UDP-4-XylAz did provide azide-dependent labeling, and the azide replacing the C-4 hydroxyl group inhibits extension of the GAG polysaccharide beyond the initial xylose unit. Therefore, UDP-4-XylAz functions as a
selective metabolic inhibitor of GAG biosynthesis. The phenotypic consequences of aberrant GAG production were ascertained and these results add to our understanding of the importance of GAGs during normal vertebrate development.
This dissertation is dedicated to my family.
Table of Contents

List of Figures v
List of Schemes viii
List of Tables viii
Acknowledgements ix

Chapter 1. Imaging Glycans in Animals with Bioorthogonal Chemistry 1
  Introduction 2
  Principles 3
  Carbonyl/α-effect nucleophiles 4
  Staudinger Ligation 5
  Azide-Alkyne cycloaddition 8
  Extending in vivo glycan visualization to glycosaminoglycans 12
  Conclusion 13
  References 15

Chapter 2. Synthesis of Azido-UDP-Xylose Analogs 18
  Introduction 19
  Results and Discussion 24
Chapter 3. Probing Glycosaminoglycans in Developing Zebrafish Embryos

Introduction

Results and Discussion

UDP-2-XylAz and UDP-3-XylAz do not provide azide-dependent labeling

UDP-4-XylAz provides azide-dependent labeling

UDP-4-XylAz delivers 4-XylAz to sites of GAG modification

UDP-4-XylAz functions as a metabolic inhibitor of GAG biosynthesis

Aberrant GAG glycosylation induced by UDP-4-XylAz causes dramatic phenotypes in developing zebrafish embryos

Conclusion

Experimental Methods

General materials and methods
Zebrafish stocks and husbandry  76

Confocal microscopy imaging of labeling with each UDP-XylAz derivative  76

Flow cytometry measurement of labeling with each UDP-XylAz derivative  77

XylT1 MO activity verification using RT-PCR  77

XylT1 MO attenuation of the UDP-4-XylAz signal  78

Determining the effect of the XylT MO on fucose expression  78

XylT activity assay with UDP-4-XylAz  79

Determining if β4GalT7 can elaborate a 4-XylAz modified peptide  79

The effect of UDP-4-XylAz on GAG expression and composition  80

Determining the effect of UDP-4-XylAz on alcian blue and alizarin red staining  80

Determining the effect of UDP-4-XylAz on zebrafish embryo gross morphology by live imaging  80

References  81
## List of Figures

| Figure 1.1 | Metabolic labeling of cell-surface glycans followed by bioorthogonal reaction with imaging probes enables visualization in living systems | 3 |
| Figure 1.2 | Carbonyl condensation reaction partners and product, with corresponding reaction mechanism | 4 |
| Figure 1.3 | Pictet-Spengler ligation forms a covalent adduct which is hydrolytically stable | 5 |
| Figure 1.4 | The Staudinger Ligation proceeds between an azide and phosphine | 6 |
| Figure 1.5 | Smart phosphine probes enhance signal to noise for azide-dependent labeling with the Staudinger ligation | 7 |
| Figure 1.6 | Copper-catalyzed azide-alkyne cycloaddition enables rapid bioconjugation | 8 |
| Figure 1.7 | Incorporation of an alkyne into an eight membered ring enhances its reactivity sufficiently so that it does not require toxic copper for catalysis | 9 |
| Figure 1.8 | Panel of the most commonly used cyclooctynes | 10 |
| Figure 1.9 | Metabolic incorporation and click imaging in zebrafish embryos | 11 |
| Figure 1.10 | Azidosugars used image different glycan types in live developing zebrafish embryos | 12 |
| Figure 1.11 | Tumor imaging in mice using nanoparticles modified with cyclooctynes and fluorophores | 13 |
| Figure 1.12 | Glycosaminoglycans have numerous functions, especially during development | 14 |
| Figure 2.1 | Xylose initiates diverse glycosaminoglycan structures and can be targeted for metabolic replacement | 19 |
| Figure 2.2 | UDP-Xylose (UDP-Xyl) biosynthesis |
| Figure 2.3 | Panel of UDP-XylAz analogs considered for synthesis |
| Figure 2.4 | Potential decomposition pathway of 5-XylAz |
| Figure 2.5 | Chemical modulators of GAG biosynthesis are limited |
| Figure 2.6 | Steric hindrance by the axial methoxy group is hypothesized to be the primary cause for the poor reactivity of the trifluoromethanesulfonate 7 |
| Figure 2.7 | Hypothesized explanation for preferential $\beta$ phosphate formation by 15 |
| Figure 2.8 | The primary, C-5 hydroxyl group of xylofuranose 21 is preferentially protected with the bulky silyl group over that of the secondary, C-3 hydroxyl group |
| Figure 2.9 | Steric contributions from the isopropylidene protecting group prevent Re-face attack by the reducing hydride |
| Figure 2.10 | Mechanism for the rearrangement of 3-azidoxylfuranose to its pyranose form |
| Figure 2.11 | Silyl substitution proceeds through a trigonal bipyramidal intermediate |
| Figure 2.12 | Resonance can explain the unique reactivity of the anomeric position of sugars |
| Figure 3.1 | Hypothesized scheme for metabolic labeling of glycosaminoglycans |
| Figure 3.2 | Azide-dependent labeling was not observed in zebrafish embryos injected with UDP-2-XylAz or UDP-3-XylAz |
| Figure 3.3 | Flow cytometry experimental setup that enables detection of azide-dependent labeling on zebrafish embryos |
| Figure 3.4 | Fluorescence signal of enveloping layer cells from zebrafish embryos injected with each UDP-XylAz analog |
| Figure 3.5 | Azide-dependent labeling is observed for embryos injected with UDP-4-XylAz |
Figure 3.6  UDP-4-XylAz functions as a visualizable inhibitor of GAG biosynthesis  66

Figure 3.7  Schematic of XylT assay  67

Figure 3.8  XylT1 can use UDP-4-XylAz as a substrate while XylT2 cannot  68

Figure 3.9  XylT1 MO at a dose of 6 ng effectively inhibits the production of XylT1 mRNA  69

Figure 3.10  Flow cytometry revealed that the XylT1 MO attenuates the UDP-4-XylAz signal  70

Figure 3.11  Fucosylation is unaffected by injection of XylT1 MO  71

Figure 3.12  4-XylAz prevents elaboration by β4GalT7  72

Figure 3.13  Effects of UDP-4-XylAz on GAG production  73

Figure 3.14  UDP-4-XylAz injection causes embryological defects  74
List of Schemes

**Scheme 2.1** Retrosynthetic scheme for the production of UDP-XylAz analogs 24

**Scheme 2.2** Synthesis of C-2 azide-modified methylxyloside 26

**Scheme 2.3** Synthesis of phenylthioxyloside modified with an azide at the C-2 position 26

**Scheme 2.4** Synthesis of bisacetal protected 2-azidoxylose-1-phosphate 27

**Scheme 2.5** Synthesis of UDP-2-XylAz 29

**Scheme 2.6** Synthesis of 3-azidoxylofuranose 24 30

**Scheme 2.7** Synthesis of UDP-3-XylAz 33

**Scheme 2.8** Synthesis of benzoyl-protected 4-XylAz-1-diallylphosphate 34

**Scheme 2.9** Synthesis of UDP-4-XylAz 35

List of Tables

**Table 3.1** XylT and β-actin primers used to amplify mRNA sequences 69
Acknowledgments

I am profoundly grateful to all the people who helped shape my PhD at UC Berkeley. First and foremost, I owe an enormous amount of gratitude to my advisor, Carolyn Bertozzi. Carolyn possesses one of the most creative and intelligent minds I've ever known and it was a pleasure learning from her. Further, Carolyn established an extremely unique lab setting that facilitates creativity and instills independence in all graduate students who are privileged to perform research in this environment. I will always be grateful for my time with Carolyn.

Another essential player in my scientific development was Sharon Amacher. Sharon is one of the kindest and most patient mentors with whom I have ever come in contact. Sharon’s breadth and depth of knowledge regarding zebrafish biology is paralleled by none, and it was a privilege to study under her tutelage.

I owe enormous gratitude to my graduate student and post-doctoral mentors, whose influence and guidance was essential to my development as a scientist. Karen Dehnert was the key person in my scientific education, most especially in zebrafish biology. In addition to her guidance in science, Karen has been a great friend outside of the lab, continuously aiding in my personal and professional development. John Jewett and Ellen Sletten were my synthetic chemistry mentors whose wisdom extends beyond that of bench science, and they too have remained great friends and key fixtures in my personal development. Mike Boyce is one of the most interesting, entertaining, and impressive people I’ve ever known. His knowledge of “biology” has no match, and without it my research progress would have undoubtedly been greatly retarded. In addition to his scientific advice, Mike has always been willing to lend an ear and dispense thoughtful and informative personal advice; I am fortunate to have his friendship. Ben Swarts is one of my closest friends to this day, and he too was instrumental in my development as a scientist and person, always willing to set aside a few minutes to delve into an interesting scientific or personal problem that I was facing. Brian Belardi and Jason Hudak, though only one to two years ahead of me in graduate school, respectively, were completely essential in guiding my experimental design in both carbohydrate chemistry and biological evaluation of my synthetic compounds. Brian’s guidance extended beyond science and he was also key in my personal development.

Beyond the mentorship provided by the people described above, I would be remiss to not mention the friendships I have gained in graduate school whose longevity is without question. Gabriela de Almeida, Evan Adams, Frances Rodriguez-Rivera, Sloan Siegrist, Bastiaan Buddingh, Doug Fox, Steve Canham, Neil Rumachik, Mason Apple, Nathan Yee, Lissette Andres, Ellie Smith, Ian Blong, Kanna Palaniappan, Kimberly Beatty, Mark Breidenbach, Kamil Godula, Shana Topp, Lauren Wagner, Chelsea Gordon, Paresh Agarwal, Peter Robinson, Hitomi Nakamura, Yisu Han, and Ioannis Mountziaris were essential contributors to my happiness on a daily basis. Lastly, I will be forever grateful for the friendship offered by Anthony Baldridge in times of need.
The speed at which I progressed through my thesis project could not have been attained had it not been for the impressive and ceaseless effort by my collaborators. Dorothe Spillmann, though several time zones away, was always prompt and thorough with experiments, analysis, and responses to questions. Joachim Kuhn provided essential experiments to validate the functions of the compounds that I synthesized and found creative uses for these compounds that would have otherwise been inconceivable by myself. Lastly, Nicholas Derr, in collaboration with Sharon Amacher, supplied data without which my thesis work could not have been published.

In addition to the scientific and personal guidance provided by the scientists mentioned above, I could not have persevered through graduate school and developed personally without the following people. My parents, Nancy and Dusty Beahm, in addition to my grandmother Joan Beahm, provided endless love and support throughout my graduate career. My siblings Alyson Young, Andrea Beahm, Brian Beahm, Kelly Randall, and Kendal Scullion never missed a phone call and stood by me through the toughest of times. Lastly, Rachel Seele was the most instrumental person in my personal development; I will forever be grateful for her contribution to my life.

Thank you all.
Chapter 1

Imaging glycans in animals with bioorthogonal chemistry
Chapter 1. Imaging Glycans in Animals with Bioorthogonal Chemistry

Introduction

The cell surface glycome comprises diverse glycans that contribute to virtually all aspects of a cell’s social life. Glycans contribute directly to cell-cell and cell-matrix interactions, affect the organization and turnover kinetics of cell surface molecules and modulate the activity of signaling receptors. Accordingly, they are central players in cell differentiation, proliferation and migration and in the translation of signals from the tissue microenvironment to cellular responses. The cell surface glycome can also be considered a complex data set that reports on the underlying cell’s physiology. Governed by numerous inputs – expression of glycosyltransferase genes, protein and lipid scaffold availability, nutrient status, and secretory pathway activity, for example – the composition of the cell-surface glycome changes in response to cellular changes in a manner that has diagnostic value. The unique glycomic signatures of cancer cells, embryonic stem cells and activated leukocytes and endothelial cells, in particular, have attracted much attention as targets for diagnostic probes.

The ability to visualize changes in cell surface glycosylation as a function of space and time, ideally in live organisms, is therefore an important goal toward advancing our understanding of glycobiology. While molecular imaging techniques targeting proteins have advanced considerably in recent decades (e.g., genetically encoded fluorescent proteins, bioluminescence methods, antibody-targeted diagnostics), it is only recently that the challenge of in vivo glyc an imaging has begun to be met. Initial work focused on using glycan-binding proteins such as antibodies and lectins to direct a conjugated fluorescent probe to glycan structures of interest. These techniques were quite effective for profiling global glycosylation changes on cultured cells and tissue slices but have limited applicability for imaging in live organisms. Antibodies and lectins often have low binding affinities for their glycan epitopes, poor tissue penetrance, and, in the case of lectins, can be toxic to cells.

The historic dearth of tools available for in vivo glycan profiling spurred the development of more recently reported chemical approaches. This chapter focuses on the combined use of metabolic labeling and bioorthogonal chemistry as a means to visualize glycans in vivo (Figure 1.1). The two-step process begins with administration of a simple sugar analog functionalized with a bioorthogonal chemical group. The sugar is processed by cells within the organism, and its metabolic products are integrated into complex glycans ultimately displayed on cell surfaces. Next, a probe molecule (e.g., fluorophore, luminescent nanoparticle, radiolabel, MRI contrast reagent) bearing complementary bioorthogonal functionality is administered to the organism. The selective, and ideally very fast, reaction among the two bioorthogonal groups delivers the probe to the glycans of interest, enabling their visualization using various imaging modalities. This chapter aims to educate readers on the history of bioorthogonal reaction development and modern applications toward imaging glycans in animals.
Principles

Initial inspiration for the metabolic labeling strategy came from the work of Werner Reutter and coworkers demonstrating that unnatural analogs of N-acetylmannosamine (ManNAc), a metabolic precursor to sialic acid, are converted to the corresponding unnatural sialic acid derivatives in cultured cells and in rodents. This was an important discovery, since it showed that the sialic acid biosynthetic pathway could tolerate unnatural precursors and deliver them to a locale, the cell surface, accessible to myriad probe reagents.

The next conceptual challenge was to identify functional groups that could be integrated into such unnatural ManNAc analogs and that could undergo highly selective conjugation reactions with probes. For the chemical reactions to occur in vivo, the reagents would have to be bioorthogonal, meaning that they must neither interact nor interfere with surrounding biological functionalities. Moreover, their mutually selective reaction must proceed under physiological conditions: in water, at 37 °C and pH 7, and with no toxic byproducts. Additionally, the reaction partner found on the metabolic precursor sugar, referred to as the “reporter” group, must be small enough to enable processing by biosynthetic enzymes. Finally, the reaction must have inherently fast kinetics. This requirement stems from the relatively low concentrations of glycan-associated reporter group and circulating probe molecule that can be achieved in living systems. As well, observing dynamic changes in glycosylation that occur over short time periods mandates the use of reactions that occur on a fast time scale.

Figure 1.1: Metabolic labeling of cell-surface glycans followed by bioorthogonal reaction with imaging probes enables visualization in living systems. Blue hexagon, natural sugar; red hexagon, unnatural sugar; X, reporter group; Y, complementary reactive group; green star, fluorophore.
Carbonyl/α-effect nucleophiles

The first reaction used for probing glycans in living systems was the condensation of carbonyls and α-effect nucleophiles. While the condensation between ketones or aldehydes and amines is typically sluggish, the corresponding reaction with α-effect nucleophiles such as aminooxy or hydrazide compounds are much faster (Figure 1.2). The nitrogen atoms of aminooxy and hydrazide groups have enhanced nucleophilicity due to their direct attachment to another heteroatom.

Typically, the ketone has been used as the metabolic label and the hydrazide for probe delivery. In an early example by Bertozzi and coworkers, a ketone-functionalized analog of ManNAc, N-levulinoylmannosamine (ManLev), was used as the metabolic precursor following Reutter and coworkers' precedent. The electrophilic ketone was chosen as the chemical reporter group because a hydrazide-based reporter would readily react with other biological ketones and aldehydes present in the cytosol (e.g., pyruvate and free monosaccharides). Although ketones and aldehydes are found within cells and in circulation, bringing into question the true bioorthogonality of this reaction, they are not otherwise present on cell surfaces where the probe ligation reaction occurs. Accordingly, it was demonstrated that treatment of cells with ManLev produced N-levulinoyl sialylated glycans on cell surfaces, enabling their visualization with hydrazide-functionalized imaging probes. While ManLev with free hydroxyl groups was administered to cells at millimolar concentrations to achieve metabolic labeling, micromolar concentrations could be employed using the peracetylated analog. Presumably, peracetylated ManLev diffused through membranes more efficiently than

![Figure 1.2: Carbonyl condensation reaction partners and product, with corresponding reaction mechanism.](image-url)
The Pictet-Spengler ligation forms a covalent adduct which is hydrolytically stable. The hydrophilic free sugar, and was subsequently deacetylated by cytosolic esterases prior to metabolism.

While useful for some applications, condensation reactions of carbonyls and α-effect nucleophiles were not particularly useful for imaging glycans in animals. These reactions require acidic pH (i.e., 5.5-6.5) for efficient labeling, which is impossible to achieve in vivo. As well, the hydrazide or oxime bond formed in the reaction is susceptible to hydrolysis. Recently, a carbonyl condensation reaction referred to as the Pictet-Spengler ligation that produces a hydrolytically-stable C-C bonded adduct was reported (Figure 1.3). But no matter the stability of the product, nor the kinetics that can be achieved at physiological pH, ketone and aldehyde reporter groups suffer from the fundamental problem that they are not unique in a biological setting. Thus, an important component of this in vivo glycan imaging approach is the development of bioorthogonal reactions among functionalities that have no counterpart whatsoever in animals. The first reaction that met this challenge was the Staudinger ligation.

Staudinger ligation

The Staudinger ligation is highly selective reaction between azide and phosphine reagents, resulting in the formation of an amide bond. The development of this reaction was inspired by the classic Staudinger reduction of azides to the corresponding amines (Figure 1.4). In the Staudinger ligation, the nucleophilic phosphine first attacks the electrophilic azide (Figure 1.4). Following extrusion of dinitrogen, an aza-ylide intermediate forms and rapidly reacts in intramolecular fashion with a nearby ester group to form a stable amide bond.

For glycan imaging, the azide has been used as the reporter group while the phosphine was used to deliver the imaging probe. The azide has proven to be a quintessential reporter group due to its small size, chemical stability in biological settings, unique reactivity, and absence from biology. Prior to its use in metabolic labeling, the azide was widely employed in synthetic chemistry as protected amine equivalent. Accordingly, there are many known methods for installing azides in small molecule substrates, particularly sugars where the azide has long been used in the construction of aminosugar analogs.

The Staudinger ligation reaction circumvented many of the limitations of ketone and aldehyde-based condensations. The ligation proceeds efficiently at neutral pH, and the generated amide bond is stable in vivo. These features, in combination with its exceptional selectivity and biocompatibility, enabled the Staudinger ligation to be the first bioorthogonal chemical reaction performed in live mice. In this study, mice were injected with peracetylated N-azidoacetylmannosamine (ManNAz) once daily for several
Figure 1.4: The Staudinger Ligation proceeds between an azide and phosphine. The reaction partners are shown in the top panel and the reaction mechanism is shown in the bottom panel.

days, then subsequently administered a biotinylated phosphine probe. After a few hours, organs were harvested and the Staudinger ligation product was detected by flow cytometry analysis of splenocytes and by Western blotting of organ lysates.

Unfortunately, the Staudinger ligation did not prove to be very useful for in vivo glycan imaging, due primarily to its relatively slow kinetics (second-order rate constant $k \approx 10^{-3}$ M$^{-1}$s$^{-1}$). There are two strategies for accelerating reactions with inherently slow kinetics: elevate temperature and increase reagent concentrations. For in vivo imaging, temperature cannot be altered outside the physiological range. The concentration of metabolically azide-labeled glycans in vivo is low and also not subject to much external control. The only variable that can be experimentally altered, therefore, is probe concentration administered to the organism. However, for fluorescent probes, high concentrations are difficult to achieve due to solubility limits, and if injected at high levels into animals, unreacted probe molecules are difficult to wash away. The consequence is high background fluorescence that obscures any fluorophore chemically attached to glycans of interest. Attempts were made to design phosphine reagents that react more rapidly with azides. Unfortunately, phosphines with enhanced azide reactivity were also more susceptible to spontaneous air oxidation, a significant liability for biological imaging applications.
Other work focused on designing phosphine reagents that underwent an enhancement in fluorescence during the Staudinger ligation; in principle, such "fluorogenic" probes could be used at higher concentrations without the burden of washing away unreacted probe and associated background fluorescence. Two design concepts were explored for producing fluorogenic phosphine probes.\(^7,8\) In the first, the mechanism for fluorescence turn-on relied on photo-induced electron transfer (PET) quenching by the phosphorous atom's lone pair of electrons (Figure 1.5A).\(^7\) During the

**Figure 1.5:** Smart phosphine probes enhance signal to noise for azide-dependent labeling with the Staudinger ligation. The first fluorogenic phosphine probe relied on PET fluorescence quenching (A), while the second harnessed FRET fluorescence quenching.
Staudinger ligation, the phosphine is converted to a phosphine oxide that lacks a lone pair of electrons and therefore no longer quenches fluorescence. Unfortunately, air oxidation of the phosphine also gave rise to fluorescence turn-on, limiting its use for biological imaging studies.

A second fluorogenic probe was designed based on fluorescence resonance energy transfer (FRET) quenching (Figure 1.5B).\textsuperscript{8} For this probe, a fluorophore was conjugated to the phosphine and a FRET quencher was attached through the ester linkage. Cleavage of the ester during the Staudinger ligation reaction released the quencher and activated fluorescence. This probe was used to image glycans on live cells, however the slow reaction rate necessitated multi-hour incubations, which, although never directly tested, would likely undermine in vivo imaging due to competing liver metabolism of the phosphine probe. Thus, faster reaction rates were urgently sought to realize the goal of in vivo glycan imaging using azidosugar reporters.

### Azide–Alkyne cycloaddition

Two years after the Staudinger ligation was described, the fast and highly chemoselective copper-catalyzed azide–alkyne cycloaddition (CuAAC) reaction was reported (Figure 1.6).\textsuperscript{9–11} Prior to this, the cycloaddition reaction between azides and alkynes to form triazoles was well-known in the organic chemistry literature.\textsuperscript{12} This classic transformation is thermodynamically favorable ($\Delta G^\circ \approx -61$ kcal mol$^{-1}$) but under standard conditions it does not proceed readily due to a high kinetic barrier. In CuAAC, the copper(I) catalyst lowers the kinetic barrier by combining with the alkyne to form a reactive copper acetylide species. This intermediate reacts rapidly with azides to form the 1,4-triazole product. Sharpless and coworkers coined the term “click chemistry” to refer to such fast, selective and reliable transformations, and CuAAC is often considered to be the quintessential click chemistry reaction.

