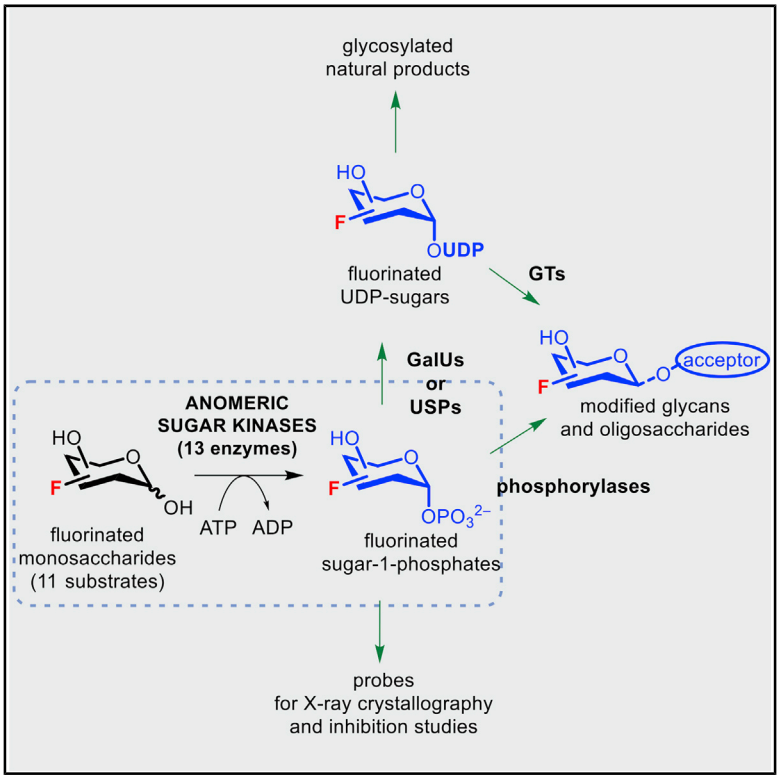
Brief Communication

Cell Chemical Biology

Profiling Substrate Promiscuity of Wild-Type Sugar Kinases for Multi-fluorinated Monosaccharides

Graphical Abstract



Highlights

- Anomeric kinases screened for activity against fluorinated monosaccharides
- Phosphorylation of 11 substrates including di- and tetrafluorinated sugars were observed
- Crystal structure of BiGalK shows active-site interactions with 2FGal
- This work has practical applications in medicinal chemistry and chemical biology

Authors

Tessa Keenan, Fabio Parmeggiani, Julien Malassis, ..., Bruno Linclau, Sabine L. Flitsch, Martin A. Fascione

Correspondence

sabine.flitsch@manchester.ac.uk (S.L.F.), martin.fascione@york.ac.uk (M.A.F.)

In Brief

Keenan, Parmeggiani et al. report the exploitation of wild-type anomeric sugar kinases to access fluorinated monosaccharide-1-phosphates. They demonstrate the effective phosphorylation of 11 different fluorinated monosaccharides, including di- and tetrafluorinated sugars. This work has practical applications in medicinal chemistry and chemical biology, particularly in the synthesis of fluorinated glycans.

Keenan et al., 2020, Cell Chemical Biology *27*, 1199–1206
September 17, 2020
https://doi.org/10.1016/j.chembiol.2020.06.005

CellPress

Cell Chemical Biology

CellPress

Brief Communication

Profiling Substrate Promiscuity of Wild-Type Sugar Kinases for Multi-fluorinated Monosaccharides

Tessa Keenan,[1,5] Fabio Parmeggiani,[2,5] Julien Malassis,[3] Clement Q. Fontenelle,[3] Jean-Baptiste Vendeville,[3] Wendy Offen,[1] Peter Both,[2] Kun Huang,[2] Andrea Marchesi,[2] Alex Heyam,[1] Carl Young,[4] Simon J. Charnock,[4] Gideon J. Davies,[1] Bruno Linclau,[3] Sabine L. Flitsch,[2,*] and Martin A. Fascione[1,6,*]

[1]Department of Chemistry, University of York, Heslington, York YO10 5DD, UK
[2]Manchester Institute of Biotechnology (MIB), School of Chemistry, University of Manchester, 131 Princess Street, Manchester M1 7DN, UK
[3]School of Chemistry, University of Southampton, Highfield, Southampton SO17 1BJ, UK
[4]Prozomix Limited, Haltwhistle Industrial Estate, Haltwhistle, Northumberland NE49 9HA, UK
[5]These authors contributed equally
[6]Lead Contact
*Correspondence: sabine.flitsch@manchester.ac.uk (S.L.F.), martin.fascione@york.ac.uk (M.A.F.)
https://doi.org/10.1016/j.chembiol.2020.06.005

SUMMARY

Fluorinated sugar-1-phosphates are of emerging importance as intermediates in the chemical and biocatalytic synthesis of modified oligosaccharides, as well as probes for chemical biology. Here we present a systematic study of the activity of a wide range of anomeric sugar kinases (galacto- and *N*-acetylhexosamine kinases) against a panel of fluorinated monosaccharides, leading to the first examples of polyfluorinated substrates accepted by this class of enzymes. We have discovered four new *N*-acetylhexosamine kinases with a different substrate scope, thus expanding the number of homologs available in this subclass of kinases. Lastly, we have solved the crystal structure of a galactokinase in complex with 2-deoxy-2-fluorogalactose, giving insight into changes in the active site that may account for the specificity of the enzyme toward certain substrate analogs.