The speed of CuAAC dwarfs that of the Staudinger ligation by over 3 orders of magnitude. However, copper(I) is highly toxic to cells at concentrations required for catalysis, largely precluding the use of CuAAC for imaging live cells and organisms. CuAAC was, however, used to image glycans on fixed cells, and enabled visualization of a snapshot of glycans in the process of intracellular trafficking.\textsuperscript{13} Recently, a copper ligand was reported that enables in vivo CuAAC labeling of glycans on cells and in developing zebrafish embryos.\textsuperscript{14} While promising, this reaction still must be performed for a short duration, suggesting that toxicity during longer copper(I) exposures is an issue.

![Figure 1.6: Copper-catalyzed azide-alkyne cycloaddition enables rapid bioconjugation.](image-url)
An alternative means for enhancing the azide-alkyne cycloaddition reaction kinetics has been identified that does not require a toxic metal catalyst: ring strain. Alkynes constrained in 8-membered rings, i.e., cyclooctynes, possess considerable ring strain (~19 kcal/mol) due to bond angle distortion from the ideal of 180° down to about 155°. As a result, cyclooctynes are high-energy substrates for cycloaddition with azides, during which much of this strain energy is released (Figure 1.7). Unsubstituted cyclooctynes react spontaneously with azides to form triazole products, whereas their linear unstrained counterparts require elevated temperatures or pressures.

The reactivity of cyclooctynes in this “strain-promoted azide-alkyne cycloaddition (SPAAC)”, also sometimes referred to as “copper-free click chemistry”, can be tuned over at least 3 orders of magnitude. Unsubstituted cyclooctynes typically react with azides with second-order rate constants on par with that of the Staudinger ligation. However, introduction of two fluorine atoms next to the alkyne, as in the difluorinated cyclooctyne “DIFO”, increases the second-order rate constant 60-fold (Figure 1.8). This improvement was sufficient to enable glycan imaging after reacting azidosugar-treated cells with a DIFO-fluorophore conjugate for one minute.

Further enhancements in cyclooctyne reactivity were achieved by fusing the ring to two benzene rings, as in the dibenzocyclooctynes (DIBO, Figure 1.8), dibenzo-aza-cyclooctynes (DIBAC, Figure 1.8) and bisaryl-aza-cyclooctynones (BARAC, Figure 1.8). As well, a substantive kinetic enhancement was observed upon fusion of the cyclooctyne to a cyclopropyl ring, generating a bicyclononyne (BCN, Figure 1.8) structure. The increased reactivity of these various ring-fused cyclooctynes may be due to enhanced ring strain. Considerable computational work has now been performed to understand the relationship of cyclooctyne structure and azide cycloaddition reactivity,

![Reaction](image)

**Mechanism**

![Mechanism](image)

**Figure 1.7**: Incorporation of an alkyne into an eight membered ring enhances its reactivity sufficiently so that it does not require toxic copper for catalysis. The top panel shows the reaction partners and product and the bottom panel describes the mechanism of the reaction.
particularly by Houk, Goddard and Alabugin.\textsuperscript{17–21} DIBO, DIBAC, BARAC and BNC react with azides 100-1000-times faster than the parent unsubstituted cyclooctyne, and can be used to image glycans on cultured cells at sub-micromolar concentrations. Nowadays, many of these cyclooctyne-probe conjugates are available from commercial sources.

The advent of copper-free click chemistry enabled numerous recent \textit{in vivo} imaging studies using zebrafish as a model organism. With different azide-functionalized metabolic precursors, cell-surface sialylated glycans, mucin-type O-glycans, and fucosylated glycans have been imaged in developing zebrafish embryos.\textsuperscript{22–25} The rapid kinetics of the reaction enabled differentiation of glycans with distinct spatiotemporal expression patterns. In some studies, glycans were temporally resolved by pulse-chase experiments; embryos were incubated or injected with azidosugar then reacted with a cyclooctyne-fluorophore conjugate at a certain time point (Figure 1.9). After a second period of azidosugar incubation, the embryos were reacted with a spectrally distinct cyclooctyne-fluorophore at a later time point. Through this workflow, the spatiotemporal dynamics of different glycan subtypes were analyzed (Figure 1.10). These investigations revealed that cells from different tissues of the embryo internalize their glycans at different rates. For example, during the sixty-sixth hour post fertilization (hpf), little internalization of mucin-type O-glycans was observed in eye cells, while almost complete internalization was observed in pectoral fin cells \textsuperscript{22}. It was also found that glycan expression in the different anatomical structures of the embryo is time-dependent; e.g. probing for sialylated glycans revealed their time-dependent expression in the unique structures of the olfactory organ \textsuperscript{25}. A surprising finding was that cell-surface glycans undergo a rapid reorganization at the cleavage furrow of dividing cell during early embryogenesis (10 hpf) \textsuperscript{23}.

\begin{figure}[h]
\centering
\includegraphics[width=0.8\textwidth]{cyclooctynes.png}
\caption{Panel of the most commonly used cyclooctynes. Their reactivity with azides was greatly enhanced over that of unmodified cyclooctyne using ring-strain and/or electronics. Figure shows DIFO, DIBO, DIBAC, BARAC and BCN.}
\end{figure}
Figure 1.9: Metabolic incorporation and click imaging in zebrafish embryos. After introducing azidosugar by incubation or microinjection, embryos are allowed to develop and incorporate the analog into cell surface glycans. Following this, glycans are detected using fluorophore-modified cyclooctyne probes.

While informative, these studies were limited to visualizing of glycans on the outermost layer of embryo cells. Attempts to image internal zebrafish glycans revealed that the reaction rate is not the only parameter that determines labeling efficiency. Visualization of internal zebrafish glycans requires tissue access by cyclooctyne probes, which appears to be limited, at least for the commercial fluorophores that have been tested. As well, efforts to use cyclooctyne probes to image azidosugar-labeled glycans in mice were initially hampered by physical properties of those probes. For example, while DIFO conjugates react efficiently with cell-surface azidosugars in culture cells, in mice such conjugates have limited bioavailability due to their tight and possibly covalent association with serum albumin. Thus a major future goal is to optimize cyclooctyne-probe conjugates for pharmacokinetic properties including tissue distribution, metabolic stability and bioavailability in serum.

Nonetheless, cyclooctyne-azide cycloaddition chemistry has been employed for glycan imaging in mice, with a focus on tumor detection. Kim and coworkers showed that cyclooctyne-nanoparticle conjugates can detect azido sialic acid residues metabolically introduced onto on tumor cells in live mice. With several cyclooctynes per nanoparticle, the resulting multivalency was thought to increase the nanoparticle’s reaction efficiency with cell-surface glycans. Further, each nanoparticle was outfitted with multiple fluorophores, generating a bright signal and enhancing detection sensitivity. Still, to achieve detectable labeling above background, the azidosugar
Figure 1.10: Azidosugars used image different glycan types in live developing zebrafish embryos. A) Peracetylated N-acetylgalactosamine (GalNAz) labels mucin-type O-glycans. B) Peracetylated ManNAz labels sialylated glycans. C) 6-Azidofucose labels fucosylated glycans.

substrate, ManNAz, had to be directly injected into the tumor rather than administered systemically; thus, it is unclear how much of the observed signal resulted from reaction of the nanoparticles with unprocessed ManNAz that was not associated with cell-surface glycans.

Extending in vivo glycan visualization to glycosaminoglycans

Although metabolic incorporation of the azide functional group into glycans followed by detection with a fluorophore-functionalized cyclooctyne probe has enabled the in vivo detection of mucin-type O-glycans, sialic acids, and fucosylated glycans, the extension of this technology to visualizing glycosaminoglycans (GAGs) had not been accomplished at the outset of this work.

Glycosaminoglycans, linear polysaccharides composed of repeating disaccharides, are important components of animal cell surfaces and the extracellular matrix. The most widely studied GAGs are heparan sulfate (HS) and chondroitin sulfate (CS), which have distinct disaccharide repeats but share a common core tetrasaccharide that links them to an underlying protein scaffold (Figure 2.1). Variable sulfation along a GAG chain can generate an enormous array of structures, allowing them to interact with diverse biomolecules. For example, GAG binding to signaling molecules, including members of the Hedgehog, TGFβ and FGF families, and to extracellular matrix molecules and basement membrane components such as laminin, is essential for proper animal development (Figure 1.12).<sup>29,30</sup> Accordingly, mutations in the GAG biosynthetic machinery can lead to dramatic phenotypes.<sup>31</sup> For example, mice
Figure 1.11: Tumor imaging in mice using nanoparticles modified with cyclooctynes and fluorophores. Mice are injected intratumorally with ManNAz. After being given time to metabolically incorporate the unnatural sugar, mice are injected intravenously with cyclooctyne- and fluorophore-modified nanoparticles, and imaged. Lacking Ext1 or Ext2, the enzymes responsible for polymerizing HS, die in early development from defective gastrulation.\textsuperscript{32–34} GAG biosynthesis mutations in zebrafish embryos have been shown to cause craniofacial developmental defects.\textsuperscript{35} Their crucial biological functions, especially during animal development, make GAGs an important class of glycans to visualize \textit{in vivo}.

Conclusion

Over the past decade, significant progress has been made in the development of bioorthogonal reactions for imaging glycans in model organisms. Following the observation that ketone-modified metabolic precursors could be processed by endogenous biosynthetic pathways and detected on live cells, significant work has gone into developing reactions with improved properties. Now, azide-functionalized sugars and cyclooctyne-probe as well as tetrazine-probe conjugates are all available from commercial sources, making \textit{in vivo} glycan imaging accessible to the broader community of biologists. Early applications have focused on studies of glycan expression during zebrafish embryogenesis and tumor detection in murine xenograft.
Figure 1.12: Glycosaminoglycans have numerous functions, especially during development. For example, GAGs function as co-receptors for ligands and their receptor target (A), they modulate the expression of signaling peptides by binding them until signaled to release and concurrently protect them from proteolysis (B), and they facilitate intercellular communication (C).

models, but there is a rich future for expansion of the metabolic labeling/bioorthogonal chemistry approach into other systems. In vivo models of inflammation, wound healing and microbial infection are all attractive venues for imaging the dynamics of cell surface glycosylation associated with disease.
References


(9) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. Angew. Chem. Int. Ed. 2002, 41, 2596.

(10) Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. Angew. Chem. 2002, 114, 2708.


Chapter 2

Synthesis of Azido-UDP-Xylose Analogs
Chapter 2. Synthesis of Azido-UDP-Xylose Analogs*

Introduction

In order to metabolically label GAGs, a monosaccharide that would specifically target them was selected. To determine the appropriate monosaccharide to derivatize with an azide, the sugar composition of GAGs was consulted. All three major families of GAGs possess a conserved core tetrasaccharide directly linked to a serine residue on the polypeptide backbone. This core glyccan is composed of a glucuronic acid (GlcA)-\(\beta(1,3)\)-galactose (Gal)-\(\beta(1,3)\)-galactose (Gal)-\(\beta(1,3)\)-xylose (Xyl) backbone (Figure 2.1). The major GAG classes include heparan sulfate (HS), chondroitin sulfate (CS), and dermatan sulfate (DS) and are differentiated by the distal repeating disaccharide unit. HS contains GlcA-\(\beta(1,4)\)-N-acetylglucosamine (GlcNAc)-\(\alpha(1,4)\) as well as IdoA-\(\alpha(1,4)\)-GlcNAc-\(\alpha(1,4)\), CS possess GlcA-\(\beta(1,3)\)-N-acetylgalactosamine (GalNAc)-\(\beta(1,4)\), and DS is composed of the CS disaccharide as well as IdoA-\(\alpha(1,3)\)-GalNAc-\(\beta(1,4)\).

Thus, structural analysis of the monosaccharides found in GAGs indicates that the candidates for unnatural metabolite incorporation are limited to GlcA, Gal, GalNAc, GlcNAc, and Xyl. IdoA is generated by epimerization of GlcA after GlcA is incorporated into GAGs, and therefore IdoA cannot function as a metabolic label for GAGs. GlcA, Gal, GalNAc, and GlcNAc abound in other types of glycans and are not likely to label

**Figure 2.1:** Xylose initiates diverse glycosaminoglycan structures and can be targeted for metabolic replacement. Heparan sulfate and chondroitin/dermatan sulfate structures. \(R^1 = H\) or \(SO_3^-\); \(R^2 = Ac\) or \(SO_3^-\). GlcA, glucuronic acid; IdoA, iduronic acid; GlcNAc, \(N\)-acetylglucosamine; GalNAc, \(N\)-acetylgalactosamine; Gal, Galactose; Xyl, Xylose; asterisk, position at which epimerization can occur.

GAGs specifically. GlcA is found in glucuronylation, Gal and GlcNAc are found in mucin-type O-glycans, N-glycans, and glycolipids, and GalNAc is found in mucin-type O-glycans and glycolipids. Furthermore, unnatural, azide-containing variants of GalNAc and GlcNAc have previously been shown to label mucin-type O-glycans and GlcNAc modified proteins. None of these monosaccharides could serve as a specific metabolic label for GAGs. Xylose, however, is found almost exclusively in GAGs. Xylose has been detected in other glycans, including a trisaccharide found on EGF-like repeat regions of certain proteins, and a recently discovered GAG-like glycan which decorates α-dystroglycan, but these specific glycans are thought to be rare relative to GAGs. Thus, xylose was selected as the monosaccharide for targeting GAGs.

GAG biosynthesis is initiated by the action of a xylosyltransferase, which appends a Xyl monosaccharide to a serine residue of a protein using the activated xylosyl donor uridine 5’-diphosphate xylose (UDP-Xyl). UDP-Xyl is biosynthesized from the highly abundant donor hexose, UDP-glucose (UDP-Glc) (Figure 2.2). Beginning in the cytosol, the 6-position of UDP-Glc is oxidized to a carboxylic acid by the enzyme UDP-Glc dehydrogenase to generate UDP-GlcA. After being transported into the Golgi, decarboxylation of the 6-position of UDP-GlcA, catalyzed by UDP-Xyl synthase, yields UDP-Xyl. UDP-Xyl is then appended to the protein core at a serine residue via the action of a xylosyltransferase (XylT1 or XylT2), initiating GAG biosynthesis (Figure 2.1 and 3.1). Since there is no known salvage pathway to biosynthesize UDP-Xyl from free xylose in vivo, unnatural UDP-Xyl derivatives must be synthesized in order to label GAGs metabolically.

Lack of structural information for the active sites of the XylTs made rational design of the UDP-Xyl analogs impossible. Instead a panel of xylose derivatives was designed and synthesized subject to three criteria: (i) chemical stability; and minimal perturbation of the natural sugar’s (ii) size and (iii) shape. When only considering the

![Figure 2.2: UDP-Xylose (UDP-Xyl) biosynthesis](image)

Figure 2.2: UDP-Xylose (UDP-Xyl) biosynthesis. UDP-Xyl biosynthesis begins from UDP-glucose (UDP-Glc), which is oxidized by UDP-Glc dehydrogenase to furnish UDP-glucuronic acid (UDP-GlcA), which is subsequently decarboxylated by UDP-Xyl synthase to give UDP-Xyl.
affect on the size and shape of the natural sugar, the four possible UDP-XylAz derivatives are those in which the azide is found at the C-2 (1), C-3 (2), C-4 (3), or C-5 (4) position (Figure 2.3). After factoring in the stability of the compounds, an azide at the C-5 position of xylose was rejected due to the fact that the sugar can form a dialdehyde and expel the azide group (Figure 2.4). Therefore, the target compounds were UDP-2-XylAz (1), UDP-3-XylAz (2), and UDP-4-XylAz (3) (Figure 2.3).

Prior to their synthesis it was important to consider the advantages and disadvantages of each unnatural analog. While the introduction of an azide into UDP-Xyl at any position could prevent recognition by the XylTs, UDP-2-XylAz and UDP-4-XylAz have particular features that warrant further discussion. The C-2 hydroxyl group of xylose is transiently phosphorylated during construction of the core tetrasaccharide. Recent studies have identified the kinase, FAM20B, and phosphatase, XYLIP, responsible for this modification of xylose and have convincingly demonstrated when the modification occurs during GAG chain construction. Additionally, studies have been performed to ascertain the purpose of this modification. Overexpression of FAM20B resulted in increased GAG production, while RNAi knockdown of FAM20B attenuated GAG expression. Interestingly, knockdown of the phosphatase also decreased the level of GAG production. Together, these findings suggest that xylose phosphorylation modulates GAG expression. Thus, eliminating the sugar’s ability to undergo this modification by replacing the C-2 hydroxyl group with an azide was expected to perturb GAG production in vivo. However, because a very small number of natural sugars are typically replaced by azidosugars it was thought that the adverse effects imparted through the loss of modification would be limited.

While UDP-2-XylAz was not expected to dramatically affect GAG biosynthesis, UDP-4-XylAz lacks the hydroxyl group needed for GAG chain elaboration and thus was

![UDP-2-XylAz (1) UDP-3-XylAz (2) UDP-4-XylAz (3) UDP-5-XylAz (4)](image)

Figure 2.3: Panel of UDP-XylAz analogs considered for synthesis. Analogs are named according to the position on the pentose ring where the azide is found; an azide at the C-2, C-3, C-4, or C-5 position correspond to UDP-2-XylAz (1), UDP-3-XylAz (2), UDP-4-XylAz (3), and UDP-5-XylAz (4), respectively.
Figure 2.4: Potential decomposition pathway of 5-XylAz.

predicted to function as a metabolic inhibitor rather than a non-perturbing metabolic label. While the goal at the outset of this work was to generate a metabolic label to visualize and probe glycosylation in its natural state, chemical tools for modulating the biosynthesis of GAGs in cells and organisms were quite limited and therefore represented a ripe area for development. The majority of work in the area of modulators of GAG biosynthesis has focused on aryl xylosides as “primers” of GAG biosynthesis.\textsuperscript{18,19} As the first sugar of the conserved core tetrasaccharide, xylose is attached through a $\beta$-linkage to serine residues on the peptide backbone (Figure 2.5A). $\beta$-xylosides with hydrophobic aglycones can mimic xylosylated proteins and compete for the elaborating enzymes (Figure 2.5B). These compounds do not technically inhibit GAG biosynthesis, but rather provide an alternative substrate for GAG elaboration and lead to the production of proteoglycans without GAGs. Further, when the GAG-modified $\beta$-xylosides are secreted into the extracellular environment they can perform other functions and confound interpretation of the observed biological effects.\textsuperscript{20} Recently, this approach was modified to generate true inhibitors of GAG biosynthesis by replacing the C-4 hydroxyl group of $\beta$-xylosides with a fluorine atom.\textsuperscript{21,22} These fluorinated xylose analogs effectively inhibit GAG biosynthesis in cultured cells.

Chain-terminating metabolic inhibitors comprise a second, though less well-developed, class of GAG biosynthesis disruptors (Figure 2.5C).\textsuperscript{23–26} These compounds are deoxy or fluorinated monosaccharides that lack the key hydroxyl group needed for polysaccharide elongation. Their metabolic conversion to nucleotide-sugar glycosyl donors and subsequent enzymatic incorporation into growing GAG chains leads to truncation and, therefore, incomplete GAG biosynthesis. A complication inherent to this approach is that most GAG monosaccharide constituents are widely distributed among other glycan types. The corresponding deoxy and fluoro analogs therefore perturb numerous glycan structures in addition to GAGs.\textsuperscript{24,27} However, because xylose is almost exclusively found in GAGs it was hypothesized that UDP-4-XylAz would selectively inhibit GAG biosynthesis.\textsuperscript{8,9}

Having determined the analogs to synthesize, a general retrosynthesis was designed (Scheme 2.1). A chemoenzymatic route was initially considered for the final step, in which XylAz-1-phosphate analogs would be ligated with uridine triphosphate (UTP) to generate each UDP-XylAz analog. However, it was unknown whether the pyrophosphorylase required for the transformation would accept all, or indeed any, of the unnatural analogs. It is possible that mutagenesis could be used to design enzymes capable of utilizing the unnatural substrates, however this process would be laborious.
Figure 2.5: Chemical modulators of GAG biosynthesis are limited. (A) Natural GAG biosynthesis; (B) aryl xyloside primers serve as alternative substrate for GAG biosynthetic enzymes, and therefore generate proteins devoid of GAGs; (C) chain-terminating monosaccharides incorporated into GAG chains prevent further chain elongation because they lack the necessary hydroxyl group.

and time-consuming, and thus it was decided that chemical synthesis should be used instead. Chemical ligation of sugar-1-phosphates with N-methylimidazole activated uridine monophosphate (UMP) is a well-established means of generating nucleotide sugars although the process is typically low yielding (~20-30%). Further, because the azidosugars are delivered to cells by microinjection into zebrafish embryos (discussed in Chapter 3) very little compound is required, and consequently a high yield is not a priority. For example, 10 mg of nucleotide sugar makes sufficient 100 mM solution to conduct more than 300 experiments each of which would inject hundreds embryos with 1-2 nL of solution.

Having settled on a plan for nucleotide sugar formation, the key step of the synthesis, prior steps were easily decided (Scheme 2.1). It was hypothesized that the sugar-1-phosphate analogs could be generated by reacting selectively protected XylAz analogs, which only contain a free hydroxyl group at the anomeric position, with a phosphoramidite reagent. Each XylAz analog would in turn be generated by activating the appropriate hydroxyl group of a xylose epimer as a leaving group and displacing it with a nucleophilic azide. In order to minimize the number of steps for each synthesis it was decided that all would begin from commercially available xylose epimers. This strategy obviates the need for additional synthetic steps to properly orient the hydroxyl groups for displacement with an azide. UDP-3-XylAz, however, was synthesized
Scheme 2.1: Retrosynthetic scheme for the production of UDP-XylAz analogs. UDP-XylAz can be generated through the ligation of XylAz-1-phosphate with UMP. XylAz-1-phosphate could be generated through the phosphorylation of XylAz with a free hydroxyl group at the C-1 position. XylAz could be generated by displacing the epimeric hydroxyl group of a xylose epimer.

starting from xylose following a known procedure for efficiently and facilely producing the selectively protected C-3 epimer of xylose, ribose, from a commercially available xylose analog.

Results and Discussion

Synthesis of UDP-2-XylAz

The initial approach for generating UDP-2-XylAz utilized the methyl D-lyxoside 5 (Scheme 2.2). An efficient method for generating the methyl D-lyxoside 5 involves stirring D-lyxose in excess acetyl chloride to afford the methyl glycoside 5 in 97% yield. Following the work of Frost and coworkers, the C-3 and C-4 hydroxyl groups of the methyl D-lyxoside were selectively protected using a butane 2,3-bisacetal protecting group to afford 6 in 52% yield. The C-2 hydroxyl group of 6 was activated as a leaving group by conversion to a trifluormethanesulfonate group, generating 7 in 55% yield. However, product 7 proved to be more robust than was intended, and no displacement occurred in the presence of lithium azide at room temperature. After varying the reaction conditions, the methyl azido xylopyranoside 8 was obtained in 70% yield by performing the reaction at 80 °C. One possible explanation for this reactivity is steric hindrance by the α-methoxy group at the anomeric position: at room temperature the ring is locked.
Figure 2.6: Steric hindrance by the axial methoxy group is hypothesized to be the primary cause for the poor reactivity of the trifluoromethanesulfonate 7.

with the methoxy group blocking attack; however, at elevated temperature, the bonds in the ring have enough rotational energy to distort and accommodate nucleophilic attack of the azide (Figure 2.6).