INTRODUCTION

Understanding the biological functions of complex oligosaccharides and the mechanisms that underpin carbohydrate recognition, synthesis, and bioconversion is an ever-expanding research area in glycobiology that has major implications in medicine and drug development. For these studies the supply of new substrates, probes, and reference materials is often a major bottleneck due to the high chemical complexity and cost of these target products. Although undoubtedly essential and powerful, the chemical synthesis of highly complex oligosaccharides such as blood group antigens or sulfated glycosaminoglycans can be technically challenging and time-consuming (Maza et al., 2012; Xia et al., 2006; Hanashima et al., 2007; Tiruchinapally et al., 2011). As a complementary approach, chemo-enzymatic strategies (Muthana et al., 2009) whereby chemical syntheses are used to generate natural and non-natural substrate analogs, before enzymatic functionalization to produce key intermediates that enable access to a range of modified glycans, have also shown great promise. A representative example of this strategy is the synthesis of modified sugar-1-phosphates. Sugar-1-phosphates are important intermediates in sugar metabolism and utilization (Varki et al., 2015-2017), being the first intermediate biosynthesized from the reducing monosaccharide or the second intermediate biosynthesized from sugar-6-phosphates. Modified sugar-1-phosphates are commonly employed as basic building blocks toward the synthesis of diverse oligosaccharides, glycoproteins, and glycosylated natural products, in their protected form as glycosyl donors for chemical syntheses or unprotected as substrates in chemo-enzymatic approaches (Plante et al., 2001; Fu et al., 2003). Fluorinated sugar-1-phosphates, in particular, are important (Figure 1), as it is well known that the presence of fluorine can influence the stereo-electronic properties of a molecule (Gillis et al., 2015; Wang et al., 2013, O'hagan, 2008). Fluorinated sugar-1-phosphates are widely employed in chemo-enzymatic strategies for the synthesis of fluorinated glycans. For example, glycoside phosphorylases that use sugar-1-phosphates in catalyzing the regio- and stereo-selective formation of glycosidic bonds (O'neill and Field, 2015) have been used for the chemo-enzymatic synthesis of fluorinated lacto- and galacto-n-bioses (galacto-n-biose/lacto-n-biose phosphorylase) (Yu et al., 2010), enabling the synthesis of fluorinated T-antigen analogs (Yan et al., 2013). Monodeoxy-monofluorinated galactose-1-phosphates as substrates for uridyltransferases (GalUs) or UDP-sugar pyrophosphorylases (USPs) have been utilized in the synthesis of various fluorinated UDP-sugars (Errey et al., 2004) and larger glycans following their coupling via glycosyltransferases (GTs) (Kodama

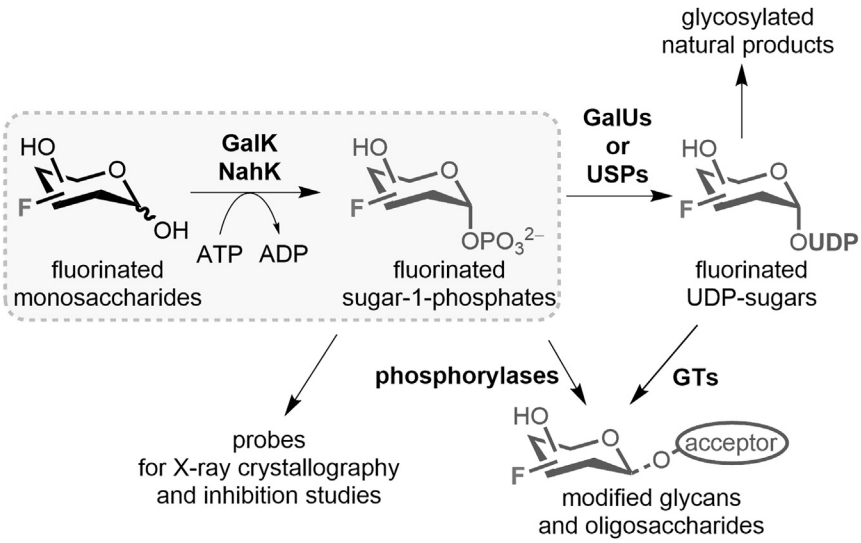


Figure 1. Kinase-Mediated Synthesis of Fluorosugar-1-Phosphates and Their Applications

et al., 1993). Similarly, a UDP-6-deoxy-6-fluoro-D-glucose was used as a donor for the synthesis of a glycosylated natural product, *N*-methylanthranilate (Caputi et al., 2013). Furthermore, the peculiar stereo-electronic changes induced by the substitution of sugar ring hydroxyls for fluorine affects how these molecules interact with enzymes, making them extremely useful probes for structural and biochemical investigations (Ioannou et al., 2011). For example, UDP 2-deoxy-2-fluorogalactose has been used to unpick mechanistic differences between α- and β-galactosyltransferases (GalTs) through enzyme inhibition studies (Burkart et al., 2000) and to investigate the catalytic mechanism of LgtC from *Neisseria meningitidis* through protein crystallography. The use of fluorinated carbohydrates in structural studies have also been key to identifying covalent reaction intermediates, as was demonstrated with the β-glucosidase Cex (White et al., 1996), and to gain insight into residues important in ligand recognition, as seen previously for mutases and synthases (Gibson et al., 2004; Van Straaten et al., 2015). Fluorinated carbohydrates have proved useful as probes to study parasite adhesion in toxoplasmosis (Allman et al., 2009). Additionally, they have been used as probes to map antibody subsites in anti-carbohydrate antibodies and in metabolic studies (Schmidt et al., 1978; Glaudemans, 1991).

Two main classes of enzymes are responsible for the *in vivo* synthesis of sugar-1-phosphates from non-activated monosaccharides: galactokinases (GalKs), which catalyze the ATP-dependent phosphorylation of α-D-galactose (Gal), and *N*-acetylhexosamine kinases (NahKs, also known as HexNAcKs), which phosphorylate both *N*-acetylglucosamine (GlcNAc) and *N*-acetylgalactosamine (GalNAc). Many well-characterized examples of GalKs have been reported (Chu et al., 2009; Dey, 1983; Schell and Wilson, 1977; Yang et al., 2003; Chen et al., 2011; Li et al., 2012, 2014; Huang et al., 2018; Keenan et al., 2019; Hartley et al., 2004; Grossiord et al., 2003) while only a few NahKs have been previously investigated (Cai et al., 2009; Li et al., 2011). Both classes are highly promiscuous, exhibiting a broad tolerance to natural and derivatized monosaccharides. Despite the interest in fluorinated sugar-1-phosphates and the several published examples of GalKs and NahKs, few examples of kinase-mediated fluorosugar phosphorylation have been demonstrated. All four monodeoxy-monofluoro galactoses (2FGal **1a**, 3FGal **2a**, 4FGal **3a**, and 6FGal **4a**) have long been recognized as substrates for yeast GalK (Thomas et al., 1974) and investigated subsequently as model substrates with EcGalK (Errey et al., 2004; Caputi et al., 2013) and LgGalK (Huang et al., 2018). Some groups have also reported the conversion of geminally fluorinated monosaccharide derivatives, such as 6,6-difluoro-D-fucose (6-deoxy-6,6-difluoro-D-galactose) (Yang et al., 2004, 2005) and *N*-trifluoroacetyl-D-glucosamine/-D-galactosamine (Li et al., 2011). However, the tolerance of anomeric kinases to vicinal dideoxy-difluorinated substrates, or monosaccharides with a higher degree of fluorination, is currently unexplored. This class of sugars is of interest in the study of protein-carbohydrate interactions due to their polar hydrophobicity, which can increase the affinity of these compounds for proteins (Biffinger et al., 2004; Kim et al., 1998; Ioannou et al., 2011; Bresciani et al., 2010; Denavit et al., 2019; Van Straaten et al., 2015, N'go et al., 2014). Here, we describe a systematic approach whereby fluorosugar reactivity of diverse anomeric kinases was explored, leading to examples of mono-, di-, and tetrafluorinated substrates that are phosphorylated by these enzymes. We also describe the three-dimensional (3D) structure of BiGalK and showcase, through ligand complexes with 2FGal and Gal, subtle changes in the active site that can occur when a sugar ring hydroxyl is replaced with fluorine