Methyl glycosides are typically reverted to free hydroxyl groups and butane 2,3-bisacetal protecting groups are typically cleaved using acidic conditions. It was therefore hypothesized that a global deprotection of 8 would provide the 2-azidoxyl derivative 9. However, attempts at this global deprotection under a variety of conditions failed to yield the desired product because the harsh acidic conditions required for cleaving the methyl glycoside decomposed the sugar (Scheme 2.2). Due to the fact that the methyl glycoside could not be deprotected, a more mild protecting group was needed.

Thioglycosides, a common anomeric protecting scheme, are stable to a variety of conditions and yet can be easily cleaved by activation with electrophiles (Scheme 2.3). In order to generate the phenylthioglycoside of lyxose, free D-lyxose was first peracetylated with acetic anhydride to yield 10 quantitatively. The peracetylated phenylthiolyxoside 11 was then generated in 87% yield by treating peracetylated lyxose with thiophenol and the Lewis acid, boron trifluoride-diethyl etherate. After treatment with sodium methoxide to cleave the acetyl protecting groups and generate the free phenylthiolyxoside 12 in 95% yield, the butane 2,3-bisacetal protecting group was used to protect the C-3 and C-4 hydroxyl groups, giving 13 in 38% yield. This yield and reported product warrant further discussion, While the starting material is a mixture of α and β anomers, only the β product is reported. This results from a practical issue with purification. Although the C-3 and C-4 hydroxyl groups of both anomer starting materials are efficiently protected during the reaction, the resulting products could not be purified to absolute purity using silica gel column chromatography. Unfortunately, the impurity could not be identified and eluent conditions were not found that could enable its separation from the desired products. Serendipitously, the β anomer readily precipitated from crude reaction mixture following concentration in vacuo, which enabled isolation of 13.

After obtaining the bisacetal protected phenylthiolyxoside 13, its C-2 hydroxyl group was activated as a leaving group by conversion to a trifluoromethanesulfonate group and displaced using lithium azide to generate the bisacetal protected phenyl-2-azido-thio-β-xyloside 14 in 85% yield over two steps. Interestingly, unlike the azide displacement reaction with the methyllyxoside 7, heat was not required for the conversion of 13 to 14; this supports our hypothesis that the axially oriented methoxy group of 7 sterically blocked nucleophilic azide attack. The thiophenol group of 14 was easily removed using N-bromosuccinimide, providing 15 in 75% yield (Scheme 2.4).
Scheme 2.2: Synthesis of C-2 azide-modified methylxyloside. Detailed experimental conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) Acetyl chloride, MeOH, 50 °C, 97%; (b) 2,2,3,3-tetramethoxybutane, trimethyl orthoformate, cat. CSA, MeOH, reflux, 52%; (c) Tf₂O, pyridine, DCM, –20 °C, 55%; (d) LiN₃, DMF, 80 °C, 70%; (e) various acidic conditions.

Scheme 2.3: Synthesis of phenylthioxyloside modified with an azide at the C-2 position. Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) Ac₂O, pyridine, 0 °C to rt, quant.; (b) PhSH, BF₃.OEt₂, DCM, 0 °C to rt, 87%; (c) NaOMe, MeOH, rt, 95%; (d) 2,2,3,3-tetramethoxybutane, trimethyl orthoformate, cat. CSA, MeOH, reflux, 38%; (e) i. Tf₂O, pyridine, DCM, -20 °C to rt; ii. LiN₃, DMF, rt, 85% (2 steps).
With the C-2 and C-3 hydroxyl groups still protected, a phosphate group could be installed at the anomeric position to give 16 in 94% yield (Scheme 2.4). The typical procedure for this transformation utilizes dibenzyl diisopropylphosphoramidite, followed by oxidation of the resulting phosphite to its corresponding phosphate, and subsequent cleavage of the benzyl groups by hydrogenolysis.\textsuperscript{32} In addition to deprotecting benzyl groups, hydrogen over palladium on activated charcoal also reduces azides to amines so this approach could not be employed. However, a similar method using diallyl diisopropylphosphoramidite had previously been shown to obviate this problem because the corresponding protected phosphate could be deprotected with tetrakis(triphenylphosphine)palladium under conditions that would not react with azides.\textsuperscript{4} This strategy was very effective at installing an allyl-protected phosphate at the anomeric position, but favored formation of the β anomer ($\alpha:\beta$ 1:1.5) where the α anomer is the one required for GAG initiation. Interestingly, the $\alpha:\beta$ ratio of the starting material contained more α anomer ($\alpha:\beta$, 1.2:1). This suggests that the α and β starting materials are rapidly interconverting, and the β anomer starting material has a lower kinetic barrier for forming the allyl phosphate product. This decreased kinetic barrier could be due to the significant 1,3-diaxial strain present when the anomeric hydroxyl group is in the axial position, which make that hydroxyl group sterically congested and disfavored (Figure 2.7). In addition to the issue of preferential β anomer formation, cleavage of the bisacetal protecting group requires harsh acidic conditions which would likely hydrolyze the anomeric phosphate group.

In order to improve the proportion of α phosphate anomer formation and eliminate the potential issue of phosphate hydrolysis, the synthetic route was reconsidered (Scheme 2.5). Bisacetal protection results in the formation of a rigid bicyclic molecule, which locks each of the axial hydrogens in their conformation (Figure 2.7A). It was thought that if the bisacetal protecting group was removed prior to phosphorylation of the anomeric position, this would attenuate the rigidity of the molecule and enable more bond rotation; thus, the 1,3-diaxial strain would be lessened for the α anomer of xyloside (18), resulting in a more favorable equilibrium for its formation, and thus a greater proportion of α phosphate would form following treatment with the phosphoramidite. Further, because the bisacetal protecting group would be removed prior to phosphorylation, the acidity needed for its cleavage would not be used in the presence of the acid-labile phosphate group. Thus, rather than cleaving the

Scheme 2.4: Synthesis of bisacetal protected 2-azidoxylose-1-phosphate. Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) NBS, H$_2$O, acetone, rt, 75%; (b) i. diallyl $N,N$-diisopropylphosphoramidite, 1$H$-tetrazole, DCM, rt; ii. mCPBA, DCM, $-40 \degree$C, 94% (2 steps).
Figure 2.7: Hypothesized explanation for preferential $\beta$ phosphate formation by 15. 1,3-diaxial interactions are thought to explain why bisacetal-protected xyloside 15 prefers $\beta$ anomer formation relative to acetyl-protected 18 when reacted with diallyl $N,N$-diisopropylphosphoramidite.

thiophenyl group of 14, first the bisacetal protecting group was removed using TFA in water, and the resulting free hydroxyl groups were protected with acetals using acetic anhydride and pyridine, giving 17 in 95% yield over two steps. Following this, the thiophenol group was removed using electrophilic bromine donated by $N$-bromosuccinimide (NBS) to furnish 18 in 92% yield. Diallyl $N,N$-diisopropylphosphoramidite was then used to generate the diallyl phosphite, which was immediately oxidized to the phosphate using meta-chloroperoxybenzaldehyde (mCPBA) to give the diallyl phosphate 19 in 29% yield with an anomeric ratio of $\alpha:\beta$, 1.1:1. The improved $\alpha:\beta$ ratio over that achieved using the bisacetal protecting group ($\alpha:\beta$, 1:1.5) supports our hypothesis regarding the cause of preferential $\beta$ anomer formation. Importantly, the reported yield is only for the $\alpha$ anomer, and it is only low due to the purification. The anomeric mixture could be purified away from other impurities using silica gel column chromatography, but the separation of $\alpha$ and $\beta$ anomers required reverse phase HPLC.

Finally, UDP-2-XylAz (1) was produced in two steps from diallyl phosphate 19, closely following the procedure developed by Hang and coworkers. Tetrakis(triphenylphosphine)palladium (Pd(PPh$_3$)$_4$) was first used to cleave the allyl groups of 19, exposing the phosphate oxygens for the ligation reaction. In a separate flask, the electrophilically activated UMP-$N$-methylimidazolide 19a was prepared, and subsequently combined with the newly formed sugar-1-phosphate to produce the nucleotide sugar. Following acyl group deprotection, 1 was obtained in 26% yield over three steps. While this yield may seem low at first glance, it is actually quite reasonable considering that the yields for ligations between pure sugar-1-phosphates and activated UMP are typically between 20-30%. This suggests that both phosphate and acyl group deprotections were nearly quantitative. Further, this yield reflects a multistep purification required to attain the purity necessary for use with zebrafish embryos (discussed in Chapter 3).
Scheme 2.5: Synthesis of UDP-2-XylAz (1). Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) i. TFA, H$_2$O, rt; ii. Ac$_2$O, pyridine, 0 °C to rt, 95% (2 steps); (b) NBS, H$_2$O, acetone, rt, 92%; (c) i. diallyl N,N-diisopropylphosphoramidite, 1H-tetrazole, DCM, rt; ii. mCPBA, DCM, –40 °C, 29% (2 steps); (d) i. Pd(PPh$_3$)$_4$, sodium p-toluenesulfinate, THF, MeOH, rt; ii. UMP-N-methylimidazolide, MeCN, 0 °C; iii. NEt$_3$, MeOH, H$_2$O, rt, 26%.

Although obtaining 1 in 26% yield was moderately good at the time of this synthesis, noteworthy improvements have been made to nucleotide sugar synthesis in recent years. Future implementations could utilize the 2011 method of Tsukamoto and Kahne that produces yields between 70-80% using an N-methylimidazolium catalyst.$^{33}$ Alternately, the method of Mohamady and Taylor further improved yields to greater than 90% using sulfonyl imidazolium salts.$^{34}$ While high yielding, these reactions, like that used in the synthesis of UDP-2-XylAz, require strict anhydrous conditions and are therefore synthetically challenging and unapproachable for researchers who lack synthetic prowess. However, in 2013, Tanaka and coworkers showed that the coupling reaction could be very routinely performed in water with moderate yields using an in situ generated 2-imidazoyl-1,3-dimethylimidazolinium reagent.$^{35}$ These improvements to nucleotide sugar synthesis will undoubtedly result in the synthesis of many novel unnatural nucleotide sugars, and enable creative applications of these compounds. For example, because unnatural nucleotide sugars need only be recognized by glycosyltransferases, unlike unnatural peracetylated monosaccharides that must traverse additional enzymes to first be processed into nucleotide sugars, it is likely that larger and more complex unnatural functional groups will be delivered to cell surface glycans.

**Synthesis of UDP-3-XylAz**

Unlike the C-2 and C-4 modified UDP-XylAz analogs, the synthesis of UDP-3-XylAz began with a protected xylose monosaccharide, rather than its C-3 epimer, ribose (Scheme 2.6). Xylose was used in lieu of ribose because an efficient, high-yielding
synthesis already existed for generating selectively protected ribose, with only a free hydroxyl group at the desired C-3 position.\textsuperscript{36}

The synthesis began from the xylofuranose 20, whose C-1 and C-2 hydroxyl groups were masked with an isopropylidene protecting group (Scheme 2.6). This compound was used because in one single step the C-3 hydroxyl group could be isolated based on the differential reactivity of primary and secondary alcohols; primary alcohols are more reactive as nucleophiles than secondary alcohols due to reduced steric congestion (Figure 2.8). This selectivity is especially heightened when using bulky electrophiles, thus the sterically encumbered tert-butyl(chloro)diphenylsilane (TBDPS) was used to selectively protect the primary C-5 hydroxyl group, giving 21 in quantitative yield. Next, because the remaining C-3 hydroxyl group was improperly oriented for S\textsubscript{N}2 displacement with an azide, it was inverted. This was achieved in two steps, starting with a Swern oxidation to convert the alcohol group of 21 to a ketone, furnishing 22 in 97% yield. The resulting ketone was reduced with sodium borohydride to give the selectively protected ribose 23 with a free hydroxyl group only at the C-3 position in 67% yield; the selectivity of this reduction is imparted by steric blocking of the re-face by the isopropylidene protecting group, resulting in addition of hydride from the si-face (Figure 2.9).

Having isolated and properly oriented the C-3 hydroxyl group of xylose, giving the ribose compound 23, an azide could be installed at the C-3 position to give the fully protected 3-XylAz analog 24. This was achieved in two steps in one pot by first activating the hydroxyl group as a leaving group by converting it to a

\begin{align*}
\text{HO-}\begin{array}{c}
\text{HO} \\
\text{TBDPSO}
\end{array}
\text{20} & \xrightarrow{\text{a}} 
\begin{array}{c}
\text{HO} \\
\text{TBDPSO}
\end{array} & \text{HO} \\
\text{TBDPSO}
\text{21} & \xrightarrow{\text{b}} 
\begin{array}{c}
\text{HO} \\
\text{TBDPSO}
\end{array} & \text{HO} \\
\text{TBDPSO}
\text{22} & \xrightarrow{\text{c}} 
\begin{array}{c}
\text{N}_3 \\
\text{TBDPSO}
\end{array}
\text{24} & \xrightarrow{\text{d}} 
\begin{array}{c}
\text{HO} \\
\text{TBDPSO}
\end{array} & \text{HO} \\
\text{TBDPSO}
\text{23} & \xrightarrow{\text{e}} 
\begin{array}{c}
\text{HO} \\
\text{TBDPSO}
\end{array} & \text{HO} \\
\text{TBDPSO}
\text{19}
\end{align*}

\textbf{Scheme 2.6: Synthesis of 3-azidoxylofuranose 24.} Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) TBDPSCI, imidazole, DMF, 0 °C, quant.; (b) i. oxalyl chloride, DMSO, DCM, −60 °C; ii. NE\textsubscript{t}\textsubscript{3}, DCM, −60 °C to rt, 97% (2 steps); (c) NaBH\textsubscript{4}, EtOH, H\textsubscript{2}O, 0 °C, 67%; (d) i. Tf\textsubscript{2}O, pyridine, DCM, −20 °C; ii. LiN\textsubscript{3}, DMF, rt, 71% (2 steps).
Figure 2.8: The primary, C-5 hydroxyl group of xylofuranose 21 is preferentially protected with the bulky silyl group over that of the secondary, C-3 hydroxyl group. Primary hydroxyl groups have less steric bulk than secondary hydroxyl groups, and thus a lower kinetic barrier to formation of the silyl ether when reacted with an electrophilic silyl reagent.

Figure 2.9: Steric contributions from the isopropylidene protecting group prevent Re-face attack by the reducing hydride.
trifluoromethanesulfonate, using trifluoromethanesulfonic anhydride (Tf₂O), which was subsequently displace with nucleophilic azide using lithium azide (LiN₃) to give 24 in 71% yield over two steps.

In order to be utilized by xylosyltransferases, xylose must be in its pyranose form. Thus, xylofuranose 24 was fully deprotected, allowed to rearrange to the pyranose and subsequently protected (Scheme 2.7). This was realized by first deprotecting the silyl protecting group using tetrabutylammonium fluoride as a nucleophilic fluoride source to give 25 in 91% yield. Although the fluoride anion is not typically considered a good nucleophile, it has an affinity for silicon and its small atomic radius enables its penetration of sterically congested silyl ethers. Interestingly, the mechanism for the deprotection of silyl ethers is not S_N2, like one might expect, but proceeds through a trigonal bipyramidal intermediate (Figure 2.11). Following deprotection of the silyl ether, the isopropylidene protecting group was removed under aqueous, acidic conditions using trifluoracetic acid. The resulting free xylofuranose then rearranged to its pyranose form (Figure 2.10) and was protected with acetyl groups using acetic anhydride, affording the peracetylated azidoxylopyranose 26 in 78% yield over two steps. Because hydroxyl phosphorylation is only desired at the anomeric position, the C-1 acetyl group was selectively deprotected using benzylamine, providing 27 in 82% yield. Although the nitrogen of benzylamine can add in to any of the acyl groups, resulting in their deprotection, it is selective for the anomeric acetyl group due to the heightened reactivity of that acyl group. The anomeric position has enhanced reactivity due to its proximity to the endocyclic oxygen, which provides it with unique resonance structures and therefore reactivity (Figure 2.12).

The selectively protected 3-XylAz analog 27 was converted to UDP-3-XylAz (2) following the same procedure that was used in the synthesis of UDP-2-XylAz (1). First, the anomeric hydroxyl group was phosphorylated by treating 27 with diallyl N,N-diisopropylphosphoramidite to give the diallyl phosphite, which was immediately oxidized, using mCPBA, to the corresponding phosphate 28 in 64% yield over two steps (Scheme 2.7). The diallyl phosphate of compound 28 was then deprotected using Pd(PPh₃)₄ and reacted with the electrophilically activated UMP-N-methylimidazolide 19a. Following deprotection and an extensive purification, UDP-3-XylAz (2) was

Figure 2.11: Silyl substitution proceeds through a trigonal bipyramidal intermediate.
Scheme 2.7: Synthesis of UDP-3-XylAz (2). Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) TBAF, THF, rt, 91%; (b) i. TFA, H₂O, 0 °C; ii. Ac₂O, pyridine, 0 °C to rt, 78% (2 steps); (c) benzylamine, THF, 80 °C, 82%; (d) i. diallyl N,N-diisopropylphosphoramidite, 1H-tetrazole, DCM, rt; ii. mCPBA, DCM, −40 °C, 64% (2 steps); (e) i. Pd(PPh₃)₄, sodium p-toluenesulfinate, THF, MeOH, rt; ii. UMP-N-methylimidazolide, MeCN, 0 °C; iii. NEt₃, MeOH, H₂O, rt, 8.4% (3 steps).

obtained in 8.4% yield over three steps. The cause of this low yield is open to speculation, though it is worth noting that minor degradation occurred during lyophilization of the compound. Because lyophilization is not typically considered a harsh treatment this may indicate UDP-3-XylAz has poor inherent stability and the degradative reaction observed after lyophilization could in fact have occurred during purification prior to lyophilization.

Synthesis of UDP-4-XylAz

The synthesis of UDP-4-XylAz began from the C-4 epimer of xylose, L-arabinose (Scheme 2.8). Using a known procedure, L-arabinose was selectively benzoylated using benzoyl chloride to afford 1,2,3-tri-O-benzoyl-L-arabinose (29) in 29% yield. While

Figure 2.12: Resonance can explain the unique reactivity of the anomeric position of sugars.
low yielding, this procedure allowed isolation of the C-4 hydroxyl group in a single step from an inexpensive, commercially available substrate. This selectivity is attributable to the steric effect of the axial hydrogen, which impedes the axial C-4 hydroxyl group from attacking bulky benzoyl chloride molecules (Scheme 2.8). While 29 is the main tribenzoylated product formed other benzyolated products are also formed, especially perbenzyolated arabinose. Similar to the other XylAz syntheses previously described, the tribenzoylated compound 29 was converted to perbenzyolated azidoxyloside 30 in two steps. This was achieved by activating the free hydroxyl group of 29 as a leaving group using Tf₂O, and subsequently displacing it with an azide using LiN₃ to give 30 in 85% yield over two steps.

Following generation of perbenzyolated xyloside 30, the benzyol group at the anomeric position was selectively removed using ethylenediamine to afford 31 in 43% yield. While low yielding, a significant amount of the starting material (52%) could be recovered; thus this is still a viable method for removal of the anomeric benzyol group. The resulting hydroxyl group was reacted with diallyl N,N-diisopropylphosphoramidite, followed by mCPBA to give the diallylphosphate 32 in 77% yield. Unfortunately, this procedure resulted in almost exclusive formation of the β anomer, and because the α anomer is required, the synthetic route had to be reconsidered. Before a new route could be determined, the cause of predominant β anomer formation was contemplated, and determined to be most likely attributable to both sterics and electronics. Specifically, the electron donating property of the C-2 benzoyl group’s phenyl ring make its ester oxygens more electron dense than those of an acetyl group, and they are therefore more prone to donate into the antibonding orbital of the β C-O bond and stabilize that.

Scheme 2.8: Synthesis of benzoyl-protected 4-XylAz-1-diallylphosphate (32).
Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) BzCl, pyridine, −35 °C to rt, 29%; (b) i. Tf₂O, pyridine, DCM, −20 °C; ii. LiN₃, DMF, rt, 85% (2 steps); (c) ethylenediamine, acetic acid, THF, 0 °C to rt, 43%; (d) i. diallyl N,N-diisopropylphosphoramidite, 1H-tetrazole, DCM, rt; ii. mCPBA, DCM, −40 °C, 45% (2 steps).
conformation. Additionally, the bulky benzoyl groups likely interact with the axial hydroxyl group when it sits in the α conformation and destabilize its formation; however, because the benzoyl groups are in the equatorial orientation it is unlikely that this effect is large.

In order to mitigate these effects, the benzoyl groups were replaced with smaller, less electron donating acetyl groups prior to phosphorylation (Scheme 2.9). This was facilely achieved by deprotecting the benzoyl groups of 30 by transesterification with methoxide in methanol, followed by peracetylation with acetic anhydride to afford peracetylated xyloside 33 in 80% yield over two steps. After deprotecting the anomeric position with hydrazine acetate to give the selectively protected 4-XylAz compound 34 in 54% yield, the C-1 hydroxyl group was phosphorylated with diallyl \( N,N \)-diisopropylphosphoramidite and immediately oxidized to the diallyl phosphate using mCPBA to afford 35 in 61% yield over two steps in an \( \alpha:\beta \) ratio of 2.1:1. This improvement in the \( \alpha:\beta \) ratio supports the hypothesized reason for preferential β anomer formation for the benzoylated derivative 31. Although moderately high yielding, the desired α anomer was only obtained in 31% yield because HPLC was required for its separation from the β anomer. Finally, UDP-4-XylAz (3) was generated in two steps by first treating 35 with Pd(Ph\(_3\))\(_4\) to cleave the allyl groups from the phosphate, and then reacting the resulting deprotected phosphate with the activated UMP-\( N \)-methylimidazolide 19a. After cleaving the acyl protecting groups with methoxide, UDP-4-XylAz (3) was obtained in 22% yield over three steps.

Scheme 2.9: Synthesis of UDP-4-XylAz (3). Detailed reaction conditions can be found in the Materials and Methods section. Reagents/Conditions: (a) i. NaOMe, MeOH, rt; ii. Ac\(_2\)O, pyridine, 0°C to rt, 80% (2 steps); (b) hydrazine acetate, DMF, rt, 54%; (c) i. diallyl \( N,N \)-diisopropylphosphoramidite, 1H-tetrazole, DCM, rt; ii. mCPBA, DCM, −40 °C, 31% (2 steps); (d) i. Pd(Ph\(_3\))\(_4\), sodium \( p \)-toluenesulfinate, THF, MeOH, rt; ii. UMP-\( N \)-methylimidazolide, MeCN, 0 °C; iii. NEt\(_3\), MeOH, water, rt, 22% (3 steps).
Conclusion

Starting from xylose epimers or a partially protected xylose monosaccharide, UDP-2-XylAz (1), UDP-3-XylAz (2) and UDP-4-XylAz (3) were synthesized. While these syntheses were not efficient in terms of overall yield, they were concise (6-9 steps) and enabled the generation of target compounds in less than one month each. Armed with these UDP-xylose derivatives, their ability to be incorporated into cell surface glycosaminoglycans of developing zebrafish could be determined.
Experimental Methods

General materials and methods

Chemical reagents were obtained from commercial sources and used without further purification unless otherwise noted. All anhydrous solvents used were either purchased from commercial providers or used from an aluminum column solvent purification system. All reactions were performed under nitrogen in oven-dried glassware. Analytical thin layer chromatography was performed on glass-backed SiliCycle® silica gel 60 Å plates and visualized by staining with 10% H2SO4 in ethanol and/or absorbance of UV light. Flash chromatography was performed using SiliCycle® SilicaFlash® P60 silica gel. All solvents were removed under reduced pressure using a rotary evaporator unless otherwise noted. 1H-NMR and 13C-NMR spectra were obtained using Bruker AV-500 or AV-600 instruments and have shifts recorded in ppm (δ), while coupling constants (J) are reported in hertz (Hz). High resolution electrospray ionization (HR ESI) mass spectra were obtained from the UC Berkeley Mass Spectrometry Facility. HPLC was performed on a Varian Pro Star or Varian Prep Star instrument using a C18 column.

UDP-2-XylAz synthetic procedures

Methyl D-lyxopyranoside (5).37 To a stirring solution of D-lyxose (600 mg, 4.0 mmol, 1.0 eq.) in 30 mL of dry MeOH was added dropwise acetyl chloride (0.60 mL, 8.5 mmol, 2.1 eq.). The reaction mixture was heated at 50 °C for 15 h, then cooled to 0 °C and neutralized by the addition of 0.5 g of solid Ag2CO3. Stirring was continued for 2 h at this temperature, at which time the reaction mixture was filtered through Celite® and concentrated in vacuo to furnish crude product 5. Crude 5 was purified by silica gel column chromatography (CH2Cl2:MeOH, 20:1 → 15:1 → 10:1) to afford 5 (640 mg, 3.9 mmol, 97%) as a mixture of anomers (α:β, 7:1). Compound 5α: 1H NMR (600 MHz, MeOD): δ 3.41 (dd, 1H, J = 10.9, 9.2 Hz), 3.63-3.66 (m, 2H), 3.74-3.75 (m, 1H), 3.76-3.80 (m, 1H), 4.56 (d, 1H, J = 2.8 Hz); 13C NMR (150 MHz, MeOD): δ 55.5, 64.0, 68.6, 71.6, 72.7, 103.1; Compound 5β: 1H NMR (600 MHz, MeOD): δ 3.25 (m, 1H), 3.45 (s, 3H), 3.59-3.61 (m, 1H), 3.74-3.75 (m, 1H), 3.85-3.86 (m, 1H), 3.94 (dd, 1H, J = 12.0, 3.5 Hz), 4.50 (d, 1H, J = 2.4 Hz); 13C NMR (150 MHz, MeOD): δ 56.6, 63.2, 69.1, 69.8, 73.8, 102.8; IR: 3367, 2926, 1445, 1132, 1055, 1007, 973 cm⁻¹; HRMS (ESI): Calcd. for C6H12O5Na [M+Na]+ 187.0577, found 187.0576.
3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-methyl-α-D-lyxopyranoside (6). To a solution of 2,2,3,3-tetramethoxybutane (210 mg, 1.2 mmol, 1.2 eq.) and 5 (160 mg, 1.0 mmol, 1.0 eq.) in 5 mL of dry MeOH was added trimethyl orthoformate (0.44 mL, 4.0 mmol, 1.2 eq.). To this solution was added camphorsulfonic acid (12 mg, 0.050 mmol, 0.050 eq.) and the reaction mixture was heated to reflux for 16 h. After cooling to room temperature, the reaction mixture was treated with solid 0.5 g of NaHCO₃, filtered and concentrated. The crude compound was then purified by silica gel column chromatography (hexanes:EtOAc, 4:1 → 3:1 → 2:1) to afford 6 (150 mg, 0.52 mmol, 52%). ¹H NMR (600 MHz, CDCl₃): δ 1.27 (s, 3H), 1.31 (s, 3H), 2.80 (s, 1H), 3.23 (s, 3H), 3.26 (s, 3H), 3.35 (s, 3H), 3.58-3.65 (m, 2H), 3.90-3.92 (m, 2H), 4.13-4.16 (m, 1H), 4.65 (s, 1H); ¹³C NMR (150 MHz, CDCl₃): δ 17.7, 17.7, 47.9, 48.0, 54.9, 60.4, 62.8, 68.6, 69.6, 99.7, 100.5, 101.1; IR: 3457, 2992, 2948, 2834, 1378, 1124, 1050, 875, 773 cm⁻¹; HRMS (ESI): Calcd. for C₁₂H₂₂O₇Na [M+Na]⁺ 301.1258, found 301.1263.

2-O-Triflate-3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-methyl-α-D-lyxopyranoside (7). To a stirring solution of 6 (670 mg, 2.4 mmol, 1.0 eq.) and pyridine (0.39 mL, 4.8 mmol, 2.0 eq.) in 8.5 mL of CH₂Cl₂ at -20 °C was added trimethanesulfonic anhydride (0.49 mL, 2.9 mmol, 1.2 eq.) dropwise. The mixture was allowed to stir for 0.5 h, and was then concentrated in vacuo. Purification with silica gel column chromatography (hexanes:EtOAc, 20:1 → 10:1 → 5:1) afforded 7 (550 mg, 1.3 mmol, 55%). ¹H NMR (600 MHz, CDCl₃): δ 1.26 (s, 3H), 1.27 (s, 3H), 3.25 (s, 3H), 3.26 (s, 3H), 3.40 (s, 3H), 3.60-3.70 (m, 2H), 4.07-4.12 (m, 2H), 4.80 (s, 1H), 4.88 (s, 1H); ¹³C NMR (150 MHz, CDCl₃): δ 17.7, 17.7, 47.9, 48.0, 54.9, 60.4, 62.8, 68.6, 69.6, 98.9, 99.9, 100.7; IR: 2924, 2852, 1415, 1210, 1145, 1070, 997, 773 cm⁻¹; HRMS (ESI): Calcd. for C₁₃H₂₂O₇F₃SNa [M+Na]⁺ 433.0751, found 433.0765.

2-Azido-2-deoxy-3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-methyl-α-D-xylopyranoside (8). To a stirring solution of 7 (100 mg, 0.25 mmol, 1.0 eq.) in 0.95 mL
of dry DMF was added lithium azide (49 mg, 1.0 mmol, 4.0 eq.). This mixture was stirred for 14 h at 80 °C. After cooling to ambient temperature, the reaction mixture was diluted with CH₂Cl₂ and washed once with 10 mL of sat. NaHCO₃ (aq.) and then 10 mL of brine. Lastly the mixture was dried over 1 g of Mg₂SO₄, filtered and concentrated in vacuo. Purification with silica gel column chromatography (hexanes:EtOAc, 20:1 → 10:1 → 5:1) afforded 8 (53 mg, 0.17 mmol, 70%). ¹H NMR (600 MHz, CDCl₃): δ 1.30 (s, 3H), 1.35 (s, 3H), 3.28 (s, 3H), 3.35 (s, 3H), 3.41 (s, 3H), 3.58 (dd, 1H, J = 5.4, 10.5 Hz), 3.66 (dd, 1H, J₁ = J₂ = 10.8 Hz), 3.83–3.78 (m, 1H), 4.18 (dd, 1H, J = 9.6, 10.7 Hz), 4.73 (d, 1H, J = 3.6 Hz); ¹³C NMR (150 MHz, CDCl₃): δ 17.6, 17.8, 48.1, 48.4, 55.3, 59.7, 60.3, 66.8, 68.1, 99.3, 99.9, 100.3; IR: 2993, 2950, 2837, 2105, 1467, 1144, 1044, 665 cm⁻¹; HRMS (ESI): Calcd. for C₁₂H₂₁O₆N₃Li [M+Li]⁺ 310.1585, found 310.1586.

1,2,3,4-tetra-O-acetyl-D-lyxopyranoside (10). To a stirring solution of D-lyxose, (2.2 g, 15 mmol, 1.0 eq.) in 36 mL of anhydrous pyridine at 0 °C, acetic anhydride (18 mL, 190 mmol, 13 eq.) was added dropwise. After stirring for 14 h, the reaction mixture was diluted with 20 mL of water and concentrated in vacuo. The crude product was purified by silica gel column chromatography (hexanes:EtOAc, 10:1 → 5:1 → 2:1), affording 10 (quant.) as a mixture of anomers. Compound 10α: ¹H NMR (600 MHz, CDCl₃): δ 5.98 (d, 1H, J = 2.8 Hz), 5.35–5.34 (m, 1H), 5.20–5.19 (m, 1H), 5.02–4.99 (m, 1H), 4.16 (dd, 1H, J = 12.6, 3.4 Hz), 2.09 (s, 3H), 2.08 (s, 3H), 2.06 (s, 6H); ¹³C NMR (150 MHz, CDCl₃): δ 169.7, 169.7, 169.4, 168.8, 89.5, 67.6, 67.5, 66.1, 61.7, 20.7, 20.7, 20.6, 20.5; Compound 10β: ¹H NMR (600 MHz, CDCl₃): δ 5.85 (d, 1H, J = 3.3 Hz), 5.22 (dd, 1H, J = 9.1, 3.5 Hz), 5.11–5.10 (m, 1H), 5.08–5.04 (m, 1H), 3.87 (dd, 1H, J = 11.5, 5.0 Hz), 3.58 (dd, 1H, J = 11.5, 9.0 Hz), 2.03 (s, 3H), 2.01 (s, 3H), 1.95 (s, 3H), 1.93 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 169.4, 169.4, 169.2, 168.1, 90.2, 67.9, 66.1, 61.5, 59.9, 13.8; IR: 2986, 1748, 1372, 1224, 1054 cm⁻¹; HRMS (ESI): Calcd. for C₁₃H₁₈O₉Na [M+Na]⁺ 341.0854, found 341.0857.

Phenyl 2,3,4-tri-O-acetyl-1-thio-D-lyxopyranoside (11). After thiophenol (0.50 mL, 4.9 mmol, 1.5 eq.) was added to a stirring solution of 10 (1.0 g, 3.3 mmol, 1.0 eq.) in 10 mL of DCM at 0 °C, boron trifluoride diethyl etherate (0.84 mL, 6.5 mmol, 2.0 eq.) was added dropwise. The reaction mixture was allowed to slowly warm to ambient temperature as it stirred for 18 h. After quenching the reaction mixture with 10 mL of sat. NaHCO₃ (aq.), and washing with 10 mL of sat. NaCl, the organic layer was extracted with DCM and concentrated in vacuo. The crude product was then purified using silica column chromatography (hexanes:EtOAc, 20:1 → 10:1 → 5:1), affording 11 (1.1 g, 2.8
mmol, 87%) as a mixture of anomers. Compound 11α: 1H NMR (600 MHz, CDCl₃): δ 7.49–7.47 (m, 2H), 7.34–7.28 (m, 3H), 5.59–5.58 (m, 1H), 5.21–5.11 (m, 3H), 4.38 (dd, 1H, J = 12.2, 3.7 Hz), 3.52–3.48 (m, 1H), 2.16 (s, 3H), 2.10 (s, 3H), 2.08 (s, 3H); Compound 11β: 1H NMR (600 MHz, CDCl₃): δ 7.49–7.47 (m, 2H), 7.33–7.30 (m, 3H), 5.41–5.40 (m, 1H), 5.33–5.30 (m, 2H), 5.17–5.16 (m, 1H), 4.07 (dd, 1H, J = 11.8, 8.0 Hz), 3.96 (dd, 1H, J = 11.8, 4.7 Hz), 2.12 (s, 3H), 2.09 (s, 3H), 2.07 (s, 3H); Compound 11α/β: 13C NMR (151 MHz, CDCl₃): δ 169.90, 169.89, 169.73, 169.66, 169.62, 169.54, 134.25, 132.83, 131.99, 131.68, 129.06, 129.03, 127.90, 127.71, 86.05, 85.40, 69.82, 69.39, 69.02, 68.36, 67.36, 67.31, 63.48, 61.95, 20.78, 20.77, 20.74, 20.73, 20.66, 20.60; IR: 3059, 2980, 2940, 2876, 2360, 2341, 1750, 1224, 1060, 745 cm⁻¹; HRMS (ESI): Calcd. for C₁₇H₂₀O₇Sn [M+Na]⁺ 391.0822, found 391.0823.

Phenyl thio-D-lyxopyranoside (12). To a stirring solution of 11 (5.2 g, 14 mmol, 1.0 eq.) in 140 mL of dry MeOH, 25% NaOMe in MeOH (0.32 mL, 1.4 mmol, 0.10 eq.) was added. This reaction mixture was allowed to stir at rt for 2 h, at which time it was neutralized to pH 7 with Dowex 50WX8-400 ion-exchange resin. Concentrating this solution in vacuo afforded 12 (3.2 g, 13 mmol, 95%) as a mixture of anomers, which was used without further purification. Compound 12α: 1H NMR (600 MHz, CDCl₃): δ 7.46–7.43 (m, 2H), 7.27–7.22 (m, 1H), 7.22–7.18 (m, 2H), 5.02 (s, 1H), 4.09–4.08 (m, 1H), 4.07–4.04 (m, 1H), 3.83–3.79 (m, 1H), 3.53 (dd, 1H, J = 8.1, 3.2 Hz), 3.22 (dd, 1H, J = 11.5, 8.4 Hz); 13C NMR (150 MHz, CDCl₃): δ 137.5, 131.5, 129.9, 127.9, 90.1, 75.1, 73.5, 72.3, 68.5; Compound 12β: 1H NMR (600 MHz, CDCl₃): δ 7.49–7.47 (m, 2H), 7.32–7.29 (m, 2H), 7.27–7.24 (m, 1H), 5.28 (d, 1H, J = 4.1 Hz), 4.00–3.98 (m, 1H), 3.85–3.82 (m, 2H), 3.78–3.75 (m, 1H), 3.72–3.71 (m, 1H); 13C NMR (150 MHz, CDCl₃): δ 135.8, 132.5, 130.0, 128.3, 89.8, 72.9, 72.0, 69.1, 65.9; IR: 3367, 2929, 1641, 737 cm⁻¹; HRMS (ESI): Calcd. for C₁₁H₁₄O₄Sn [M+Na]⁺ 265.0505, found 265.0508.

Phenyl 3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-thio-D-lyxopyranoside (13). To a solution of 12 (5.8 g, 24 mmol, 1.0 eq.) and 2,2,3,3-tetramethoxybutane (5.1 g, 29 mmol, 1.2 eq.) in 100 mL of dry MeOH was added trimethyl orthoformate (11 mL, 94 mmol, 4.0 eq.). Camphorsulfonic acid (280 mg, 1.2 mmol, 0.050 eq.) was added and the reaction mixture was heated to reflux for 16 h. After cooling the reaction mixture to 0 °C, it was treated with 0.5 g of solid NaHCO₃, filtered and concentrated in vacuo to give a crude yellow oil. Addition of MeOH to the oil precipitated pure β anomer 13 (1.73 g, 4.85 mmol, 38%); the yield was calculated from the amount of the β starting material. 1H
NMR (500 MHz, CDCl₃): δ 7.52–7.51 (m, 2H), 7.33–7.27 (m, 3H), 4.84 (s, 1H), 4.25 (ddd, 1H, J₁ = J₂ = 10.4, J₃ = 4.9 Hz), 4.20 (m, 1H), 4.03 (dd, 1H, J = 10.8, 5.0 Hz), 3.69 (dd, 1H, J₁ = J₂ = 10.8 Hz), 3.26 (s, 6H), 2.49 (s, 1H), 1.34 (s, 3H), 1.28 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 134.4, 131.2, 128.9, 127.3, 100.5, 99.7, 88.2, 71.6, 70.8, 68.0, 62.4, 48.0, 47.8, 17.6, 17.5; IR: 3477, 3070, 2994, 2931, 2868, 2831, 1585, 1380, 1125, 1041, 746 cm⁻¹; HRMS (ESI): Calcd. for C₁₇H₂₄O₆SNa [M+Na]⁺ 379.1186, found 379.1190.

Phenyl 2-azido-2-deoxy-3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-thio-D-xylopyranoside (14). To a stirring solution of 13 (40 mg, 0.11 mmol, 1.0 eq.) and anhydrous pyridine (18 mL, 0.22 mmol, 2.0 eq.) in 0.5 mL of DCM at -20 °C was added trimethanesulfonic anhydride (23 mL, 0.14 mmol, 1.2 eq.) dropwise. After stirring for 3.5 h, solvent was removed in vacuo. After dissolving the remaining reaction mixture in 0.5 mL of DMF, lithium azide (22 mg, 0.45 mmol, 4.0 eq.) was added and the solution was allowed to stir at rt for 1.5 h. The crude product was diluted with DCM, washed with 1 mL of sat. NaHCO₃ (aq.), washed with 2 mL of brine, dried over 1 g of solid MgSO₄, filtered and concentrated in vacuo, affording 14 (37 mg, 0.096 mmol, 85%), was used without further purification. ¹H NMR (600 MHz, CDCl₃): δ 7.57–7.56 (m, 2H), 7.34–7.33 (m, 3H), 4.34 (d, 1H, J = 9.8 Hz), 3.97 (dd, 1H, J = 10.9, 4.8 Hz), 3.71 (ddd, 1H, J₁ = J₂ = 10.2, J₃ = 4.7 Hz), 3.67–3.64 (m, 1H), 3.41–3.36 (m, 2H), 3.32 (s, 3H), 3.25 (s, 3H), 1.33 (s, 3H), 1.27 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 134.0, 130.8, 129.0, 128.6, 100.3, 99.7, 87.2, 73.6, 67.9, 65.7, 61.4, 48.1, 48.0, 17.6, 17.5; IR: 2992, 2919, 2835, 2110, 1474, 1036, 882, 749 cm⁻¹; HRMS (ESI): Calcd. for C₁₇H₂₃O₅N₃SLi [M+Li]⁺ 388.1513, found 388.1511.

2-Azido-2-deoxy-3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-D-xylose (15). To a stirring solution of 14 (110 mg, 0.30 mmol, 1.0 eq.) in 4 mL of water and 20 mL of acetone was added N-bromosuccinimide (120 mg, 2.4 mmol, 8.0 eq.), resulting in a yellow solution. Once the solution became colorless, additional N-bromosuccinimide (120 mg, 2.4 mmol, 8.0 eq.) was added. The reaction mixture was allowed to stir until it became colorless again, at which time it was quenched with 10 mL of sat. NaHCO₃, extracted with CH₂Cl₂, and dried over 1 g of solid Mg₂SO₄, filtered and concentrated in vacuo. After purification with silica gel column chromatography (hexanes:EtOAc, 10:1 → 5:1), 15 (65 mg, 0.22 mmol, 75%) was obtained as a mixture of anomers (α:β, 1.2:1).
Compound 15α: $^1$H NMR (600 MHz, CDCl$_3$): $\delta$ 1.29 (s, 3H), 1.34 (s, 3H), 3.26 (s, 3H), 3.35 (s, 3H), 3.36-3.40 (m, 1H), 3.54-3.60 (m, 1H), 3.76-3.81 (m, 1H), 3.86-3.92 (m, 1H), 4.19-4.22 (m, 1H), 5.24-5.25 (m, 1H); $^{13}$C NMR (150 MHz, CDCl$_3$): $\delta$ 17.5, 17.7, 48.0, 48.5, 59.8, 64.1, 66.7, 67.9, 92.7, 99.8, 100.3; Compound 15β: $^1$H NMR (600 MHz, CDCl$_3$): $\delta$ 1.27 (s, 3H), 1.34 (s, 3H), 3.26 (s, 3H), 3.30 (s, 3H), 3.36-3.40 (m, 2H), 3.54-3.60 (m, 1H), 3.76-3.81 (m, 1H), 3.86-3.92 (m, 1H), 4.19-4.22 (m, 1H), 5.24-5.25 (m, 1H); $^{13}$C NMR (150 MHz, CDCl$_3$): $\delta$ 17.4, 17.6, 48.3, 53.4, 60.6, 63.8, 65.9, 71.3, 97.3, 99.7, 100.1; IR: 3425, 2951, 2836, 2109, 1377, 1145, 1036, 935 cm$^{-1}$; HRMS (ESI): Calcd. for C$_{11}$H$_{19}$O$_6$N$_3$Li $[M+Li]^+$ 296.1428, found 296.1429.