RESULTS

A panel of seven published GalKs—EcGalK (Yang et al., 2003), PfGalK (Hartley et al., 2004), SpGalK (Chen et al., 2011), LlGalK (Grossiord et al., 2003), LgGalK (Huang et al., 2018), BiGalK (Li et al., 2012), and ScGalK (Keenan et al., 2019)—and two published NahKs, BliNahK (Li et al., 2011) and BllNahK (Li et al., 2011), was chosen. Due to the lower diversity available in the latter family, four putative NahK homologs (49%–54% sequence identity) from bacterial species (Data S1) were also selected and included in the panel: ApNahK (Genbank: WP_091282512), CxNahK (Genbank: WP_045094128), EtNahK (Genbank: ODM08532), and TbNahK (Genbank: WP_062612735). The kinase panel was investigated for the conversion of a series of deoxyfluorinated galactoses, glucoses, and a mannose derivative. While some examples of kinase-mediated phosphorylation of monofluorinated galactoses have been reported previously (Huang et al., 2018; Thomas et al., 1974; Errey et al., 2004; Caputi et al., 2013), we rationalized that the inclusion of these compounds in our extensive screen could help to identify kinases

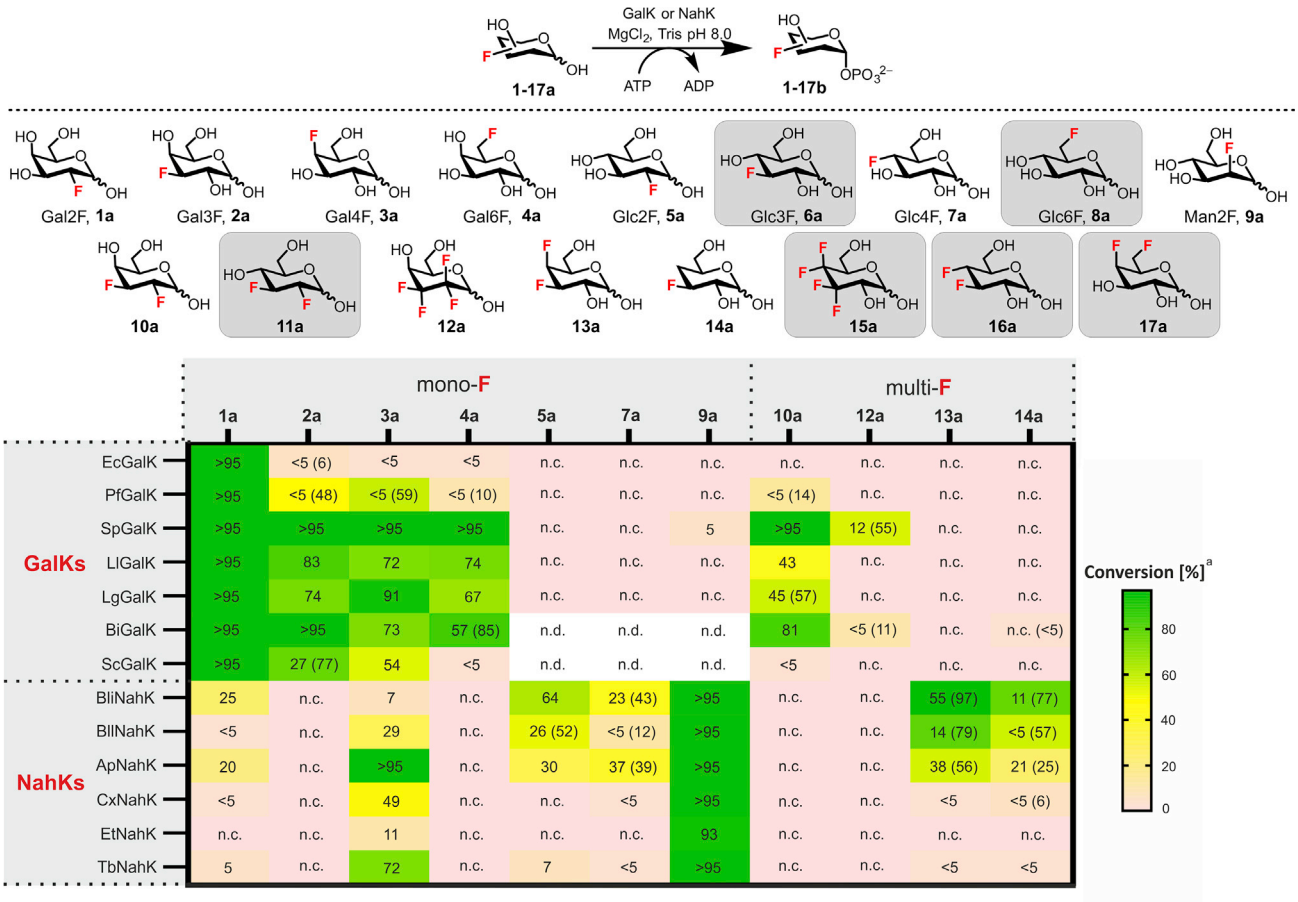


Figure 2. Screening of Fluorinated Monosaccharides against GalKs and NahKs

[a]Determined by ^{19}F NMR (n.c. = no conversion, n.d. = not determined). Optimised conversions shown in brackets. Limit of quantitation estimated at 5% conversion. Compounds 6a, 8a, 11a, 15a, 16a and 17a (grey) were not substrates for any kinases tested.

with higher activity against these, as well as those with a broader substrate scope. Kinase-mediated conversion of the substrates to the sugar-1-phosphate was determined by ^{19}F nuclear magnetic resonance (NMR) (Figure 2 and Data S1) by relative integration of the corresponding resonances, which were typically baseline-separated. Phosphorylation at the anomeric position of the monosaccharides was evident, as shown by the formation of only a single new species in the ^{19}F NMR, in comparison with two species that were present in the reference spectrum of the monosaccharide, where the anomeric center is in the hemiacetal form. Formation of the sugar-1-phosphate was further validated by high-resolution mass spectrometry (HRMS) analysis (Data S1). As expected, the GalKs demonstrated a strict preference for deoxyfluorinated galactoses. Deoxyfluorination at the 2-position was the most tolerated, as demonstrated by the near full conversion of 2FGal (**1a**) by all of the GalKs. All GalKs displayed some activity toward 3FGal (**2a**), 4FGal (**3a**), and 6FGal (**4a**), although there were clearly differences in the activity profiles toward these substrates between the different enzymes. For example, BiGalK catalyzed nearly full conversion of 3FGal but moderate conversions of 4FGal and 6FGal. In contrast, ScGalK catalyzed moderate conversion of 4FGal but low-level conversions of 3FGal and 6FGal. Overall, SpGalK catalyzed the highest conversions of all deoxyfluorinated galactoses (>95%) as well as the conversion of 2FMan (**9a**) at low levels, further indicating its promiscuity. The NahKs displayed a broader fluorinated substrate scope, demonstrating the conversion of some deoxyfluorinated galactoses, glucoses, and the mannose derivative. Most NahKs displayed moderate activity toward 2- and 4-deoxyfluorinated analogs while displaying little or no activity toward their 3- and 6-deoxyfluorinated counterparts. As expected, most of the NahKs phosphorylated both galacto- and gluco-configured substrates. This could be explained by a lack of hydrogen-bonding interactions made between the enzyme active-site residues and the 4-OH, as is shown in the crystal structure of BliNahK in complex with GlcNAc (Wang et al., 2014) (Figure S1). All NahKs catalyzed the highest conversions of 2FMan (**9a**). In previous studies by Li et al. (2011), BllNahK and BliNahK were shown to be highly promiscuous enzymes, not only demonstrating the efficient conversion of GalNAc and GlcNAc and derivatives but also mannose and its C-2, C-4, and C-6 derivatives, albeit with lower specific activities. They were also shown to convert glucose and galactose, though to a lesser extent than mannose. Their findings are consistent with our NahK screening results, which suggest that 2FMan (**9a**) is a better substrate for NahKs than any of the galacto- or gluco-configured substrates.

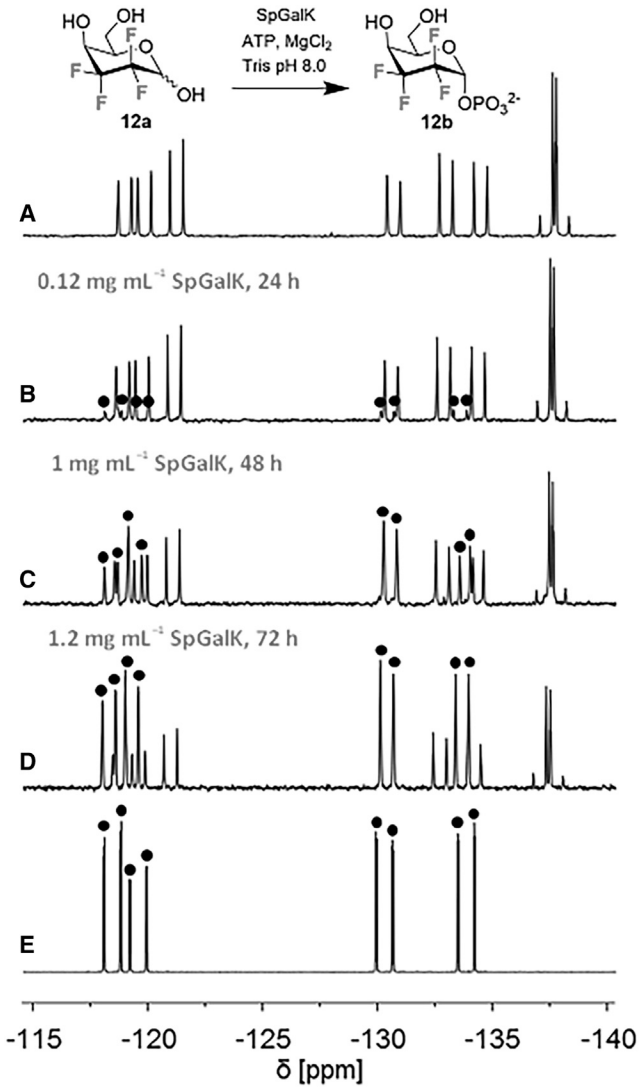


Figure 3. ^{19}F NMR Analysis of SpGalK-Mediated Phosphorylation of Substrate 12a

Peaks corresponding to product **12b** are marked by filled circles.
(A) Reference spectrum of **12a**.
(B) Reaction under screening conditions.
(C and D) Reactions with increased loading and reaction times.
(E) Reference spectrum of **12b**.

The panel of kinases was then investigated for activity toward vicinal dideoxygenated monosaccharides having di- and tetrafluorinated motifs (Figure 2 and Data S1). Interestingly, most of the GalKs phosphorylated 2,3-dideoxy-2,3-difluorogalactose (**10a**) (Figure S3) but not 3,4-dideoxy-3,4-difluorogalactose (**13a**) or 4,6-dideoxy-4,6-difluorogalactose (**17a**), despite being able to convert monofluorinated monosaccharides 3FGal, 4FGal, and 6FGal. Notably, SpGalK and BiGalK also converted 2,3-dideoxy-2,2,3,3-tetrafluoro-D-*threo*-hexopyranose (**12a**) but not 3,4-dideoxy-3,3,4,4-tetrafluoro-D-*erythro*-hexopyranose (**15a**). Although only trace amounts of product was observed under normal screening conditions, reaction optimization by increasing the concentration of SpGalK and the reaction time improved the conversion to 55% (Figure 3). The chemically synthesized sugar phosphate, 2,3-dideoxy-2,2,3,3-tetrafluoro-D-*threo*-hexopyranose phosphate (**12b**) (N'go et al., 2014), was used as a ^{19}F NMR standard to validate the presence of the correct product. The findings that GalKs could only catalyze the conversion of vicinal di- and tetrafluorinated galactoses modified at the 2- and 3-positions further support observations made in the deoxyfluorinated monosaccharide screening, which highlighted the high activity of GalKs (particularly SpGalK and BiGalK) toward 2- and 3-deoxyfluorinated galactose derivatives. On the other hand, the GalKs did not show turnover of the gluco-configured substrates **11a** and **16a** or the vicinal 3,4-dideoxy-3-fluoro sugar (**14a**), further emphasizing the importance of galacto-configuration at the 4-position in determining GalK substrate preference. In contrast to the GalKs, the NahKs were moderately active toward 3,4-dideoxy-3,4-difluorogalactose (**13a**) and the 3,4-dideoxy-3-fluoro sugar (**14a**), and conversions of up to 97% and 77% could be obtained for **13a** and **14a**, respectively (Figure 2) but displayed no activity toward 3,4-dideoxy-3,4-difluoroglucose (**16a**). Despite increasing the enzyme loading and reaction times, NahKs failed to reveal any activity toward analogs **10a**, **11a**, **12a**, **15a**, or **17a**. Considering that a single fluorine on a monosaccharide substrate can lead to reduced enzyme catalytic efficiency (Council et al., 2020), remarkably within this panel of eight heavily fluorinated monosaccharides, four of them could be effectively phosphorylated (>50% conversion) by at least one of the kinases investigated.