2-Azido-2-deoxy-3,4-O-(2',3'-dimethoxy-2',3'-dimethylbutane)-D-xylose-diallylphosphate (16). To a stirring solution of 15 (5.0 mg, 0.017 mmol, 1.0 eq.) and 1H-tetrazole (6.1 mg, 0.087 mmol, 5.0 eq.) in 1 mL of CH$_2$Cl$_2$ at ambient temperature was added diallyl diisopropylphosphoramidite (14 µL, 0.054 mmol, 3.1 eq.) dropwise. After stirring for 5 min, the mixture was brought to -40 °C and mCPBA (14.9 mg, 0.0865 mmol, 5.0 eq.) was added. This was allowed to stir for 10 min and was then transferred to an ice bath where it was allowed to warm to room temperature over 1 h. The reaction mixture was then diluted with 5 mL of CH$_2$Cl$_2$, washed once with 5 mL of 10% Na$_2$SO$_4$ (aq.), once with 5 mL of sat. NaHCO$_3$ (aq.), and once with 5 mL of water. After drying the organic layer over 1 g of solid Na$_2$SO$_4$, it was filtered and concentrated in vacuo. The resulting crude material was purified using silica gel column chromatography (hexanes:EtOAc, 10:1 → 5:1 → 3:1), affording 16 (7.3 mg, 0.016 mmol, 94%) as a mixture of anomers (α:β, 1:1.5). Compound 16α: $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 1.27 (s, 3H), 1.33 (s, 3H), 3.26 (s, 3H), 3.33 (s, 3H), 3.50-3.53 (m, 1H), 3.71 (m, 1H), 4.12-4.13 (m, 1H), 4.56-4.60 (m, 4H), 5.23-5.27 (m, 2H), 5.34-5.38 (m, 2H), 5.68-5.70 (m, 1H), 5.90-5.95 (m, 2H); Compound 16β: $^1$H NMR (500 MHz, CDCl$_3$): $\delta$ 1.29 (s, 3H), 1.34 (s, 3H), 3.52 (s, 3H), 3.29 (s, 3H), 3.43-3.46 (m, 1H), 3.51 (dd, 1H, J = 10.6, 7.8 Hz), 3.60 (m, 1H), 3.80 (dd, 1H, J$_1$ = J$_2$ = 10.3, J$_3$ = 5.0 Hz), 3.93 (dd, 1H, J = 11.2, 5.1 Hz), 4.56-4.60 (m, 4H), 4.95 (m, 1H), 5.23-5.27 (m, 2H), 5.34-5.38 (m, 2H), 5.90-5.95 (m, 2H); $^{13}$C NMR (150 MHz, CDCl$_3$): $\delta$ 17.4, 17.6, 48.0, 48.1, 63.0 (d, J = 8.9 Hz), 64.8, 65.5, 68.2 (t, J = 3.9 Hz), 71.2 (d, J = 1.5 Hz), 98.1 (d, J = 5.1 Hz), 99.7, 100.2, 118.4 (d, J = 10.4 Hz), 132.1 (dd, J = 14.3, 7.4 Hz); IR: 2952, 2111, 1378, 1277, 1139, 1027, 958, 732 cm$^{-1}$; HRMS (ESI): Calcd. for C$_{17}$H$_{28}$O$_9$N$_3$PNa $[M+Na]^+$ 472.1455, found 472.1460.
Phenyl 2-azido-2-deoxy-3,4-di-O-acetyl-1-thio-β-D-xylopyranoside (17). To a stirring solution of 9:1 TFA:H₂O (30 mL) was added 14 (1.9 g, 4.9 mmol, 1.0 eq.). After stirring at rt for 15 min, the reaction mixture was coevaporated with toluene, and dissolved in 10 mL of anhydrous pyridine. The solution was then cooled to 0 °C and acetic anhydride (5 mL, 53 mmol, 11 eq.) was added dropwise. After stirring for 17 h, the reaction was diluted with 5 mL of H₂O, concentrated in vacuo and purified using silica gel column chromatography (hexanes:EtOAc, 20:1 → 10:1 → 5:1) affording 17 (1.63 g, 4.64 mmol, 95%). ¹H NMR (400 MHz, CDCl₃): δ 7.58–7.56 (m, 2H), 7.36–7.35 (m, 3H), 5.07 (t, J = 9.2 Hz, 1H), 4.87 (td, J = 9.5, 5.3 Hz, 1H), 4.50 (d, J = 9.5 Hz, 1H), 4.17 (dd, J = 11.6, 5.3 Hz, 1H), 3.38–3.33 (m, 2H), 2.09 (s, 3H), 2.07 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 169.9, 169.7, 133.8, 130.6, 129.1, 128.7, 86.6, 73.4, 68.5, 66.2, 62.3, 20.7, 20.6; IR: 3062, 2925, 2868, 2109, 1755, 1745 cm⁻¹; HRMS (ESI): Calcd. for C₁₅H₁₇O₅N₃SLi [M+Li]⁺ 358.1043, found 358.1040.

2-azido-2-deoxy-3,4-di-O-acetyl-β-D-xylopyranoside (18). To a stirring solution of 17 (1.6 g, 4.6 mmol, 1 eq.) in 275 mL of acetone:H₂O (5:1), N-bromosuccinimide (4.6 g, 93 mmol, 20 eq.) was added in 3 equal portions over the course of 2 h. After the reaction mixture was allowed to stir at rt for 1 h, it was washed with 100 mL of sat. NaHCO₃ (aq.), extracted with 300 mL of DCM, dried over 3 g of solid MgSO₄, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to afford 18 (1.11 g, 4.29 mmol, 92%) as a mixture of anomers. Compound 18α: ¹H NMR (600 MHz, CDCl₃): δ 5.52 (t, J = 9.6 Hz, 1H), 5.33 (t, J = 3.6 Hz, 1H), 4.97–4.93 (m, 1H), 3.90 (t, J = 10.7 Hz, 1H), 3.82 (dd, J = 11.1, 5.8 Hz, 1H), 3.37–3.32 (m, 1H), 2.92 (dd, J = 3.9, 1.2 Hz, 1H), 2.12 (s, 3H), 2.05 (s, 3H); Compound 18β: ¹H NMR (600 MHz, CDCl₃): δ 5.03 (t, J = 9.7 Hz, 1H), 4.97–4.93 (m, 1H), 4.68 (dd, J = 7.8, 5.1 Hz, 1H), 4.09 (dd, J = 11.7, 5.5 Hz, 1H), 3.42 (dd, J = 10.1, 7.8 Hz, 1H), 3.40 (d, J = 5.1 Hz, 1H), 3.37–3.32 (m, 1H), 2.12 (s, 3H), 2.04 (s, 3H); Compound 18α/β: ¹³C NMR (150 MHz, CDCl₃): δ 170.4, 170.3, 170.2, 96.4, 91.8, 71.9, 69.7, 69.2, 68.9, 64.5, 62.4, 61.2, 58.2, 20.4, 20.4, 20.4; IR: 3446, 2953, 2360, 2111, 1750, 1369, 1240, 1053 cm⁻¹; HRMS (ESI): Calcd. for C₉H₁₃O₆N₅Li [M+Li]⁺ 266.0959, found 266.0955.
2-azido-2-deoxy-3,4-di-O-acetyl-1-(diallylphosphate)-α-D-xylopyranoside (19). 1H-tetrazole (180 mg, 2.6 mmol, 5.0 eq.) and 18 (130 mg, 0.52 mmol, 1.0 eq.) were suspended in 5 mL of dry toluene, and sonicated for 1 h to ensure that large clumps of 1H-tetrazole were broken into fine particles so it could dissolve readily in the reaction solvent. After removing toluene, the reaction mixture was dissolved in 5 mL of DCM, and allowed to stir for a few minutes. Diallyl N,N-diisopropylphosphoramidite (0.42 mL, 1.6 mmol, 3.1 eq.) was then added dropwise, and the solution was allowed to stir at rt for 45 min. After cooling to −40 °C, in a MeCN/CO₂(s.) bath, mCPBA (450 mg, 2.6 mmol, 5.0 eq.) was added and the reaction was allowed to stir for 30 min. The reaction mixture was then diluted with 100 mL of DCM and washed sequentially with 100 mL of 10% Na₂SO₃ (aq.), 100 mL of sat. NaHCO₃ (aq.), and 100 mL of water. After drying the DCM layer over 5 g of Na₂SO₄ (s.), the crude product was concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to afford the desired product 19 (63 mg, 0.150 mmol, 29%). ¹H NMR (600 MHz, CDCl₃): δ 5.99–5.87 (m, 2H), 5.76 (dd, J = 6.4, 3.3 Hz, 1H), 5.43 (t, J = 9.9 Hz, 1H), 5.37 (ddq, J = 17.1, 7.0, 1.5 Hz, 2H), 5.29–5.23 (m, 2H), 4.98 (ddd, J = 10.8, 9.4, 5.9 Hz, 1H), 4.64–4.54 (m, 4H), 3.90 (ddd, J = 11.2, 5.9 Hz, 1H), 3.79 (t, J = 11.0 Hz, 1H), 3.54 (dt, J = 10.4, 3.1 Hz, 1H), 2.10 (s, 3H), 2.02 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 169.8, 169.7, 132.0 (d, J = 7.3), 132.0 (d, J = 5.2 Hz), 118.7, 118.5, 95.5 (d, J = 5.5 Hz), 69.8, 68.7 (d, J = 5.4 Hz), 68.5, 68.5 (d, J = 5.4 Hz), 61.2 (d, J = 8.9 Hz), 60.2, 20.6, 20.6; IR: 2956, 2899, 2112, 1755, 1369, 1237, 1025, 958 cm⁻¹; HRMS (ESI): Calcd. for C₁₅H₂₂O₉N₃PNa [M+Na]⁺ 442.0986, found 442.0981.

Uridine 5'-monophosphate di-triethylammonium salt (19a). HNEt₃⁺-DOWEX-50WX8 resin was generated by stirring 10 g of H⁺-DOWEX-50WX8 overnight in 750 mL of 5% NEt₃·HCl, then washing several times with water. After suspending the resin in 65 mL of water, 1 g of uridine 5'-monophosphate disodium salt was added and the reaction mixture was allowed to stir overnight. After filtering off the resin, the sample was frozen using dry ice and lyophilized to give 19a.
UDP-2-XylAz (1). To a 1 mL stirring solution of THF:MeOH (1:1) was added 19 (20 mg, 0.048 mmol, 1 eq.), sodium p-toluensulfinate (16 mg, 0.093 mmol, 2.0 eq.), and tetrakis(triphenylphosphine)-palladium(0) (2.8 mg, 0.0024 mmol, 0.05 eq.). After allowing the reaction mixture to stir at rt for 1.5 h, it was concentrated and coevaporated with dry toluene (3x). The crude material was then resuspended in 0.35 mL of MeCN, 20 pellets of 4 angstrom molecular sieves were added and the reaction mixture was cooled to 0 °C. In a separate flask, trifluoroacetic anhydride (40 µL, 0.29 mmol, 6.0 eq.) in 0.10 mL MeCN was cooled to 0 °C and added dropwise to the flask containing 19a. After allowing the reaction to stir at room temperature for 5 minutes, N₂ was blown over the solution for 5-10 min to remove TFA and TFAA, and the reaction mixture was cooled to 0 °C. In a separate flask, N-methylimidazole (14 µL, 0.17 mmol, 3.6 eq.) and NEt₃ (40 µL, 0.29 mmol, 6.0 eq.) were dissolved in 0.1 mL MeCN and cooled to 0 °C, then added to the flask containing 19a. Stirring for 5 to 10 minutes at 0 °C provided the activated UMP, which was then added dropwise to the flask containing 19. After allowing for 3 h at 0 °C, the reaction mixture was concentrated in vacuo, resuspended in 2 mL of MeOH:H₂O:NEt₃ (5:2:1) and allowed to stir for 18 h at rt. The crude product was then purified by HPLC, using a 250 X 10 mm C-18 column and 10 mM tributylammonium bicarbonate [A] and MeOH [B] as the mobile phases. 4 L of A were prepared by bubbling CO₂ through a solution of 10 mM tributylamine in water until all tributylamine dissolved into the water. Once the crude product was dissolved in 0.5 mL of A, it was injected onto the HPLC and purified under the following conditions: 0 → 10 min [100:0, A:B], 10 → 35 min [70:30, A:B], 35 → 60 min [50:50, A:B]. The pure product was then washed with DCM to remove excess tributylammonium bicarbonate (3X) and then purified by size exclusion chromatography using a p2 column. Finally, the purified compound was passed through a sodium ion exchange resin, and concentrated in vacuo to give 1 (7.5 mg, 0.023 mmol, 26%). ¹H NMR (500 MHz, D₂O): δ 7.96 (d, J = 8.1 Hz, 1H), 5.98–5.96 (m, 2H), 5.56 (dd, J = 7.2, 3.4 Hz, 1H), 4.37–4.34 (m, 2H), 4.28–4.24 (m, 2H), 4.20–4.17 (m, 1H), 3.82 (dd, J = 11.2, 5.6 Hz, 1H), 3.76–3.67 (m, 2H), 3.61–3.55 (m, 2H); ¹³C NMR (150 MHz, D₂O): δ 166.3, 151.8, 141.5, 102.6, 95.6 (d, J = 7.0 Hz), 88.5, 83.1 (d, J = 9.5 Hz), 73.8, 71.8, 71.7 (d, J = 8.4 Hz), 69.4, 64.7 (d, J = 5.1 Hz), 60.9, 60.5; ³¹P NMR (243 MHz, D₂O): δ –10.51 (d, J = 21.0 Hz), –12.30 (d, J = 20.7 Hz); HRMS (ESI): Calcd. for C₁₄H₂₀O₁₅N₅P₂ [M–H]⁻ 560.0437, found 560.0441.
**UDP-3-XylAz synthetic procedures**

5-O-TBDPS-1,2-O-isopropylidene-α-D-xylofuranose (21). To a solution of TBDPSCI (3.59 mL, 13.8 mmol, 1.31 eq.) in 50 mL of DMF at 0 °C was added 1,2-O-isopropylidene-α-D-xylofuranose (20, 2.00 g, 10.5 mmol, 1.0 eq.) and imidazole (1.79 g, 29.8 mmol, 2.8 eq.). After stirring for 10 min at 0 °C, the reaction mixture was warmed to rt and allowed to stir for 1 h. The reaction mixture was poured into 50 mL of ice water and diluted with 100 mL of EtOAc. After extracting the organic layer, it was washed sequentially with 100 mL of sat. NH₄Cl, 100 mL of sat. NaHCO₃, and 100 mL of brine. The resulting product was then dried over solid Na₂SO₄, filtered, concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 9:1 → 6:1 → 4:1) to afford pure 21 (quant.). ¹H NMR (600 MHz, CDCl₃): δ 7.72 (dt, J = 6.7, 1.5 Hz, 2H), 7.68 (dt, J = 6.7, 1.5 Hz, 2H), 7.47–7.43 (m, 2H), 7.43–7.36 (m, 4H), 6.01 (d, J = 3.7 Hz, 1H), 4.55 (d, J = 3.7 Hz, 1H), 4.38 (t, J = 2.6 Hz, 1H), 4.16–4.10 (m, 3H), 4.08 (d, J = 3.1 Hz, 1H), 1.47 (s, 3H), 1.33 (s, 3H), 1.05 (s, 9H).

5-O-TBDPS-1,2-O-isopropylidene-α-D-xylo-3-ulose (22). To a stirring solution of oxalyl chloride (1.06 mL, 12.5 mmol, 1.1 eq.) in 30 mL of dry DCM at −60 °C (CHCl₃/CO₂(s.)), was added DMSO (1.76 mL, 25.0 mmol, 2.2 eq.) slowly. A solution of 21 (4.87 g, 11.4 mmol, 1.0 eq.) in 10 mL of dry DCM was then added dropwise over 5 min, where it was allowed to continue stirring at −60 °C for 0.5 h. After the addition of NEt₃ (5.2 mL, 37.5 mmol, 3.3 eq.) in 8.5 mL of dry DCM, the mixture was allowed to stir at −60 °C for an additional 1 h. The reaction mixture was then allowed to warm to rt and stir 2.5 h, at which point it was poured into 50 mL of water and extracted with 50 mL of CHCl₃ (2X). After washing the organic layer with brine and drying it over solid Na₂SO₄, 22 was obtained (4.71 g, 11.0 mmol, 97%) and used without further purification. ¹H NMR (600 MHz, CDCl₃): δ 7.69 (dd, J = 6.7, 1.6 Hz, 2H), 7.62 (dd, J = 6.8, 1.5 Hz, 2H), 7.47–7.36 (m, 6H), 6.27 (d, J = 4.5 Hz, 1H), 4.44 (d, J = 4.5, 1H), 4.40 (s, 1H), 3.92 (dd, J = 11.1, 1.9 Hz, 1H), 3.87 (dd, J = 11.1, 2.2 Hz, 1H), 1.49 (s, 3H), 1.48 (s, 3H), 1.02 (s, 9H).
5-O-TBDPS-1,2-O-isopropylidene-α-D-ribofuranose (23). After dissolving 22 (4.71 g, 11.0 mmol, 1 eq.) in 90 mL of ethanol:water (3:1), and cooling to 0 °C, sodium borohydride (2.71 g, 71.8 mmol, 6.5 eq.) was added in four equivalents over 0.5 h. The reaction mixture was then allowed to stir for 3.5 h, at which point it was poured into 180 mL of water and extracted with EtOAc. After drying the organic layer over solid Na$_2$SO$_4$, concentrating in vacuo, and purifying by silica gel column chromatography (hexanes:EtOAc, 15:1 → 10:1 → 5:1), the desired product, 23 (3.17 g, 7.39 mmol, 67%), was obtained. $^1$H NMR (500 MHz, CDCl$_3$): δ 7.76–7.70 (m, 4H), 7.46–7.36 (m, 6H), 5.87 (d, $J = 3.8$ Hz, 1H), 4.61 (dd, $J = 5.1$, 3.8 Hz, 1H), 4.20–4.15 (m, 1H), 4.00–3.99 (m, 1H), 3.93–3.86 (m, 2H), 2.46 (dd, $J = 10.0$, 4.6 Hz, 3H), 1.58 (s, 3H), 1.40 (s, 3H), 1.09 (s, 9H).

3-azido-3-deoxy-5-O-TBDPS-1,2-O-isopropylidene-α-D-xylofuranose (24). To a stirring solution of 23 (3.168 g, 7.39 mmol, 1 eq.) in 130 mL of DCM was added anhydrous pyridine (1.2 mL, 14.8 mmol, 2 eq.). After cooling the reaction to −20 °C, trifluoromethanesulfonic anhydride (1.5 mL, 8.87 mmol, 1.2 eq.) was added dropwise. After stirring for 20 min, the reaction was concentrated in vacuo and resuspended in 130 mL of DMF. Lithium azide (1.45 g, 29.6 mmol, 4 eq.) was then added to the reaction mixture and it was allowed to stir for 7.5 h at rt. After washing the crude product with sat. NaHCO$_3$ and brine, it was dried over MgSO$_4$ (s.), concentrated in vacuo and purified by silica gel column chromatography (hexanes:EtOAc, 20:1 → 10:1 → 5:1) to give pure 24 (2.37 g, 5.22 mmol, 71%). $^1$H NMR (600 MHz, CDCl$_3$): δ 7.67 (dd, $J = 14.6$, 6.6, Hz, 4H), 7.46–7.35 (m, 6H), 5.83 (d, $J = 3.6$ Hz, 1H), 4.62 (d, $J = 3.7$ Hz, 1H), 4.32–4.29 (m, 1H), 4.09 (d, $J = 3.2$ Hz, 1H), 3.91 (dd, $J = 10.3$, 5.3 Hz, 1H), 3.85 (dd, $J = 10.3$, 8.5 Hz, 1H), 1.49 (s, 3H), 1.32 (s, 3H), 1.07 (s, 9H); $^{13}$C NMR (151 MHz, CDCl$_3$): δ 135.6, 135.5, 129.9, 129.8, 127.8, 112.2, 104.8, 83.4, 79.2, 66.3, 61.2, 26.8, 26.6, 26.3, 19.2; IR: 3072, 3050, 2959, 2933, 2890, 2858, 2107, 1428 1113, 1021 cm$^{-1}$; HRMS (ESI): Calcd. for C$_{24}$H$_{31}$O$_{4}$N$_{3}$SiNa [M+Na]$^+$ 476.1976, found 476.1975.
3-azido-3-deoxy-1,2-O-isopropylidene-α-D-xylofuranose (25). To a stirring solution of 24 (2.34 g, 5.15 mmol, 1 eq.) in 75 mL of THF was added TBAF·3H₂O (2.53 g, 8.03 mmol, 1.56 eq.) in five equal portions over the course 30 min. After stirring for 1 h at rt, the reaction mixture was concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to afford 25 (1.00 g, 4.67 mmol, 91%).

**1H NMR** (600 MHz, CDCl₃): δ 5.93 (d, J = 3.7 Hz, 1H), 4.68 (d, J = 3.7 Hz, 1H), 4.37–4.35 (m, 1H), 4.02 (d, J = 3.3 Hz, 1H), 3.93 (dd, J = 11.5, 6.3 Hz, 1H), 3.88–3.81 (m, 1H), 1.52 (s, 3H), 1.34 (s, 3H); **13C NMR** (150 MHz, CDCl₃): δ 112.3, 104.7, 83.6, 79.4, 79.4, 66.0, 60.8, 26.6, 26.3; IR: 3426, 2989, 2939, 2108, 1377, 1216, 1018 cm⁻¹; HRMS (ESI): Calcd. for C₈H₁₃O₄N₃Li [M+Li]⁺ 222.1061, found 222.1059.

3-azido-3-deoxy-1,2,4-tri-O-acetyl-D-xylopyranoside (26). After dissolving 25 in 0 °C TFA/water (9:1) and letting it stir at this temperature for 0.5 h, the reaction mixture was coevaporated with toluene (3X). The reaction mixture was then dissolved in 10 mL anhydrous pyridine and cooled to 0 °C. After dropwise addition of 5 mL of acetic anhydride, the reaction mixture was allowed to stir for 15 h during gradual warming to rt. The reaction mixture was next diluted with 5 mL of water, concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 10:1 → 5:1 → 3:1) to give pure 26 (1.10 g, 3.66 mmol, 78%) as a mixture of anomers. Compound 26α: **1H NMR** (600 MHz, CDCl₃): δ 6.20 (d, J = 3.5 Hz, 1H), 4.88–4.80 (m, 2H), 3.94 (t, J = 10.3 Hz, 1H), 3.90 (dd, J = 11.1, 5.8 Hz, 1H), 3.59 (t, J = 10.9 Hz, 1H), 2.15 (s, 3H), 2.10 (s, 3H), 2.07 (s, 3H); **13C NMR** (150 MHz, CDCl₃): δ 169.4, 169.4, 168.7, 88.6, 69.9, 69.0, 60.5, 60.3, 20.7, 20.6, 20.4; Compound 26β: **1H NMR** (600 MHz, CDCl₃): δ 5.60 (d, J = 7.4 Hz, 1H), 4.92 (dd, J = 9.6, 7.4 Hz, 1H), 4.84–4.80 (m, 1H), 4.11 (dd, J = 11.9, 5.2 Hz, 1H), 3.71 (t, J = 9.4 Hz, 1H), 3.41 (dd, J = 11.9, 9.0 Hz, 1H), 2.08 (s, 6H), 2.06 (s, 3H); **13C NMR** (150 MHz, CDCl₃): δ 169.5, 169.0, 168.9, 92.2, 69.6, 69.1, 63.6, 62.7, 20.6, 20.6, 20.5; IR: 2919, 2851, 2107, 1765, 1759, 1755, 1745, 1738, 1728, 1368, 1215 cm⁻¹; HRMS (ESI): Calcd. for C₁₁H₁₆O₇N₃Li [M+Li]⁺ 308.1065, found 308.1063.

3-azido-3-deoxy-2,4-di-O-acetyl-D-xylopyranoside (27). To a stirring solution of 26α (264 mg, 0.876 mmol, 1 eq.) in 10 mL of THF was added benzylamine (115 µL, 1.05
mmol, 1.2 eq.). Once the reaction mixture was warmed to 80 °C it was allowed to stir for 18 h. After determining that a large amount of starting material still remained, additional benzylamine (115 µL, 1.05 mmol, 1.2 eq.) was added and the reaction mixture was allowed to continue stirring at 80 °C for an additional 7 h. The reaction mixture was then quenched with 1M HCl, extracted with DCM, and concentrated in vacuo. The desired product was then purified using silica gel column chromatography (hexanes:EtOAc, 10:1 → 2:1) to afford 27 (187 mg, 0.721 mmol, 82%). Major anomer: 1H NMR (600 MHz, CDCl3): δ 5.32 (bs, 1H), 4.75 (q, J = 9.6 Hz, 1H), 4.64 (dd, J = 10.4, 2.8 Hz, 1H), 4.01 (t, J = 10.2 Hz, 1H), 3.84 (s, 1H), 3.76 (s, 1H), 3.74 (d, J = 3.7 Hz, 1H), 2.13 (s, 3H), 2.09 (s, 3H); 13C NMR (150 MHz, CDCl3): δ 170.3, 170.0, 89.4, 71.9, 69.7, 60.0, 58.1, 20.7, 20.7; IR: 3457, 2956, 2895, 2502, 2256, 2109, 1751, 1228, 1052 cm⁻¹; HRMS (ESI): Calcd. for C₉H₁₃O₆N₃Na [M+Na]⁺ 282.0697, found 282.0693.