The preference of GalKs for galacto-configured substrates fluorinated at the 2-position, emphasized by the preference of difluorinated analog **10a** over analogs **13a** and **17a**, prompted us to investigate these interactions at the molecular level. Using BiGalK as a model example, the crystal structures of BiGalK in complex with 2FGal **1a** (PDB: 6TEQ), Gal (PDB: 6TER), and ADP (PDB: 6TEP) were solved (Table S1). Overall, the secondary structure architectures of BiGalK were nearly indistinguishable between the three ligand-bound complexes (Figure S2A). The position of the bound ADP was as expected (Figure S2B) when compared with the nucleotide-binding sites in the structures of LlGalK (Thoden and Holden, 2003) and PfGalK (Hartley et al., 2004). In the sugar-bound complexes, 2FGal (**1a**) and Gal are present as a mixture of the α and β forms. In the Gal-bound structure, the ligand is positioned through a network of hydrogen-bonding interactions (Figure 4A). The active-site residues that interact directly with Gal (R40, H47, D49, G188, D191, and Y248) are conserved in all bacterial GalKs investigated here (Figure S2C). As expected for the α-configured sugar, the 1-OH is positioned within hydrogen-bonding distance of R40 and close to D191 (within 3.2–3.6 Å). The corresponding residues in *Lactococcus lactis* GalK (R37, D186) were previously shown to have a role in catalysis by deprotonating the 1-OH of Gal (Thoden and Holden, 2003). The 2-OH forms hydrogen-bonding interactions with the side chain of D191 and some water molecules, while the 3-OH hydrogen bonds to OD2 D49, N G188, and a water molecule. D49, which forms hydrogen bonds with both the 3-OH and 4-OH, appears to be important for substrate positioning. This potentially accounts for the absence

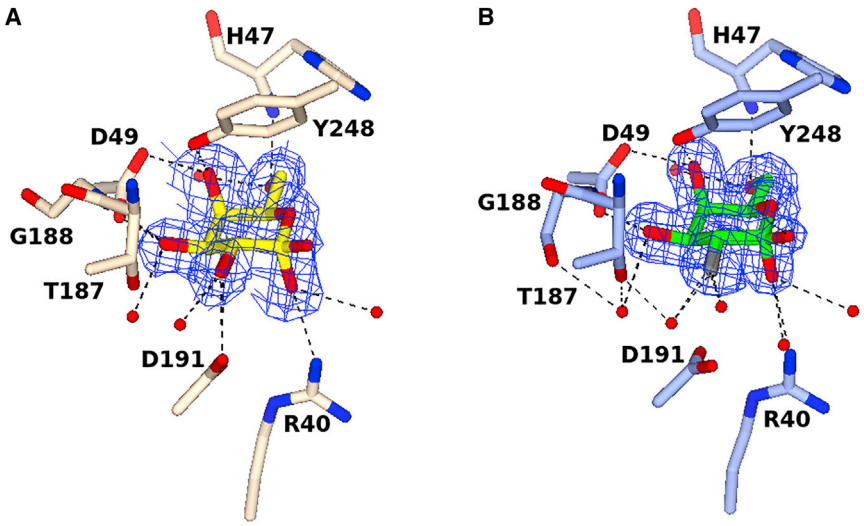


Figure 4. BiGalK Active Site Residues Interacting with Gal and 2FGal

Electron density maps are REFMAC maximum-likelihood/σA weighted $2F_o-F_c$ maps contoured at an r.m.s.d. level of 1.0 σ for both (0.31 and 0.36 electrons/Å^3 for (A) and (B), respectively). Alpha and beta anomers are modelled at 0.9/0.1 occupancy in (A) and 0.6/0.4 occupancy in (B), respectively.

of GalK activity toward **13a**, as fluorines at the 3- and 4-positions could not act as hydrogen-bond donors. Whereas the sugar molecule is relatively accessible near the 2-OH where there is a network of hydrogen-bonded water molecules, the active-site residues interact more closely around the 4-OH and 6-OH.

In the 2FGal-bound structure, the 2-fluorine can only serve as a hydrogen-bond acceptor and therefore cannot form a hydrogen bond with the side chain of D191 (Figure 4B). There are one or two additional water molecules modeled in the region close to and forming hydrogen bonds with the 2-fluorine in the 2FGal structure molecules, compared with those observed for the 2-OH group in the Gal structure. These extra water molecules may help to compensate for the loss of interactions with D191. The carbonyl group of G188 is flipped, allowing it to form a new hydrogen bond with the water molecule that interacts with the 3-OH. OG1 T187 appears to stabilize the water molecules that interact with the 2-fluorine and the 3-OH. These subtle variations in the ligand-water hydrogen-bonding network may explain why fluorination at the 2- and 3-positions is well tolerated by the enzyme. However, the kinetic parameters of BiGalK toward Gal (K_M = 1.029 mM, k_{cat}/K_M = 132.32 s^{-1} mM^{-1}) and 2FGal (K_M = 4.742 mM, k_{cat}/K_M = 13.5 s^{-1} mM^{-1}) show that even a single fluorine at the 2-position results in a near 10-fold decrease in the catalytic efficiency of the enzyme (Figure S5). One could envisage that the restricted space around the 4-OH and 6-OH prohibits access for water molecules that could form compensatory hydrogen bonds when these groups are substituted with fluorine, as in **17a**. This observation is reflected in the decreasing conversions when using BiGalK, as the number of fluorines around the substrates ring increases when comparing monofluorinated Gal2F **1a** (>95% conversion) with 2,3-dideoxy-2,3-difluoro galactose **10a** (81%), 2,3-dideoxy-2,2,3,3-tetrafluoro-D-*threo*-hexopyranose **12a** (<5%), indicating that if a fluorine is present at position-2 in the ring fluorination at other positions may be tolerated, but with an increasing penalty to reactivity. Overall, the structure showcases several residues that may be good targets for mutagenesis to expand BiGalK specificity toward fluorinated substrates. Residues H47 and Y248 that interact closely with the 4-OH and 6-OH would be good targets to explore for expanding BiGalK activity toward **17a**. When considering substrates with fluorination at the 3- and 4-positions, Y248, G188, and T187 could be targeted. Considering the apparent role of D49 for achieving correct substrate positioning, it is likely that mutations to this residue may completely eliminate enzyme activity.

DISCUSSION

In summary, we have demonstrated that wild-type anomeric sugar kinases can be exploited to access fluorinated monosaccharide-1-phosphates. We have systematically profiled kinase activity toward fluorinated analogs, demonstrating the first examples of polyfluorinated substrates accepted by this enzyme class. Our screen consisted of a large and diverse panel of kinases, highlighting that the selection of homologs with low sequence similarity can facilitate the identification of novel substrate specificities. We have shown that GalKs show a strong preference for galacto-configuration and a 4-OH. Deoxyfluorination at the 2-position was most tolerated by the GalKs, and this was further reinforced by the observation that only the vicinal dideoxy-difluorinated galactose modified at the 2- and 3-positions (**10a**) was turned over by the enzymes. Furthermore, 2,3-dideoxy-2,2,3,3-tetrafluoro-D-*threo*-hexopyranose (**12a**) was converted by SpGalK and BiGalK, with a conversion of 55% achieved under optimized reaction conditions. In contrast, NahKs displayed a broader fluorosugar substrate range than the GalKs, catalyzing the highest conversions of mannose derivative 2FMan (**9a**), but accepting both galacto- and gluco-configured substrates. They displayed a strict preference toward 2- and 4-deoxyfluorinated analogs and catalyzed the conversion of vicinal dideoxy-difluorinated sugars, 3,4-dideoxy-3,4-difluorogalactose (**13a**) and the 3,4-dideoxy-3-fluoro sugar (**14a**). Overall, we have demonstrated the effective phosphorylation of 11 different fluorinated monosaccharides and, through optimized reaction conditions, have obtained conversions of >90% for seven of these, setting the scene for further optimization by rational design and/or random mutagenesis. Through the structural characterization of a GalK in complex with 2FGal, we have gained valuable insight into subtle changes in the active site that can occur when a sugar ring hydroxyl is replaced with fluorine. These results pave the way to the one-step synthesis of useful intermediates for enzymatic glycosylations, to be explored through the synthesis of

UDP-sugars (GalUs/USPs) and their coupling via GTs, or directly via glycosyl phosphorylases. They may also serve as useful probes for chemical biology in structural studies to inform protein-ligand interactions.

SIGNIFICANCE

The development and optimization of new chemo-enzymatic strategies for the efficient sustainable synthesis of modified glycans is a relevant research area. In particular, the synthesis of fluorinated sugar-1-phosphates, which is rather challenging by traditional chemical methods, is an attractive target due to the range of applications of such molecules as key intermediates for the preparation of modified disaccharides and larger glycans, nucleotide sugars, and probes for chemical biology. Since the replacement of hydrogen atoms and hydroxyl groups with fluorine is known to influence the stereo-electronic properties of carbohydrates in a marked and sometimes subtly unpredictable way, in this work a systematic panel of 17 mono-, di-, and tetrafluorinated monosaccharides was investigated for biocatalytic 1-phosphorylation mediated by anomeric sugar kinases. 13 Different wild-type enzymes belonging to two structurally unrelated families (galactokinases [GalKs] and *N*-acetylhexosamine kinases [NahKs]) were screened, leading to the identification of suitable enzymes for the selective phosphorylation of 11 substrates tested, with good conversions in most cases. We show the first examples of the phosphorylation of polyfluorinated substrates mediated by sugar kinases, optimized to preparatively useful conversions. Furthermore, the discovery of four previously uncharacterized NahKs with a different substrate scope compared with known homologs adds valuable components to the biocatalytic toolbox for monosaccharide phosphorylation. Along with the practical applications in medicinal chemistry and chemical biology, this study also provides general information regarding the activity and specificity of the enzymes selected, as well as binding interactions (from X-ray crystallography data), which are useful information for the biochemical profiling and further improvement of sugar kinases by mutagenesis or evolution.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
  - Bacterial Strains
- METHOD DETAILS
  - General Methods
  - Chemical Synthesis
  - Determination of Fluorosugar Substrate Specificity
  - Optimisation of Substrate Conversion
  - ^{19}F NMR Analysis
  - DNS Assay
  - BiGalK Protein Production and Purification
  - ScGalK Protein Production and Purification
  - Molecular cloning and protein production for EcGalK, LlGalK, PfGalK, SpGalK, LgGalK, BllNahK, BliNahK, ApNahK, CxNahK, EtNahK and TbNahK (performed by Prozomix)
  - Determination of BiGalK Kinetic Parameters
  - Crystallization, Data Collection and Structure Solution
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.chembiol.2020.06.005.

ACKNOWLEDGMENTS

This project has been funded by the Industrial Biotechnology Catalyst (Innovate UK, BBSRC, EPSRC) to support the translation, development, and commercialization of innovative Industrial biotechnology processes: BB/M028836/1; BB/L013762/1, BB/M027791/1; BB/M02903411; BB/M028836/1; BB/M028941/1; BB/M02847X/1, EP/K039466/1, and ERC (788231-ProgrES-ERC-2017-ADG). Dr. Matt Cliff and Dr. Reynard Spiess (University of Manchester) are acknowledged for their support in NMR and HRMS analysis, respectively. The York Center of Excellence in Mass Spectrometry was created thanks to a major capital investment through Science City York, supported by Yorkshire Forward with funds from the Northern Way Initiative, and subsequent support from EPSRC (EP/K039660/1; EP/M028127/1).

AUTHOR CONTRIBUTIONS

F.P., T.K., K.H., and A.M. performed the enzymatic transformations; J.M., C.Q.F., and J.-B.V. performed chemical synthesis and characterization; T.K. and W.O. performed crystallography; P.B., C.Y., and T.K. performed sequence identification, cloning, and protein production; A.H. developed ^{19}F NMR methodology; M.A.F., S.L.F., B.L., G.J.D., and S.J.C. supervised the project; M.A.F., S.L.F., and B.L. designed the study; T.K. and F.P. wrote the paper; all authors commented on the paper.