3-azido-3-deoxy-2,4-di-O-acetyl-1-(diallylphosphate)-α-D-xylopyranoside (28). 1H-tetrazole (163 mg, 2.33 mmol, 5 eq.) and 27 (121 mg, 0.465 mmol, 1 eq.) were suspended in 5 mL of dry toluene, and sonicated for 1 h to ensure that large clumps of 1H-tetrazole were broken into fine particles so it could dissolve readily in the reaction solvent. After removing toluene, the reaction mixture was dissolved in 3 mL of DCM, and allowed to stir for a few minutes. Diallyl N,N-diisopropylphosphoramidite (0.18 mL, 0.697 mmol, 1.5 eq.) was then added dropwise and the reaction mixture was allowed to stir for 15 min at rt. After cooling to −40 °C, mCPBA (401 mg, 2.33 mmol, 5 eq.) was added and the reaction mixture was allowed to stir for an additional 30 min at this temperature. The reaction mixture was then diluted with DCM and washed sequentially with 10% Na₂SO₃, sat. NaHCO₃, and water. Finally, the DCM layer was dried over Na₂SO₄ (s.), concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to give the desired product 28 (125 mg, 0.298 mmol, 64%). 1H NMR (600 MHz, CDCl₃): δ 6.01–5.89 (m, 2H), 5.78 (dd, J = 6.7, 3.3 Hz, 1H), 5.42–5.36 (m, 2H), 5.32–5.27 (m, 2H), 4.85 (td, J = 10.5, 5.8 Hz, 1H), 4.76 (dt, J = 10.6, 3.0 Hz, 1H), 4.62–4.56 (m, 4H), 3.99 (t, J = 10.3 Hz, 1H), 3.94 (dd, J = 11.1, 5.8 Hz, 1H), 3.72 (t, J = 11.0 Hz, 1H), 2.14 (s, 3H), 2.12 (s, 3H); 13C NMR (150 MHz, CDCl₃): δ 169.6, 169.5, 132.0 (d, J = 6.6), 132.0 (d, J = 6.3 Hz), 118.8, 93.4 (d, J = 5.4 Hz), 70.7 (d, J = 7.5 Hz), 69.0, 68.6 (d, J = 5.4), 68.6 (d, J = 4.0 Hz), 60.0, 59.9, 20.6, 20.6; IR: 2955, 2110, 1755, 1369, 1268, 1219, 1041, 958 cm⁻¹; HRMS (ESI): Calcd. for C₁₅H₂₂O₉N₃PNa [M+Na]⁺ 442.0986, found 442.0977.
UDP-3-XylAz (2). To a 6.5 mL stirring solution of THF:MeOH (1:1) was added 28 (100 mg, 0.24 mmol, 1 eq.), sodium p-toluenesulfinate (85 mg, 0.48 mmol, 2.0 eq.), and tetrakis(triphenylphosphine)-palladium(0) (14 mg, 0.012 mmol, 0.05 eq.). After allowing the reaction mixture to stir at rt for 1.5 h, it was concentrated and coevaporated with dry toluene (3x). The crude material was then resuspended in 1.5 mL of MeCN, 20 pellets of 4 angstrom molecular sieves were added and the reaction mixture was cooled to 0 °C. In a separate flask, after dissolving 19a (150 mg, 0.29 mmol, 1.2 eq.) in 1.5 mL of dry MeCN under N\textsubscript{2}, N,N-dimethylaniline (0.15 mL, 1.1 mmol, 4.8 eq.) and NEt\textsubscript{3} (40 µL, 0.29 mmol, 1.2 eq.) were added and the reaction mixture was brought to 0 °C. In a separate flask, trifluoroacetic anhydride (0.20 mL, 1.4 mmol, 6.0 eq.) in 0.50 mL MeCN was cooled to 0 °C and added dropwise to the flask with 19a. After allowing the reaction to stir at room temperature for 5 minutes, N\textsubscript{2} was blown over the solution for 5-10 min to remove TFA and TFAA, and the reaction mixture was cooled to 0 °C. In a separate flask, N-methylimidazole (70 µL, 0.86 mmol, 3.6 eq.) and NEt\textsubscript{3} (0.20 mL, 1.4 mmol, 6.0 eq.) were dissolved in 0.5 mL MeCN and cooled to 0 °C, then added to the flask containing 19a. Stirring for 5 to 10 minutes at 0 °C provided the activated UMP, which was then added dropwise to the flask containing 28. After stirring for 3 h at 0 °C, the reaction mixture was concentrated in vacuo, resuspended in 13 mL of MeOH:H\textsubscript{2}O:NEt\textsubscript{3} (5:2:1) and allowed to stir for 18 h at rt. The crude product was then concentrated in vacuo, redissolved in water, washed with DCM, and concentrated again. The crude product was then purified by HPLC, using a 250 X 10 mm C-18 column and 10 mM tributylammonium bicarbonate [A] and MeOH [B] as the mobile phases. 4 L of A were prepared by bubbling CO\textsubscript{2} through a solution of 10 mM tributylamine in water until all tributylamine dissolved into the water. Once the crude product was dissolved in 0.5 mL of A, it was injected onto the HPLC and purified under the following conditions: 0 → 10 min [100:0, A:B], 10 → 35 min [70:30, A:B], 35 → 60 min [50:50, A:B]. The pure product was then washed with DCM to remove excess tributylammonium bicarbonate (3X) and then purified by size exclusion chromatography using a p2 column. Finally, the purified compound was passed through a sodium ion exchange resin, and concentrated in vacuo to give 2 (12 mg, 0.020 mmol, 8.4%). \textsuperscript{1}H NMR (600 MHz, D\textsubscript{2}O): δ 7.96 (d, J = 8.1 Hz, 1H), 5.98–5.96 (m, 2H), 5.53 (dd, J = 7.1, 3.3 Hz, 1H), 4.38–4.35 (m, 2H), 4.28–4.18 (m, 3H), 3.76–3.75 (m, 2H), 3.70–3.56 (m, 3H). \textsuperscript{13}C NMR (150 MHz, D\textsubscript{2}O): δ 166.3, 151.8, 141.6, 102.7, 94.8 (d, J = 6.6 Hz), 88.5, 83.2 (d, J = 9.0 Hz), 73.8, 70.6 (d, J = 8.7 Hz), 69.6, 67.9, 65.9, 64.9, 62.1; \textsuperscript{31}P NMR (243 MHz, D\textsubscript{2}O): δ −10.85 (d, J = 20.7 Hz), −12.68 (d, J = 20.5 Hz); HRMS (ESI): Calcd. for C\textsubscript{14}H\textsubscript{20}O\textsubscript{15}N\textsubscript{5}P\textsubscript{2} [M–H]\textsuperscript{−} 560.0437, found 560.0431.
**UDP-4-XylAz synthetic procedures**

**1,2,3-tri-O-benzoyl-L-arabinopyranoside (29).** After allowing L-arabinose (10.0 g, 66.6 mmol, 1 eq.) to fully dissolve in 750 mL of anhydrous pyridine by stirring for 4 h, the reaction mixture was cooled to –35 °C, and benzoyl chloride (24 mL, 206 mmol, 3.1 eq.) was added dropwise over 30 min. After holding the temperature at –35 °C for 2 h, the reaction was allowed to slowly warm to rt for 15 h. Most of the pyridine was then removed in vacuo, and the resulting crude mixture was dissolved in CHCl₃ and washed sequentially with 1M HCl, sat. NaHCO₃, and water. After drying the organic layer over MgSO₄(s.) and concentrating in vacuo, the desired product was purified by silica gel column chromatography (hexanes:ether, 3:1) to give 29 (8.81 g, 19.1 mmol, 29%).

\[
\begin{align*}
&\text{1H NMR (600 MHz, CDCl₃): } \delta 8.13-8.08 \text{ (m, 2H), } 8.05-8.00 \text{ (m, 2H), } 7.89-7.86 \text{ (m, 2H), } 7.66-7.60 \text{ (m, 1H), } 7.58-7.39 \text{ (m, 6H), } 7.33-7.26 \text{ (m, 2H), } 6.76 \text{ (d, } J = 3.6 \text{ Hz, 1H), } 6.01 \text{ (dd, } J = 10.5, 3.2 \text{ Hz, 1H), } 4.49-4.48 \text{ (m, 1H), } 4.27 \text{ (d, } J = 12.8, 1H), 4.05 \text{ (dd, } J = 12.9, 2.2 \text{ Hz, 1H), } 2.34-2.30 \text{ (m, 1H); } \text{13C NMR (150 MHz, CDCl₃): } \delta 165.9, 165.5, 164.7, 133.7, 133.6, 133.3, 129.9, 129.8, 129.7, 129.3, 129.0, 128.9, 128.7, 128.6, 128.4, 91.3, 70.8, 67.8, 67.3, 64.7; \text{ IR: 3502, 3064, 2937, 2362, 2342, 1729, 1266, 1112, 1007, 709 cm}^{-1}; \text{ HRMS (ESI): Calcd. for C}_{26}H₂₂O₈Cl [M+Cl]^{-} 497.1003, \text{ found 497.1001.}
\end{align*}
\]

**4-azido-4-deoxy-1,2,3-tri-O-benzoyl-D-xylopyranoside (30).** To a stirring solution of 29 (8.81g, 19.06 mmol, 1 eq.) in 500 mL of DCM under an atmosphere of N₂ (g.) was added anhydrous pyridine (3.08 mL, 38.1 mmol, 2 eq.). After cooling to –20 °C, trifluoromethanesulfonic anhydride (3.84 mL, 22.9 mmol, 1.2 eq.) was added dropwise and the reaction mixture was allowed to stir 1 h at this temperature. The reaction mixture was then concentrated in vacuo and dissolved in 100 mL of DMF. After the addition of lithium azide (3.73 g, 76.2 mmol, 4 eq.), the reaction mixture was stirred for 2.5 h at rt. The reaction mixture was then diluted with 300 mL of DCM, washed with sat. NaHCO₃ and brine, dried over Na₂SO₄(s.), and concentrated in vacuo. Finally the crude product was purified by silica gel column chromatography (hexanes:EtOAc, 10:1 → 5:1) to afford 30 (7.94g, 16.3 mmol, 85%).

\[
\begin{align*}
&\text{1H NMR (600 MHz, CDCl₃): } \delta 8.11 \text{ (d, } J = 7.1 \text{ Hz, 2H), } 8.03 \text{ (d, } J = 7.2 \text{ Hz, 2H), } 7.88 \text{ (d, } J = 7.1 \text{ Hz, 2H), } 7.64 \text{ (t, } J = 7.5 \text{ Hz, 1H), } 7.56-7.50 \text{ (m, 3H), } 7.46 \text{ (t, } J = 7.5 \text{ Hz, 1H), } 7.42 \text{ (t, } J = 7.7 \text{ Hz, 2H), } 7.30 \text{ (t, } J = 7.7 \text{ Hz, 2H), } 6.71 \text{ (d, } J = 3.6 \text{ Hz, 1H), } 6.02 \text{ (t, } J = 9.7 \text{ Hz, 1H), } 5.51 \text{ (dd, } J = 10.2, 3.6 \text{ Hz, 1H), } 4.12-4.04 \text{ (m, 2H), } 3.94-3.87 \text{ (m, 1H); } \text{13C NMR (150 MHz, CDCl₃): } \delta 165.6, 165.4, 164.5, 133.9, 133.6, 133.5, 129.9, 129.8, 128.9, 128.8, 128.7, 128.5, 128.4, 90.2, 71.0,
\end{align*}
\]
4-azido-4-deoxy-2,3-di-O-benzoyl-D-xylopyranoside (31). To a stirring solution of 30 (36 mg, 0.074 mmol, 1.0 eq.) at 0 °C in 1.7 mL of THF was added acetic acid (6 µL, 0.10 mmol, 1.4 eq.) followed by ethylenediamine (6 µL, 0.089 mmol, 1.2 eq.). This mixture was allowed to slowly warm to rt and stir for 48 h. The reaction mixture was then quenched with water and the product was extracted with DCM. The DCM layer was washed sequentially with 1 M HCl, sat. NaHCO₃, and water. After drying with Na₂SO₄ (s.), the crude product was purified by silica column chromatography (hexanes:EtOAc, 10:1 → 5:1) to give 31 (12 mg, 0.032 mmol, 43%). Significant starting material (31) was recovered (19 mg, 0.039 mmol, 52%). Major anomer: ¹H NMR (600 MHz, CDCl₃): δ 7.98 (dd, J = 7.5 Hz, 4H), 7.53 (q, J = 8.2 Hz, 2H), 7.41-7.36 (m, 4H), 5.92 (t, J = 9.5 Hz, 1H), 5.62-5.61 (m, 1H), 5.17 (dd, J = 10.1, 3.5 Hz, 1H), 4.02-3.98 (m, 1H), 3.91-3.87 (m, 2H); ¹³C NMR (150 MHz, CDCl₃): δ 166.8, 165.9, 165.7, 165.6, 133.6, 133.6, 133.5, 133.4, 129.9, 129.9, 129.8, 129.7, 129.1, 128.8, 128.7, 128.6, 128.5, 128.4, 128.4, 96.3, 90.6, 73.9, 72.9, 72.2, 70.6, 63.9, 60.5, 59.8, 59.5, 59.5, 29.7, 14.1; HRMS (ESI): Calcd. for C₁₉H₁₇O₆N₃Na [M+Na]⁺ 406.1010, found 406.1011.

4-azido-4-deoxy-2,4-di-O-benzoyl-1-(diallylphosphate)-D-xylopyranoside (32). ¹H-tetrazole (11 mg, 0.16 mmol, 5 eq.) and 31 (12 mg, 0.032 mmol, 1 eq.) were suspended in 1 mL of dry toluene, and sonicated for 1 h to ensure that large clumps of ¹H-tetrazole were broken into fine particles so it could dissolve readily in the reaction solvent. After removing toluene, the reaction mixture was dissolved in 0.5 mL of DCM, and allowed to stir for a few minutes. Diallyl N,N-diisopropylphosphoramidite (25 µL, 0.097 mmol, 3.0 eq.) was then added dropwise and the reaction mixture was allowed to stir for 15 min at rt. After cooling to −40 °C, mCPBA (29 mg, 0.16 mmol, 5 eq.) was added and the reaction mixture was allowed to stir for an additional 1 h at this temperature. The reaction mixture was then diluted with DCM and washed sequentially with 10% Na₂SO₄, sat. NaHCO₃, and water. Finally, the DCM layer was dried over Na₂SO₄ (s.), concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to give 32 (7.9 mg, 0.015 mmol, 45%). Major anomer: HRMS (ESI): Calcd. for C₂₅H₂₆O₉N₃NaP [M+Na]⁺ 566.1294, found 566.1299.
4-azido-4-deoxy-1,2,3-tri-O-acetyl-D-xylopyranoside (33). To a stirring solution of 30 (7.94 g, 16.3 mmol, 1 eq.) in 250 mL of anhydrous MeOH under N₂, was added 25% MeONa in MeOH (350 µL, 1.63 mmol, 0.1 eq.). After stirring for a 3 h at rt, 5 still remained so additional 25% MeONa in MeOH (350 µL, 1.63 mmol, 0.1 eq.) was added and the reaction mixture was allowed to continue stirring for 1 h. After the reaction was neutralized with DOWEX 50WX8-400 ion-exchange resin and filtered, all MeOH was removed in vacuo, and the resulting crude mixture was dissolved in 120 mL of anhydrous pyridine. The reaction mixture was then cooled to 0 °C, and acetic anhydride (60 mL, 0.636 mol, 39 eq.) was added dropwise. The reaction mixture was then allowed to stir for 13 h as it slowly warmed to rt. After diluting the reaction with 60 mL of water and concentrating in vacuo, the crude material was dissolved in CHCl₃, washed with 1M HCl and water, and dried over MgSO₄(s). Finally, the crude product was concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 10:1 → 5:1) to give 33 (3.90 g, 12.9 mmol, 80%) as a mixture of anomers. Compound 33α: ¹H NMR (600 MHz, CDCl₃): δ 6.18 (d, J = 3.7 Hz, 1H), 5.33 (t, J = 9.9 Hz, 1H), 4.95 (dd, J = 10.3, 3.9 Hz, 1H), 3.83 (dd, J = 11.4, 5.5 Hz, 1H), 3.75–3.66 (m, 1H), 3.60 (t, J = 11.3 Hz, 1H), 2.12 (s, 3H), 2.07 (s, 3H), 1.97 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 169.7, 169.7, 168.8, 89.2, 70.3, 69.3, 61.4, 58.9, 20.7, 20.5, 20.2; Compound 33β: ¹H NMR (600 MHz, CDCl₃): δ 5.61 (d, J = 7.7 Hz, 1H), 5.10 (t, J = 9.1 Hz, 1H), 4.98–4.96 (m, 1H), 4.05 (dd, J = 12.0, 5.2 Hz, 1H), 3.76–3.66 (m, 1zH), 3.40 (dd, J = 12.1, 10.3 Hz, 1H), 2.07 (s, 3H), 2.07 (s, 3H), 2.01 (s, 3H); ¹³C NMR (150 MHz, CDCl₃): δ 169.8, 169.4, 168.9, 92.1, 72.9, 70.0, 63.9, 58.4, 20.8, 20.6, 20.4; IR: 2957, 2925, 2854, 2109, 1755, 1370, 1211, 1071, 937 cm⁻¹; HRMS (ESI): Calcd. for C₁₁H₁₅O₇N₃Na [M+Na]⁺ 324.0802, found 324.0798.

4-azido-4-deoxy-2,3-di-O-acetyl-D-xylopyranoside (34). To a stirring solution of 33 (3.90 g, 12.9 mmol, 1 eq.) in 65 mL of DMF was added hydrazine acetate (1.31 g, 14.2 mmol, 1.1 eq.). After stirring for 1 h at rt, the reaction mixture was diluted with 100 mL of EtOAc, and washed with sat. NaCl and water. The crude reaction mixture was then dried over Na₂SO₄ (s.), concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to afford 34 (1.80 g, 6.96 mmol, 54%) as a mixture of anomers. Compound 34α: ¹H NMR (600 MHz, CDCl₃): δ 5.32 (t, J = 9.8 Hz, 1H), 5.25 (d, J = 3.5 Hz, 1H), 4.72–4.67 (m, 1H), 3.72 (t, J = 11.3 Hz, 1H), 3.67–3.62 (m, 1H), 3.60–3.52 (m, 1H), 3.60 (s, 3H), 1.98 (s, 3H); Compound 34β: ¹H NMR (600 MHz, CDCl₃): δ 5.00 (t, J = 9.5 Hz, 1H), 4.72–4.67 (m, 1H), 4.59 (d, J = 7.8 Hz, 1H), 3.96 (dd, J = 12.0, 5.4 Hz, 1H), 3.67–3.62 (m, 1H), 3.22 (t, J = 11.4 Hz, 1H), 2.01 (s, 3H), 1.97 (s, 3H); Compound 34α/β: ¹³C NMR (150 MHz, CDCl₃): δ 170.4, 170.0, 95.4,
4-azido-4-deoxy-2,3-di-O-acetyl-1-(diallylphosphate)-α-D-xylopyranoside (35). 1H-tetrazole (3.38 g, 48.2 mmol, 5 eq.) and 34 (2.50 g, 9.64 mmol, 1 eq.) were suspended in 100 mL of dry toluene, and sonicated for 1 h to ensure that large clumps of 1H-tetrazole were broken into fine particles so it could dissolve readily in the reaction solvent. After removing toluene, the reaction mixture was dissolved in 95 mL of DCM, and allowed to stir for a few minutes. Diallyl N,N-diisopropylphosphoramidite (7.87 mL, 29.9 mmol, 3.1 eq.) was then added dropwise and the reaction mixture was allowed to stir for 45 min at rt. After cooling to –40 °C, mCPBA (8.32 g, 48.2 mmol, 5 eq.) was added and the reaction mixture was allowed to stir for an additional 30 min at this temperature. The reaction mixture was then diluted with DCM and washed sequentially with 10% Na2SO3, sat. NaHCO3, and water. Finally, the DCM layer was dried over Na2SO4 (s.), concentrated in vacuo, and purified by silica gel column chromatography (hexanes:EtOAc, 5:1 → 2:1) to give the desired product 35 (1.25 g, 2.97 mmol, 31%).

1H NMR (600 MHz, CDCl3): δ 5.97–5.87 (m, 2H), 5.78 (dd, J = 7.0, 3.3 Hz, 1H), 5.41–5.31 (m, 3H), 5.29–5.24 (m, 2H), 4.88 (dt, J = 10.1, 2.9 Hz, 1H), 4.60–4.50 (m, 4H), 3.91–3.82 (m, 1H), 3.75–3.67 (m, 2H), 2.09 (s, 3H), 2.03 (s, 3H); 13C NMR (150 MHz, CDCl3): δ 169.6, 169.5, 131.8 (d, J = 7.1 Hz), 118.6, 118.4, 94.0 (d, J = 4.9 Hz), 69.9 (d, J = 7.2 Hz), 69.7, 68.5 (d, J = 5.6 Hz), 68.3 (d, J = 5.4 Hz), 61.0, 58.7, 20.5, 20.4; IR: 2922, 2851, 2110, 1755, 1370, 1233, 1025, 957 cm–1; HRMS (ESI): Calcd. for C15H22O9N3PNa [M+Na]+ 442.0986, found 442.0979.

UDP-4-XylAz (3). To a 6.5 mL stirring solution of THF:MeOH (1:1) was added 35 (105 mg, 0.25 mmol, 1 eq.), sodium p-toluenesulfinate (89 mg, 0.50 mmol, 2.0 eq.), and tetrakis(triphenylphosphine)-palladium(0) (14 mg, 0.0125 mmol, 0.05 eq.). After allowing the reaction mixture to stir at rt for 1.5 h, it was concentrated and coevaporated with dry toluene (3x). The crude material was then resuspended in 1.5 mL of MeCN, 20 pellets of 4 angstrom molecular sieves were added and the reaction mixture was cooled to 0 °C. In a separate flask, after dissolving 19a (158 mg, 0.30 mmol, 1.2 eq.) in 1.5 mL of dry MeCN under N2, N,N-dimethylaniline (0.15 mL, 1.2 mmol, 4.8 eq.) and NEt3 (40 µL, 0.30 mmol, 1.2 eq.) were added and the reaction mixture was brought to 0 °C. In a separate flask, trifluoroacetic anhydride (0.21 mL, 1.5 mmol, 6.0 eq.) in 0.5 mL MeCN
was cooled to 0 °C and added dropwise to the flask with 19a. After allowing the reaction to stir at room temperature for 5 minutes, N₂ was blown over the solution for 5 min to remove TFA and TFAA, and the reaction mixture was cooled to 0 °C. In a separate flask, N-methylimidazole (72 µL, 0.90 mmol, 3.6 eq.) and NEt₃ (0.21 mL, 1.5 mmol, 6.0 eq.) were dissolved in 0.5 mL MeCN and cooled to 0 °C, then added to the flask containing 19a. Stirring for 5 to 10 minutes at 0 °C provided the activated UMP, which was then added dropwise to the flask containing 35. After stirring for 3 h at 0 °C, the reaction mixture was concentrated in vacuo, resuspended in 13 mL of MeOH:H₂O:NEt₃ (5:2:1) and allowed to stir for 18 h at rt. The crude product was then concentrated in vacuo, redissolved in water, washed with DCM, and concentrated again. The crude product was then purified by HPLC, using a 250 X 10 mm C-18 column and 10 mM tributylammonium bicarbonate [A] and MeOH [B] as the mobile phases. 4 L of A were prepared by bubbling CO₂ through a solution of 10 mM tributylamine in water until all tributylamine dissolved into the water. Once the crude product was dissolved in 0.5 mL of A, it was injected onto the HPLC and purified under the following conditions: 0 → 10 min [100:0, A:B], 10 → 35 min [70:30, A:B], 35 → 60 min [50:50, A:B]. The pure product was then washed with DCM to remove excess tributylammonium bicarbonate (3X) and then purified by size exclusion chromatography using a p2 column. Finally, the purified compound was passed through a sodium ion exchange resin, and concentrated in vacuo to give 1 (32.8 mg, 0.0542 mmol, 22%). ¹H NMR (500 MHz, D₂O): δ 7.95 (d, J = 8.1 Hz, 1H), 6.00–5.93 (m, 2H), 5.61 (dd, J = 7.4, 3.3 Hz, 1H), 4.39–4.33 (m, 2H), 4.29–4.14 (m, 3H), 3.85 (dd, J = 10.3, 8.9 Hz, 1H), 3.80–3.70 (m, 2H), 3.70–3.62 (m, 1H), 3.33 (dt, J = 10.3, 3.0 Hz, 1H); ¹³C NMR (150 MHz, D₂O): δ 166.8, 152.4, 142.3, 103.2, 95.1 (d, J = 6.3 Hz), 89.0, 83.9 (d, J = 9.7 Hz), 74.4, 71.4, 70.3, 70.0, 65.5, 63.5 (d, J = 8.7 Hz), 62.9; ³¹P NMR (243 MHz, D₂O): δ −11.58 (d, J = 19.4 Hz), −13.55 (d, J = 19.1 Hz); HRMS (ESI): Calcd. for C₁₄H₂₀O₁₅N₅P₂ [M–H]⁻ 560.0437, found 560.0440.
References


Chapter 3

Probing Glycosaminoglycans in Developing Zebrafish Embryos
Chapter 3. Probing Glycosaminoglycans in Developing Zebrafish Embryos*

Introduction

The ability to deliver nucleotide sugars to cells could greatly increase the diversity of unnatural functionalities that can be expressed on cell surfaces. Unlike previous studies by our group, which used peracetylated sugars, negatively charged nucleotide sugars, such as the UDP-XylAz analogs previously described, will not passively diffuse across the cell membrane into the cytoplasm. Fortunately, at the outset of this work, our lab had recently discovered a feasible alternative to this problem; researchers showed that UDP-GalNAz could be microinjected into zebrafish embryo yolk sacs at the one-to-four cell stage and be taken up by the first cells of the developing organism. This method ensures that every subsequent cell of the developing embryo will contain the unnatural azido analog (Figure 3.1). However, this method only guarantees that the nucleotide sugar gains access to the cytoplasm of zebrafish cells. Notably, azide-dependent labeling of cell surface zebrafish glycans was observed when UDP-GalNAz was injected into zebrafish embryos, therefore providing evidence that unnatural nucleotide sugars are accepted by natural sugar transporters and can translocate from the cytoplasm to the Golgi. When applying this strategy to probe glycosaminoglycans (GAGs), it was hypothesized that after unnatural UDP-XylAz derivatives were delivered to all the cells of the embryo, they would be accepted by natural transporters, and thereby gain access to the Golgi. Once in the Golgi, if recognized by the xylosyltransferases 1 and 2 (XylT1 and XylT2), the unnatural UDP-XylAz analogs would be used to install XylAz into cell surface GAGs. In addition to providing a platform for delivering hydrophilic and charged nucleotide sugars to cells, zebrafish have many features that make them an ideal system for imaging GAGs.

Zebrafish have proven extremely valuable for developmental studies because of their ideal rate of development, optical translucency, and amenability to forward and reverse genetics. The ideal model organism should develop rapidly enough so that meaningful experiments can be performed in a timely manner. If the organism develops too rapidly it is extremely difficult or even impossible to observe developmental changes; however, if development is too slow, experiments become time consuming and impede research progress. After only one day post-fertilization, zebrafish embryos have a beating heart and circulating red blood cells. By four days, embryos have developed a functional liver, pancreas, and intestines. In addition to their superb time frame of development, the fact that zebrafish embryos can be kept translucent using small molecule additives renders them highly suitable for molecular imaging. This has proven to be particularly valuable for exploring the dynamic biological processes associated with development. Lastly, zebrafish is a genetically tractable organism, amenable to forward and reverse genetic manipulation. Forward genetics (i.e. after embryos are exposed to mutagens, a particular phenotype can be correlated to its genetic association) has been used to probe gene function in zebrafish for several years. While reverse genetics has been available using morpholino oligonucleotides

Figure 3.1: Hypothesized scheme for metabolic labeling of glycosaminoglycans. UDP-XylAz is microinjected into zebrafish embryos at the 1-4 cell stage and actively transported into the cytosol of the first cells of the embryo. Transporters deliver unnatural analogs to the Golgi lumen where they are utilized by xylosyltransferases to append XylAz to protein scaffolds. Depending on the position in which the azide is installed, GAG polymerization occurs or is inhibited and the resulting azide-modified biomolecule is transported to cell surfaces. Once present on the cell surface, azide-modified glycans can be detected through a Cu-free azide-alkyne cycloaddition reaction to reveal organismal locations of glycosaminoglycans.
for many years, reliable methods to generate specific mutant zebrafish lines have only recently become available. These features combined have made zebrafish the most utilized organism for studying vertebrate development. Because GAGs have essential functions in modulating signaling events during vertebrate development, primarily through their interactions with signaling proteins, zebrafish serve as a great platform for probing this type of glycosylation using unnatural sugars in combination with bioorthogonal chemistry.

Results and Discussion

UDP-2-XylAz and UDP-3-XylAz do not provide azide-dependent labeling

Both of the azido UDP-xylene derivatives that were hypothesized to label GAGs in a non- or only slightly perturbing manner (see Chapter 2 for details), UDP-2-XylAz and UDP-3-XylAz, were initially tested for their ability to deliver azidoxylose to zebrafish GAGs. In these experiments, 25–200 pmol of each analog was injected into zebrafish embryos at the 1–4 cell stage. After 24 hours post-fertilization (hpf), embryos were incubated with difluorocyclooctyne-AlexaFluor 488 (DIFO-488, 200 µM), mounted in agarose, and imaged by confocal microscopy (Figures 3.1 and 3.2). The range of doses injected was chosen based on the morphological effects on embryo development.

![Image](image_url)

Figure 3.2: Azide-dependent labeling was not observed in zebrafish embryos injected with UDP-2-XylAz or UDP-3-XylAz. Embryos were injected with each analog at the 1-4 cell stage, reacted with DIFO-488 at 24 hpf, and subsequently imaged. Top row, DIFO-488 channel; bottom row, brightfield channel. Left column, vehicle-injected; center column, UDP-2-XylAz-injected; right column, UDP-3-XylAz-injected.
and/or the extent of azide-dependent labeling; 25 pmol of each analog was the initial dose tested, as this is the amount at which peracetylated N-acetylgalactosamine (Ac₄GalNAz) provides robust azide-dependent labeling when microinjected into zebrafish embryo yolk sacs. This dose was increased until morphological effects were noted in hopes of observing azide-dependent labeling. Unfortunately, no azide-dependent signal was detected by microscopy when UDP-2-XylAz or UDP-3-XylAz were injected (Figure 3.2). Flow cytometry was used to verify the absence of labeling with these UDP-XylAz analogs as it enables quantification of fluorescence and detection of much more subtle differences in signal. It can detect signal that is only slightly greater than 1-fold over background, while fluorescence imaging can only reliably detect signal that is closer to 10-fold over background. In these experiments, zebrafish embryos at the 1–4 cell stage were injected with 100 pmol of UDP-2-XylAz or 200 pmol of UDP-3-XylAz, and allowed to develop to 24 hpf (Figure 3.3). After incubating in a 50 µM solution of NHS-647 to label the embryo surface cells through reaction with their solvent-exposed lysines, the embryos were washed and dissociated into single cells. Cells were then reacted with DIFO-biotin (200 µM, 1 h, 28 °C), washed, incubated in a 5 µg/mL solution of avidin-488 (15 min, 4 °C) twice, followed by flow cytometry analysis. This experiment verified that azide-dependent labeling was not observed with these compounds with a lower limit of detection (Figure 3.4).

**Figure 3.3:** Flow cytometry experimental setup that enables detection of azide-dependent labeling on zebrafish embryos. Embryos are injected with UDP-XylAz analogs and allowed 24 h to incorporate the unnatural sugar. Embryos are then reacted with NHS-647, which non-specifically reacts with lysine amines on embryo surface cells. After dissociation into single cells, DIFO-biotin is added and Cu-free click chemistry progresses with available azides, resulting in their modification with biotin groups. Avidin-488 is then added to enable detection of biotin-modified cells, and flow cytometry is used to quantify the level of azide dependent labeling.
Figure 3.4: Fluorescence signal of enveloping layer cells from zebrafish embryos injected with each UDP-XylAz analog. This data confirms that no azide-dependent labeling is observed with UDP-2-XylAz or UDP-3-XylAz. However, it is confirmed that UDP-4-XylAz provides azide dependent labeling.

Although azide-dependent labeling was not observed with UDP-2-XylAz or UDP-3-XylAz, UDP-2-XylAz had interesting effects that warrant further discussion. Only minor morphological effects were observed with UDP-3-XylAz when it was injected at high doses (>100 pmol), which suggests that it is simply not recognized by the XylTs. Interestingly, although azide-dependent labeling was not observed with UDP-2-XylAz, it profoundly affected development at lower doses (25-50 pmol, Figure 3.2). The fact that it affects development at such low doses suggests that UDP-2-XylAz has a specific effect. There are several possible explanations that could describe the observed effect. Perhaps, UDP-2-XylAz enters but does not exit from the Golgi transporter, preventing natural UDP-xylose from entering this subcellular compartment. It is also possible that UDP-2-XylAz is bound by one or both of the XylTs, but cannot be used as a substrate, and therefore inhibits enzyme function. Lastly, UDP-2-XylAz may in fact be utilized by one or both of the XylTs and incorporated into GAGs, but having an azide in place of the C-2 hydroxyl group prevents phosphorylation, which is required for modulating GAG expression (discussed in Chapter 2). While this third option is perhaps the most compelling, it is also the least likely because if the unnatural sugar were incorporated into cell surface GAGs, it should be detectable using cyclooctyne probes. However, it is also possible that the azide would be efficiently blocked from access by the cyclooctynes by the bulky GAG chains or that azide modification at the C-2 position of xylose would prevent canonical transport of the protein to the cell surface. Further investigations will be required to ascertain the fate and effects of UDP-2-XylAz following its injection into zebrafish embryos.
Figure 3.5: Azide-dependent labeling is observed for embryos injected with UDP-4-XylAz. Following injection with UDP-4-XylAz, embryos are given 24 h to incorporate the unnatural sugar into cell-surface glycans, and subsequently incubated with a DIFO-488 to detect azide-modified glycans on the surface of the embryo. Left panel, vehicle-injected embryos; right panel, UDP-4-XylAz-injected embryos; top row, DIFO-488 signal; bottom row, brightfield signal; Scale bars: 10x, 200 µm; 20x, 50 µm.

UDP-4-XylAz provides azide-dependent labeling

While no labeling was detected on embryos injected with UDP-2-XylAz or UDP-3-XylAz, azide-dependent labeling was observed for embryos injected with UDP-4-XylAz (Figure 3.5). In these experiments, 50 pmol of UDP-4-XylAz was injected into zebrafish embryos at the 1–4 cell stage. After 24 hours post-fertilization (hpf), embryos were incubated with DIFO-488 (200 µM), mounted in agarose, and imaged by confocal microscopy (Figures 3.1 and 3.5). Images reporting on the DIFO-488 channel clearly show that when UDP-4-XylAz is injected, signal is observed; by using a higher magnification lens (20x) it was clear that signal emanated from the proper cellular location, the cell surface. Flow cytometry, which provides more sensitive detection than fluorescence microscopy, confirmed this finding (Figure 3.4). Because the goal of this work at the outset of the project was to generate tools for probing glycosaminoglycans in their natural state, this at first seemed like a minor success. However, after discovering that tools for modulating GAG expression were quite limited, this became an exciting discovery. UDP-4-XylAz would not only enable attenuation of GAG expression, but bioorthogonal chemistry could be used to determine the organismal sites of inhibition (Figure 3.6). This feature can expedite research, as it allows investigators to rapidly confirm the inhibitory effects spatially and temporally.
Figure 3.6: UDP-4-XylAz functions as a visualizable inhibitor of GAG biosynthesis. After microinjection of UDP-4-XylAz into the zebrafish yolk sac at the 1-4 cell stage, 4-XylAz is transferred to the protein substrates of the XylTs and subsequently inhibits GAG biosynthesis, because the necessary C-4 hydroxyl group is replaced with an azide. Following their transport to the cell surface, 4-XylAz-modified proteins can be visualized using fluorophore-modified cyclooctyne probes, enabling facile discernment of the organismal sites of GAG inhibition.

**UDP-4-XylAz delivers 4-XylAz to sites of GAG modification**

Having confirmed that UDP-4-XylAz delivers azides to the cell surfaces of zebrafish embryos, it was important to next establish that the observed signal emanates from 4-XylAz present at sites of GAG glycosylation. To this end, we performed biochemical assays with human xylosyltransferases 1 and 2 (XylT1 and 2), which are known to initiate GAG glycosylation. In these experiments, an acceptor peptide was incubated with UDP-4-XylAz and either XylT1 or XylT2, after which the glycopeptide product was analyzed by mass spectrometry (Figure 3.7). More specifically, following its injection into the mass spectrometer, electrospray ionization is used to cleave the glycosidic bond, which results in concomitant loss of azidoxylose from the peptide. The mass transition corresponding to loss of azidoxylose is detected using the multiple reaction-monitoring mode (MRM) of the spectrometer and quantified to verify the presence or absence of the glycosylated peptide (Figure 3.8). While XylT1 transferred 4-XylAz to the acceptor peptide, XylT2 was inactive on the substrate (Figure 3.8A and 3.8C). As positive controls, analogous experiments were run in parallel using UDP-xylose in place of UDP-4-XylAz; as was expected, both XylT1 and XylT2 readily transferred xylose to the peptide substrate (Figure 3.8). Further, a negative control was
The acceptor peptide was incubated with UDP-4-XylAz and one of the XylTs. After removing impurities, the sample containing the potentially xylosylated peptide was injected into a mass spectrometer and analyzed for the presence of the modified peptide using MRM monitoring.

performed with heat killed XylT1, which predictably resulted in no glycosylation (Figure 3.8B). As human XylT1 is highly homologous to its zebrafish counterpart, sharing 75% identity, this result supports the claim that XylT1 is able to transfer UDP-4-XylAz onto proteoglycans in zebrafish.²⁰,²¹

Having validated that XylT1 can transfer UDP-4-XylAz to an acceptor peptide in vitro, we disrupted translation of XylT1 in zebrafish using a previously reported splice-blocking morpholino oligonucleotide (MO) to study the effect on 4-XylAz labeling.²² The XylT1 MO is a splice-blocking MO that functions by disrupting canonical splicing of the XylT1 pre-mRNA, and results in exclusion of exon 2, the largest XylT mRNA exon. Because the mature mRNA differs in size when the MO is effective, the MO’s success can be determined using RT-PCR. To confirm the success of the MO and determine an optimal dosage for disrupting XylT1 expression, 0, 2, or 6 ng of the XylT1 MO were microinjected in zebrafish embryos at the 1–4 cell stage. At 24 hpf, total RNA was isolated from MO-injected embryos and controls, purified, and used to generate total cDNA. RT-PCR was then performed to detect the presence or absence of exon 2, and thus the effectiveness of the MO (Figure 3.9). Oligonucleotide primers were designed to amplify exon 2 and short segments of the surrounding introns (Figure 3.10). From control-injected embryos, we detected a 851 bp product, the expected size for the normal exon 2-containing transcript amplicon (inclusion product), whereas from embryos injected with 2 or 6 ng of XylT1 MO, we observed a dose-dependent decrease in the 851 bp amplicon, and a concomitant increase in a 328 bp amplicon, the expected
Figure 3.8: XylT1 can use UDP-4-XylAz as a substrate while XylT2 cannot. While this mass transition was observed when the peptide was incubated with XylT1 (A), it was not observed when incubated with heat-killed XylT1 (B) or XylT2 (C). As the glycosylated, or unglycosylated, peptide pass through the spectrometer, their presence, or absence, is detected by the presence, or absence, of the mass transition (top right corner of each chromatogram) corresponding to loss of XylAz from the peptide. Note: As a positive control for both XylT1 and XylT2 activity, an analogous experiment to that described above was performed in which UDP-xylose was used in place of UDP-4-XylAz; both enzymes readily transferred xylose to the acceptor peptide (data not shown).

After assessing the effectiveness of the XylT1 MO, it was used to show that UDP-4-XylAz is incorporated into zebrafish proteoglycans. To this end, embryos were injected with 50 pmol of UDP-4-XylAz alone in the presence or absence of 6 ng of XylT1 MO and were allowed to develop to 24 hpf. Enveloping layer cells were labeled with NHS-647 and the embryos were dissociated into single cells, incubated with DIFO-biotin
Figure 3.9: XylT1 MO at a dose of 6 ng effectively inhibits the production of XylT1 mRNA. Embryos were injected with 0, 2, or 6 ng of XylT1 MO, and a cDNA library was generated from their mRNA. Using RT-PCR, the targeted exon was amplified to analyze the effect of the MO. While 2 ng of the XylT1 MO attenuated production of the inclusion product, 6 ng essentially eliminated the inclusion product.

Table 3.1: XylT and β-actin primers used to amplify mRNA sequences.

<table>
<thead>
<tr>
<th>Primer name</th>
<th>Primer sequence (5’ → 3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>XylT1 forward</td>
<td>ATCGGCATCAACAAAGCCTA</td>
</tr>
<tr>
<td>XylT1 reverse</td>
<td>GGATACTGGTGAGCGAGAGC</td>
</tr>
<tr>
<td>β-actin forward</td>
<td>GACATCAAGGAGAAGCTGTGC</td>
</tr>
<tr>
<td>β-actin reverse</td>
<td>GAGGAGGGCAAAACTGGTAAAC</td>
</tr>
</tbody>
</table>

followed by avidin-AlexaFluor-488, and analyzed by flow cytometry. Knockdown of XylT1 activity by MO treatment reduced UDP-4-XylAz-dependent fluorescence labeling by 30% (Figure 3.10), thus confirming that the observed signal from UDP-4-XylAz injection reports on incorporation onto proteoglycans. This partial decrease in labeling was expected because embryos have maternally provided XylT activity that cannot be affected by the XylT1 MO.\textsuperscript{22}

We next sought to ensure that the decrease in labeling upon treatment with the MO was not due to a non-specific, global perturbation of glycan biosynthesis. Fucose is a monosaccharide not found in zebrafish GAGs; its incorporation into cell-surface glycans should therefore be unaffected by the XylT1 MO. To determine the effects of
Figure 3.10: Flow cytometry revealed that the XylT1 MO attenuates the UDP-4-XylAz signal. Zebrafish embryos injected with UDP-4-XylAz or UDP-4-XylAz and 6 ng of the XylT1 MO were allowed to develop to 24 hpf. After incubating in a solution of NHS-647, the embryos were washed and dissociated into single cells. Cells were then reacted with DIFO-biotin and subsequently incubated in a solution of avidin-488 to be analyzed by flow cytometry.

the XylT1 MO on cell-surface fucosylation, a proxy for non-GAG glycosylation, we dissociated MO-treated embryos into single cells, incubated them with a biotinylated Aleuria aurantia lectin that recognizes fucose, incubated with avidin-AlexaFluor-488, and measured for fluorescence by flow cytometry (Figure 3.11).\textsuperscript{17} Comparison of embryos injected with the XylT1 MO and untreated embryos confirmed that the MO had no effect on fucosylation, and, by inference, global glycosylation.

**UDP-4-XylAz functions as a metabolic inhibitor of GAG biosynthesis**

Determination that UDP-4-XylAz effectively delivers 4-XylAz to sites of GAG modification was followed by studies to provide direct evidence that xylose with an azide replacing the C-4 hydroxyl group does in fact prevent GAG polymerization and functions as an inhibitor of GAG biosynthesis. Though unlikely, it was considered a possibility that by replacing the C-4 hydroxyl group with an azide, the elaborating enzymes may polymerize the GAG chain from the C-3 position of xylose, for example. To confirm the inhibitory effect of 4-XylAz, an in vitro experiment was performed in which a peptide modified with 4-XylAz, or natural xylose as a control, was incubated with the enzyme that elaborates natural xylose, β4-galactosyltransferase (β4GalT7),\textsuperscript{23} and its required substrate, UDP-galactose (UDP-Gal). Successful galactose modification was then
Fucosylation is unaffected by injection of XylT1 MO. Embryos were injected with 6 ng of the XylT1 MO and allowed to develop to 24 hpf. Untreated and MO-injected embryos were then dissociated into single cells and incubated with biotinylated Aleuria aurantia lectin (AAL-biotin). As a control, a second group of untreated embryos was dissociated into single cells and incubated in a solution of AAL-biotin that had been pre-incubated with fucose to remove the fucose binding ability. After washing to remove excess AAL-biotin, cells were incubated with avidin-488 and analyzed by flow cytometry. Determined by mass spectrometry; following its injection into the mass spectrometer, electrospray ionization is used to cleave the glycosidic bond, which results in concomitant loss of the glycan from the peptide. The mass transition corresponding to loss of the glycan is detected using the multiple reaction-monitoring mode (MRM) of the spectrometer (m/z 1056.2 → 909.1, z=2 for the Gal-xylose modified peptide and m/z 1068.6 → 909.1, z=2 for the Gal-4-XylAz modified peptide) and quantified to verify the presence or absence of the glycosylated peptide. Thus, this method enables determination of galactosyltransferase activity with the xylose and 4-XylAz modified peptides based on the presence or absence of the glycosylation product. While β4GalT7 readily glycosylated the xylose-modified peptide (Figure 3.12A), the mass transition corresponding to loss of Gal-4-XylAz was not observed, verifying that β4GalT7 was not able to glycosylate the 4-XylAz modified peptide (Figure 3.12B). Because human and zebrafish β4GalT7 share 66% sequence identity, this is a good indication that 4-XylAz is not elaborated with galactose in vivo.