DECLARATION OF INTERESTS

The authors declare no competing financial interests. Prozomix is a commercial enzyme producer.

Received: March 2, 2020
Revised: May 20, 2020
Accepted: June 12, 2020
Published: July 2, 2020

REFERENCES

Allman, S.A., Jensen, H.H., Vijayakrishnan, B., Garnett, J.A., Leon, E., Liu, Y., Anthony, D.C., Sibson, N.R., Feizi, T., and Matthews, S. (2009). Potent fluoro--oligosaccharide probes of adhesion in toxoplasmosis. ChemBioChem *10*, 2522–2529.

Biffinger, J.C., Kim, H.W., and Dimagno, S.G. (2004). The polar hydrophobicity of fluorinated compounds. ChemBioChem *5*, 622–627.

Bresciani, S., Lebl, T., Slawin, A.M., and O'hagan, D. (2010). Fluorosugars: synthesis of the 2, 3, 4-trideoxy-2, 3, 4-trifluoro hexose analogues of D-glucose and D-altrose and assessment of their erythrocyte transmembrane transport. Chem. Commun. *46*, 5434–5436.

Burkart, M.D., Vincent, S.P., Düffels, A., Murray, B.W., Ley, S.V., and Wong, C.-H. (2000). Chemo-enzymatic synthesis of fluorinated sugar nucleotide: useful mechanistic probes for glycosyltransferases. Bioorg. Med. Chem. *8*, 1937–1946.

Cai, L., Guan, W., Wang, W., Zhao, W., Kitaoka, M., Shen, J., O'neil, C., and Wang, P.G. (2009). Substrate specificity of N-acetylhexosamine kinase towards N-acetylgalactosamine derivatives. Bioorg. Med. Chem. Lett. *19*, 5433–5435.

Caputi, L., Rejzek, M., Louveau, T., O'neill, E.C., Hill, L., Osbourn, A., and Field, R.A. (2013). A one-pot enzymatic approach to the O-fluoroglucoside of N-methylanthranilate. Bioorg. Med. Chem. *21*, 4762–4767.

Chen, M., Chen, L.-L., Zou, Y., Xue, M., Liang, M., Jin, L., Guan, W.-Y., Shen, J., Wang, W., and Wang, L. (2011). Wide sugar substrate specificity of galactokinase from *Streptococcus pneumoniae* TIGR4. Carbohydr. Res. *346*, 2421–2425.

Chu, X., Li, N., Liu, X., and Li, D. (2009). Functional studies of rat galactokinase. J. Biotechnol. *141*, 142–146.

Council, C.E., Kilpin, K.J., Gusthart, J.S., Allman, S.A., Linclau, B., and Lee, S.S. (2020). Enzymatic glycosylation involving fluorinated carbohydrates. Org. Biomol. Chem. *18*, 3423–3451.

Denavit, V., Lainé, D., Bouzriba, C., Shanina, E., Gillon, E., Fortin, S., Rademacher, C., Imberty, A., and Giguère, D. (2019). Stereoselective synthesis of fluorinated galactopyranosides as potential molecular probes for galactophilic proteins: assessment of monofluorogalactoside-LecA interactions. Chem. Eur. J. *25*, 4478–4490.

Dey, P.M. (1983). Galactokinase of *Vicia faba* seeds. Eur. J. Biochem. *136*, 155–159.

Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. Section D: Biol. Crystallogr. *66*, 486–501.

Errey, J.C., Mukhopadhyay, B., Kartha, K.R., and Field, R.A. (2004). Flexible enzymatic and chemo-enzymatic approaches to a broad range of uridine-diphospho-sugars. Chem. Commun. 2706–2707.

Evans, P.R., and Murshudov, G.N. (2013). How good are my data and what is the resolution? Acta Crystallogr. Section D: Biol. Crystallogr. *69*, 1204–1214.

Fontenelle, C.Q., Shishmarev, D., Kuchel, P.W., and Linclau, B. (2017). The synthesis of 3,4-dideoxy-3,4-difluoro-D-glucose. Trends Carbohydr. Res. *9*, 29–34.

Fu, X., Albermann, C., Jiang, J., Liao, J., Zhang, C., and Thorson, J.S. (2003). Antibiotic optimization via in vitro glycorandomization. Nat. Biotechnol. *21*, 1467.

Gibson, R.P., Tarling, C.A., Roberts, S., Withers, S.G., and Davies, G.J. (2004). The donor subsite of trehalose-6-phosphate synthase binary complexes with UDP-glucose and UDP-2-deoxy-2-fluoro-glucose at 2 Å resolution. J. Biol. Chem. *279*, 1950–1955.

Gillis, E.P., Eastman, K.J., Hill, M.D., Donnelly, D.J., and Meanwell, N.A. (2015). Applications of fluorine in medicinal chemistry. J. Med. Chem. *58*, 8315–8359.

Glaudemans, C.P. (1991). Mapping of subsites of monoclonal, anti-carbohydrate antibodies using deoxy and deoxyfluoro sugars. Chem. Rev. *91*, 25–33.

Golten, S., Fontenelle, C.M.Q., Timofte, R.S., Bailac, L., Light, M., Sebban, M., Oulyadi, H., and Linclau, B. (2016). Enantioselective synthesis of dideoxy-tetrafluorinated hexoses. J. Org. Chem. *81*, 4434–4453.

Grossiord, B.P., Luesink, E.J., Vaughan, E.E., Arnaud, A., and de Vos, W.M. (2003). Characterization, expression, and mutation of the *Lactococcus lactis* galPMKTE genes, involved in galactose utilization via the Leloir pathway. J. Bacteriol. *185*, 870–878.

Hanashima, S., Castagner, B., Esposito, D., Nokami, T., and Seeberger, P.H. (2007). Synthesis of a sialic acid alpha(2-3) galactose building block and its use in a linear synthesis of sialyl Lewis X. Org. Lett. *9*, 1777–1779.

Harris, R.K., Becker, E.D., Cabral DE Menezes, S.M., Goodfellow, R., and Granger, P. (2002). NMR nomenclature: nuclear spin properties and conventions for chemical shifts. IUPAC recommendations 2001. Solid State Nucl. Magn. Reson. *22*, 458–483.