We next investigated the effects of UDP-4-XylAz treatment on the total amount of HS and CS per embryo, which we quantified using a previously described method.²⁴
Figure 3.12: 4-XylAz prevents elaboration by β4GalT7. Following incubation with β4GalT7 and UDP-Gal, the 4-XylAz or xylose modified peptides were injected into a mass spectrometer and subjected to ionization conditions which result in cleavage of the glycosidic bond. The mass transition corresponding to loss of the glycan (noted in the top left corner of each chromatogram) is detected and used determine the presence or absence of the glycosylated peptide, and thus the ability of the transferase to elaborate the 4-XylAz modified peptide. Although the xylose-modified peptide can be elaborated with galactose (A), the 4-XylAz-modified peptide cannot (B).

GAGs were isolated from UDP-4-XylAz-treated or untreated embryos, enzymatically degraded into disaccharides, and analyzed by reverse phase ion pairing HPLC. Treatment with UDP-4-XylAz caused a reduction in HS and CS levels by 47% and 77%, respectively (Figure 3.13, A and B). The unequal effect of the inhibitor on HS and CS levels is consistent with observations in zebrafish mutants with genetic knockdowns of the GAG biosynthesis enzymes responsible for constructing the core tetrasaccharide. We further analyzed the extent of sulfation of HS and CS in the presence and absence of the metabolic inhibitor. In line with the reduction in GAG levels, HS and CS sulfation were decreased by 53% and 83%, respectively (Figure 3.13, C and D). In parallel, we analyzed the disaccharide compositions of HS and CS in the presence and absence of UDP-4-XylAz. In agreement with its effects on total sulfation, the inhibitor did not appreciably alter the HS or CS disaccharide composition (Figure 3.13, E and F).
Figure 3.13: Effects of UDP-4-XylAz on GAG production. UDP-4-XylAz effect on HS expression (A), CS expression (B), HS total sulfation (C), CS total sulfation (D), HS disaccharide composition (E), and CS disaccharide composition (F). Grey bars, untreated; black bars, UDP-4-XylAz. Error bars denote the standard deviation from three replicate experiments. * p < 0.05, ** p < 0.01, *** p < 0.001.
Aberrant GAG glycosylation induced by UDP-4-XylAz causes dramatic phenotypes in developing zebrafish embryos

Finally, we examined the morphological effects of GAG disruption by UDP-4-XylAz injection. Zebrafish embryos treated with the metabolic inhibitor displayed an observable phenotype that is subtle at early stages (18 hpf, data not shown) and becomes more prominent as the embryo develops (Figure 3.14). UDP-4-XylAz-treated

![Image of embryos with various phenotypes](image)

**Figure 3.14: UDP-4-XylAz injection causes embryological defects.** Compared to 24 hpf controls (A), injected embryos (B) have a shortened axis, disorganized brain with enlarged hindbrain ventricle, small eyes, and abnormally shaped myotomes (see inset). By 48 hpf, notochord defects and tail blistering are prominent in treated versus control embryos (C and D). At 6 dpf, treated embryos have neurocranium defects, including a narrow ethmoid plate (arrow) and less cohesive midline cartilagenous cells (arrowhead), and the absence of a differentiated notochord (asterisk) (E and F), phenotypes strikingly similar to that of lamininb1a<sup>b1166</sup> mutant embryos (G and H). Another craniofacial defect in treated embryos is the narrowing of the lower jaw (I and J); visceral skeletal elements (K and L), including the pharyngeal arches (not shown) are largely normal. Alcian-stained cartilage elements are blue; Alizarin-stained bone elements are red. Although the alizarin staining in the lamb1a<sup>b1166</sup> mutant panel had faded significantly prior to imaging, additional preparations confirm the notochord defect described for laminin mutants.27
fish had a shortened axis, disorganized brain, small eyes, notochord defects, and abnormally shaped myotomes by 24 hpf, as well as prominent tail blistering by 48 hpf (Figure 3.14, A-D). These phenotypes are reminiscent of those observed for zebrafish with GAG or proteoglycan defects induced by genetic methods,24,25 and of zebrafish laminin (lam) mutants, particularly lama1/bashful, lamb1a/grumpy, and lamc1/sleepy (sly).26–29 By 6 dpf, craniofacial abnormalities were apparent in live UDP-4-XylAz-treated fish. At this stage, the craniofacial skeleton is comprised of a simple pattern of cartilage and bone elements; thus, we utilized stains for cartilage and bone (alcian and alizarin, respectively) to better understand the impact of UDP-4-XylAz on these structures. In inhibitor-treated embryos, alcian staining of cartilaginous elements revealed defects in the neurocranium, including a narrow ethmoid plate and less cohesive midline cartilage cells, while alizarin staining of developing bone tissues revealed a lack of notochord differentiation (Figure 3.14, E and F). A similar ethmoid plate phenotype is observed in embryos with deficiencies in Hedgehog signaling, defects in cytokine signaling, or excess retinoic acid signaling,30–32 yet the alcian stains are most strikingly similar to those of lamb1a b1166 mutants (Figure 3.14, compare E and F with G and H), especially in light of other phenotypic similarities between inhibitor-treated and laminin mutant embryos. Another craniofacial element defect in UDP-4-XylAz-injected embryos includes a narrowing of the lower jaw elements (Figure 3.14, I and J); otherwise visceral skeletal elements, including the pharyngeal arches (not shown) are largely normal (Figure 3.14, K and L). During early development, laminin and proteoglycan interactions are critical for basement membrane integrity and interactions with growth factors; the phenotypic similarities of UDP-4-XylAz-injected and lamb1a b1166 mutant embryos suggest that the inhibitor may interfere with these interactions.33

Conclusion

UDP-4-XylAz is an effective chain-terminating metabolic inhibitor of GAG biosynthesis in zebrafish. The compound has a unique attribute in that its organismal sites of inhibition can be visualized in vivo through bioorthogonal reaction with fluorescent cyclooctynes. Comparison of GAGs from UDP-4-XylAz and untreated embryos revealed differences in GAG abundance that likely cause the specific embryological defects observed. This metabolic inhibitor should aid in elucidating novel roles of GAG chains during vertebrate development.
Experimental Methods

General materials and methods

Biological reagents were obtained from commercial sources and used without further purification. Pronase (protease from *Streptomyces griseus*, type XIV), tricaine (ethyl 3-aminobenzoate methanesulfonate), phenol red, and *N*-phenylthiourea (PTU) were all obtained from Sigma-Aldrich. *N*-hydroxysuccinimidy ester of AlexaFluor-647 (NHS-647), rhodamine-dextran (dextran, tetramethylrhodamine, 10,000 MW), avidin conjugated to AlexaFluor-488 (avidin-488), TRIzol Reagent, Superscript III reverse transcriptase, and oligo(dT)20 primers were obtained from Invitrogen. Other primers were synthesized by Elim Biopharmaceuticals. Morpholino oligonucleotide reagents were designed and synthesized by Gene Tools. NucleoSpin® RNA II RNA purification kit was purchased from Macherey-Nagel. Phusion High-Fidelity DNA Polymerase was purchased from Roche Applied Science. DIFO-biotin and DIFO-488 were synthesized as previously reported. Flow cytometry was performed on a BD Biosciences FACSCalibur flow cytometer, and analyses were performed using FlowJo software. All images were obtained on a Zeiss LSM 780 NLO AxioExaminer, which were obtained on a Zeiss LSM 510 META/NLO AxioImager. All images were analyzed using Slidebook 5.0 (Intelligent Imaging Innovations).

Zebrafish stocks and husbandry

Adult wild-type AB zebrafish were kept at 28.5 °C on a 14-h light/10-h dark cycle. Embryos were obtained from natural spawning and were maintained in embryo medium (EM; 150 mM NaCl, 0.5 mM KCl, 1.0 mM CaCl₂, 0.37 mM KH₂PO₄, 0.05 mM Na₂HPO₄, 2.0 mM MgSO₄, 0.71 mM NaHCO₃ in deionized water, pH 7.4). Embryos were developmentally staged according to Kimmel and coworkers.

Confocal microscopy imaging of labeling with each UDP-XylAz derivative

Embryos at the 1–4 cell stage were injected with 2 nL of a 25 mM solution of UDP-2-XylAz, a 50 mM solution of UDP-3-XylAz, a 25 mM solution of UDP-4-XylAz, or vehicle (in deionized water with 0.2 M KCl and 2.5% rhodamine-dextran). After developing to 24 hpf, embryos were manually dechorionated with forceps and bathed in 100 µL of a 200 µM solution of DIFO-488 (in EM) for 1 h at 28.5 °C. After quenching the DIFO reaction by transferring embryos to 35x10 mm tissue culture dishes filled with EM, they were rinsed sequentially in four 100x15 mm tissue culture dishes filled with EM and deposited into 35x10 mm tissue culture dishes filled with EM until mounted. The agarose mounting solution was then prepared by melting a 1% low-melt agarose (in EM) stock solution in a microwave and combining 800 µL with 400 µL of EM in a microcentrifuge tube. Embryos were then added to the agarose mounting solution in a minimal amount of EM, then removed in the mounting solution and deposited onto a microscope slide in ~ 200 µL of the mounting solution. After properly positioning the embryos they were imaged by
confocal microscopy.

**Flow cytometry measurement of labeling with each UDP-XylAz derivative**

Embryos at the 1–4 cell stage were injected with 2 nL of a 50 mM solution of UDP-2-XylAz, a 100 mM solution of UDP-3-XylAz, or a 25 mM solution of UDP-4-XylAz (in deionized water with 0.2 M KCl and 0.2 % phenol red); these concentrations were selected because they were the concentrations at which minimal morphological effects were observed. At 10 hpf, injected and untreated embryos were transferred to 1% high-melt agarose-coated 100x15 mm tissue culture dishes filled with EM. After decanting most of the EM, a pronase solution (1mg/mL in EM) was added and the embryos were allowed to incubate in this solution for 10 min, at which time the chorions are enzymatically removed by swirling the dish. The dechorionated embryos were then transferred to 1% high-melt agarose (in EM) coated 35x10 mm tissue culture dishes filled with EM, and incubated at 28.5 °C. After developing to 24 hpf, embryos were incubated in a 50 µM solution of NHS-647 for 30 min at 28.5 °C, and then washed in EM-filled 100x15 mm tissue culture dishes. After adding embryos to a 1.5 mL microcentrifuge tube and removing as much EM as possible, 100 µL of calcium-free Ringer’s solution (116 mM NaCl, 2.6 mM KCl, 5 mM HEPES, pH 7.0) was added. Gentle pipetting with a 200 µL pipette tip was used to deyolk the embryos. The embryos were then transferred to a 24-well plate with wells filled with 1.2 mL of a 5 mM solution of EDTA in phosphate-buffered saline (PBS) pre-warmed to 28.5 °C. After mixing by pipetting with a 200 µL pipette every 5 min for 30 min to facilitate dissociation into single cells, the EDTA was quenched by the addition of 200 µL of the stop solution (30% calf serum, 6 µM CaCl2, in PBS). After pelleting the cells by centrifugation for 5 min at 350xg, the supernatant was removed, and the cells were transferred to a V-bottom 96-well plate in 200 µL of FACS buffer (1% FBS in PBS). The cells were then washed following the wash procedure: after pelleting the cells by centrifugation at 3500 RPM for 3 min, and discarding the supernatant, they were resuspended in 200 µL of FACS buffer. The cells were then pelleted, the supernatant was removed, and they were incubated in a 200 µM solution of DIFO-biotin for 1 h at 28.5 °C. After two wash steps, the cells were resuspended in a 5 µg/mL solution of avidin-488 at 4 °C for 15 min. After two wash steps the cells were again incubated in a 5 µg/mL solution of avidin-488 at 4 °C for 15 min. After two additional wash steps the cells were measured for fluorescence by flow cytometry.

**XylT1 MO activity verification using RT-PCR**

Embryos at the 1–4 cell stage were injected with 2 nL of either a 0, 1, or 3, ng/nL solution of XylT1 MO (in deionized water with 0.2 M KCl and 0.2 % phenol red). At 18 hpf, the embryos were enzymatically dechorionated with pronase (1mg/mL in EM). The total RNA was isolated from embryos using TRIzol Reagent and the Macherey-Nagel Total RNA Isolation kit. The first strand cDNA was then synthesized using Superscript III reverse transcriptase and oligo(dT)20 as the first strand primer. Finally, the cDNA was amplified by PCR using the primers shown in Supplementary Table 1 and analyzed by agarose gel electrophoresis.
Embryos at the 1–4 cell stage were injected with 2 nL of either a 25 mM solution of UDP-4-XylAz or a 25 mM solution of UDP-4-XylAz with 3 ng/nL of XylT1 MO (in deionized water with 0.2 M KCl and 0.2 % phenol red). At 10 hpf, injected and untreated embryos were transferred to 1% high-melt agarose coated 100x15 mm tissue culture dishes filled with EM. After decanting most of the EM, a pronase solution (1mg/mL in EM) was added and the embryos were allowed to incubate in this solution for 10 min, at which time the chorions are enzymatically removed by swirling the dish. The dechorionated embryos were then transferred to 1% high-melt agarose (in EM) coated 35x10mm tissue culture dishes filled with EM and incubated at 28.5 °C. After developing to 24 hpf, embryos were incubated in a 50 µM solution of NHS-647 for 30 min at 28.5 °C, and then washed in EM filled 100x15 mm tissue culture dishes. After adding embryos to a 1.5 mL microcentrifuge tube and removing as much EM as possible, 100 µL of calcium-free Ringer’s solution (116 mM NaCl, 2.6 mM KCl, 5 mM HEPES, pH 7.0) was added. Gentle pipetting with a 200 µL pipette tip was used to deyolk the embryos. The embryos were then transferred to a 24-well plate with wells filled with 1.2 mL of a 5 mM solution of EDTA in phosphate buffer saline (PBS) prewarmed to 28.5 °C. After mixing by pipetting with a 200 µL pipette every 5 min for 30 min to facilitate dissociation into single cells, the EDTA was quenched by the addition of 200 µL of the stop solution (30% calf serum, 6 µM CaCl_2, in PBS). After pelleting the cells by centrifugation for 5 min at 350xg, the supernatant was removed, and the cells were transferred to a V-bottom 96-well plate in 200 µL of FACS buffer (1% FBS in PBS). The cells were then washed following the wash procedure: after pelleting the cells by centrifugation at 3500 RPM for 3 min, and discarding the supernatant, they were resuspended in 200 µL of FACS buffer. The cells were then pelleted, the supernatant was removed, and they were incubated in a 200 µM solution of DIFO-biotin for 1 h at 28.5 °C. After two wash steps the cells were resuspended in a 5 µg/mL solution of avidin-488 at 4 °C for 15 min. After two wash steps, the cells were again incubated in a 5 µg/mL solution of avidin-488 at 4 °C for 15 min. After two wash steps, the cells were incubated in 100 µL of a 2.5 µg/mL solution of propidium iodide for 15 min at rt in the dark, and measured for fluorescence by flow cytometry. XylT1 MO sequence: 5’–CAAGACAAAGATGAAATTAGTGGTG–3’.

Determining the effect of the XylT1 MO on fucose expression

Embryos at the 1–4 cell stage were injected with 2 nL of a 3 ng/nL solution of XylT1 MO (in deionized water with 0.2 M KCl and 0.2 % phenol red). At 10 hpf, injected and untreated embryos were transferred to 1% high-melt agarose coated 100x15 mm tissue culture dishes filled with EM. After decanting most of the EM, a pronase solution (1mg/mL in EM) was added and the embryos were allowed to incubate in this solution for 10 min, at which time the chorions are enzymatically removed by swirling. The dechorionated embryos were then transferred to 1% high-melt agarose (in EM) coated 35x10mm tissue culture dishes filled with EM and incubated at 28.5 °C. After developing to 24 hpf, embryos were transferred to a 1.5 mL microcentrifuge tube, as much EM as
possible was removed, and 100 µL of calcium-free Ringer’s solution (116 mM NaCl, 2.6 mM KCl, 5 mM HEPES, pH 7.0) was added. Gentle pipetting with a 200 µL pipette tip was used to deyolk the embryos. The embryos were then transferred to a 24-well plate with wells filled with 1.2 mL of a 5 mM solution of EDTA in phosphate buffer saline (PBS) pre-warmed to 28.5 °C. After mixing by pipetting with a 200 µL pipette every 5 min for 30 min to facilitate dissociation into single cells, the EDTA was quenched by the addition of 200 µL of the stop solution (30% calf serum, 6 µM CaCl₂, in PBS). After pelleting the cells by centrifugation for 5 min at 350xg, the supernatant was removed, and the cells were transferred to a V-bottom 96-well plate in 200 µL of FACS buffer (1% FBS in PBS). The cells were then washed following the wash procedure: after pelleting the cells by centrifugation at 3500 RPM for 3 min, and discarding the supernatant, they were resuspended in 200 µL of FACS buffer. The cells were then pelleted, the supernatant was removed, and they were incubated in a 5 µg/mL solution of AAL-biotin for 45 min at 4 °C. Additionally, as a control, one subset of untreated embryo cells was incubated in a 5 µg/mL solution of AAL-biotin that had been pre-incubated in 200 mM L-fucose to abrogate binding ability. After two wash steps, the cells were resuspended in a 5 µg/mL solution of avidin-488 at 4 °C for 30 min. After two additional wash steps the cells were measured for fluorescence by flow cytometry.

**XylT activity assay with UDP-4-XylAz**

These experiments were performed following the procedure by Kuhn and coworkers. Briefly, 20 µL of a 150 µM solution of UDP-4-XylAz was combined with 5 µL of a 250 mM solution of MES buffer (pH 6.5), 5 µL of a 50 mM MgCl₂ and 50 mM MnCl₂ solution, 20 µL of a 33 µM solution of the XylT acceptor peptide (biotin-NH-QEEEGSGGGQKK(5-fluorescein)-CONH₂), and 50 µL of a 15.0 µU/L solution of XylT1, a 40.5 µU/L solution of XylT2, or a 15.0 µU/L solution of heat-killed XylT1. After incubating at 37 °C for 90 min, the reaction was halted by heating to 100 °C for 10 min. After brief centrifugation at 10,000xg to remove precipitate, 5 µL of the clear supernatant was injected into a C18 UPLC column that was directly coupled to a Quattro LC tandem mass spectrometer fitted with a Z Spray ion source. The MRM mode of the spectrometer was then used to determine the amount of the acceptor peptide with 4-XylAz covalently linked, and thus the ability of the XylTs to accept UDP-4-XylAz as a substrate.

**Determining if β4GalT7 can elaborate a 4-XylAz modified peptide**

These experiments were performed following the procedure of Kuhn and coworkers. Briefly, 20 µL of a 0.2 µM solution of a xylose modified peptide (biotin-NH-QEEEGS(Xyl)GGGQKK(5-fluorescein)-CONH₂) or 4-XylAz modified peptide (biotin-NH-QEEEGS(4-XylAz)GGGQKK(5-fluorescein)-CONH₂) was combined with 5 µL of a 250 mM solution of MES buffer (pH 6.5), 5 µL of a 50 mM MgCl₂ and 50 mM MnCl₂ solution, 20 µL of a 1 mg/mL solution of UDP-Gal, and 50 µL of a 150 µU/L solution of β4GalT7. After incubating at 37 °C for 24 h, the reaction was halted by heating to 100 °C for 10 min. After brief centrifugation at 10,000xg to remove precipitate, 5 µL of the clear supernatant was injected into a C18 UPLC column that was directly coupled to a
Quattro LC tandem mass spectrometer fitted with a Z Spray ion source. The MRM mode of the spectrometer was then used to determine the amount of each acceptor glycopeptide with galactose covalently linked, and thus the ability of β4GalT7 to elaborate a 4-XylAz-modified peptide.

The effect of UDP-4-XylAz on GAG expression and composition

Embryos at the 1–4 cell stage were injected with 2 nL of a 25 mM solution of UDP-4-XylAz (in deionized water with 0.2 M KCl and 0.2 % phenol red). At 10 hpf, UDP-4-XylAz injected and untreated embryos were enzymatically dechorionated with pronase (1mg/mL in EM). After developing to 24 hpf, embryos were transferred to microcentrifuge tubes. After removing most of the EM, 1 mL of ddH2O was added and removed three times to fully remove buffer from the embryos, and the remaining water was removed by lyophilization. The GAGs were then isolated, enzymatically cleaved into disaccharides, and analyzed by HPLC following the protocol of Ledin and coworkers, except for elution of DEAE-columns with 2 M NaCl followed by desalting on PD MiniTrap™ G-25 columns (GE Healthcare Biosciences Uppsala, Sweden) before further analysis.

Determining the effect of UDP-4-XylAz on alcian blue and alizarin red staining

These experiments were performed following a procedure adapted from Kimmel and coworkers. Embryos at the 1–4 cell stage were injected with 2 nL of a 25 mM solution of UDP-4-XylAz (in deionized water with 0.2 M KCl and 0.2 % phenol red). An Alcian/Alizarin double-stain mixture was made by mixing a stock of 0.2% Alcian Blue dye dissolved in 80% ethanol, with a stock of 0.5% Alizarin Red dye dissolved in water. The final concentrations of the double-stain mixture were 0.02% Alcian Blue, 0.5% Alizarin Red, 10mM MgCl2, and 80% ethanol. Larvae were anesthetized with tricaine at 6 days post-fertilization and collected at a maximum of 50 fish per 1.5 mL tube. Water was removed and larvae were fixed with 1mL of 2% paraformaldehyde in 1X phosphate buffered saline for 1 hour, rocking at room temperature. Fixative was removed and the larvae were washed with 50% ethanol for 10 minutes, rocking at room temperature. After the dehydrating wash, larvae were stained overnight at room temperature with the double-stain mixture, bleached with 3% H2O2/0.5% KOH for 10 minutes, and cleared with successive washes of 25% glycerol/0.1% KOH and 50% glycerol/0.1%KOH. Larvae were stored in 50% glycerol/0.1% KOH at 4°C.

Determining the effect of UDP-4-XylAz on zebrafish embryo gross morphology by live imaging

Embryos at the 1–4 cell stage were injected with 2 nL of a 25 mM solution of UDP-4-XylAz (in deionized water with 0.2 M KCl and 0.2 % phenol red). At 24 and 48 hpf, UDP-4-XylAz injected and untreated embryos were mounted in low melt agarose, as described above, and imaged by confocal microscopy.
References