Hartley, A., Glynn, S.E., Barynin, V., Baker, P.J., Sedelnikova, S.E., Verhees, C., de Geus, D., Van der Oost, J., Timson, D.J., and Reece, R.J. (2004). Substrate specificity and mechanism from the structure of *Pyrococcus furiosus* galactokinase. J. Mol. Biol. *337*, 387–398.

Huang, K., Parmeggiani, F., Pallister, E., Huang, C.J., Liu, F.F., Li, Q., Birmingham, W.R., Both, P., Thomas, B., and Liu, L. (2018). Characterisation of a bacterial galactokinase with high activity and broad substrate tolerance for chemoenzymatic synthesis of 6-aminogalactose-1-phosphate and analogues. ChemBioChem *19*, 388–394.

Ioannou, A., Cini, E., Timofte, R.S., Flitsch, S.L., Turner, N.J., and Linclau, B. (2011). Heavily fluorinated carbohydrates as enzyme substrates: oxidation of tetrafluorinated galactose by galactose oxidase. Chem. Commun. *47*, 11228–11230.

Keenan, T., Mills, R., Pocock, E., Budhadev, D., Parmeggiani, F., Flitsch, S., and Fascione, M. (2019). The characterisation of a galactokinase from *Streptomyces coelicolor*. Carbohydr. Res. *472*, 132–137.

Kim, H.W., Rossi, P., Shoemaker, R.K., and Dimagno, S.G. (1998). Structure and transport properties of a novel, heavily fluorinated carbohydrate analogue. J. Am. Chem. Soc. *120*, 9082–9083.

Kodama, H., Kajihara, Y., Endo, T., and Hashimoto, H. (1993). Synthesis of UDP-6-deoxy-and-6-fluoro-D-galactoses and their enzymatic glycosyl transfer to mono-and biantennary carbohydrate chains. Tetrahedron Lett. *34*, 6419–6422.

Li, L., Liu, Y., Wang, W., Cheng, J., Zhao, W., and Wang, P. (2012). A highly efficient galactokinase from *Bifidobacterium infantis* with broad substrate specificity. Carbohydr. Res. *355*, 35–39.

Li, S.P., Hsiao, W.C., Yu, C.C., Chien, W.T., Lin, H.J., Huang, L.D., Lin, C.H., Wu, W.L., Wu, S.H., and Lin, C.C. (2014). Characterization of *Meiothermus taiwanensis* galactokinase and its use in the one-pot enzymatic synthesis of uridine diphosphate-galactose and the chemoenzymatic synthesis of the carbohydrate antigen stage specific embryonic antigen-3. Adv. Synth. Catal. *356*, 3199–3213.

Li, Y., Yu, H., Chen, Y., Lau, K., Cai, L., Cao, H., Tiwari, V.K., Qu, J., Thon, V., and Wang, P.G. (2011). Substrate promiscuity of N-acetylhexosamine 1-kinases. Molecules *16*, 6396–6407.

Malassis, J., Vendeville, J.-B., Nguyen, Q.-H., Boujon, M., Gaignard-Gaillard, Q., Light, M., and Linclau, B. (2019). Synthesis of vicinal dideoxy-difluorinated galactoses. Org. Biomol. Chem. *17*, 5331–5340.

Maza, S., Macchione, G., Ojeda, R., López-Prados, J., Angulo, J., de Paz, J.L., and Nieto, P.M. (2012). Synthesis of amine-functionalized heparin oligosaccharides for the investigation of carbohydrate–protein interactions in microtiter plates. Org. Biomol. Chem. *10*, 2146–2163.

Mccoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674.

Mtashobya, L., Quiquempoix, L., and Linclau, B. (2015). The synthesis of mono-and difluorinated 2, 3-dideoxy-d-glucopyranoses. J. Fluorine Chem. *171*, 92–96.

Murshudov, G.N., Skubák, P., Lebedev, A.A., Pannu, N.S., Steiner, R.A., Nicholls, R.A., Winn, M.D., Long, F., and Vagin, A.A. (2011). REFMAC5 for the refinement of macromolecular crystal structures. Acta Crystallogr. D Biol. Crystallogr. *67*, 355–367.

Muthana, S., Cao, H., and Chen, X. (2009). Recent progress in chemical and chemoenzymatic synthesis of carbohydrates. Curr. Opin. Chem. Biol. *13*, 573–581.

N'go, I., Golten, S., Arda, A., Canada, J., Jimenez-Barbero, J., Linclau, B., and Vincent, S.P. (2014). Tetrafluorination of sugars as strategy for enhancing protein-carbohydrate affinity: application to UDP-Galp mutase inhibition. Chem. Eur. J. *20*, 106–112.

O'hagan, D. (2008). Understanding organofluorine chemistry. An introduction to the C–F bond. Chem. Soc. Rev. *37*, 308–319.

O'neill, E.C., and Field, R.A. (2015). Enzymatic synthesis using glycoside phosphorylases. Carbohydr. Res. *403*, 23–37.

Plante, O.J., Palmacci, E.R., Andrade, R.B., and Seeberger, P.H. (2001). Oligosaccharide synthesis with glycosyl phosphate and dithiophosphate triesters as glycosylating agents. J. Am. Chem. Soc. *123*, 9545–9554.

Platt, A., Ross, H.C., Hankin, S., and Reece, R.J. (2000). The insertion of two amino acids into a transcriptional inducer converts it into a galactokinase. Proc. Natl. Acad. Sci. U S A *97*, 3154–3159.

Potterton, L., Agirre, J., Ballard, C., Cowtan, K., Dodson, E., Evans, P.R., Jenkins, H.T., Keegan, R., Krissinel, E., and Stevenson, K. (2018). CCP4i2: the new graphical user interface to the CCP4 program suite. Acta Crystallogr. D Struct. Biol. *74*, 68–84.

Schell, M., and Wilson, D. (1977). Purification and properties of galactokinase from *Saccharomyces cerevisiae*. J. Biol. Chem. *252*, 1162–1166.

Schmidt, M.F., Biely, P., Kratky, Z., and Schwartz, R.T. (1978). Metabolism of 2-deoxy-2-fluoro-d-[^{3}H] glucose and 2-deoxy-2-fluoro-d-[^{3}H] mannose in yeast and chick-embryo cells. Eur. J. Biochem. *87*, 55–68.

Thoden, J.B., and Holden, H.M. (2003). Molecular structure of galactokinase. J. Biol. Chem. *278*, 33305–33311.

Thomas, P., Bessell, E.M., and Westwood, J.H. (1974). The use of deoxy-fluoro-D-galactopyranoses in a study of yeast galactokinase specificity. Biochem. J. *139*, 661–664.

Timofte, R.S., and Linclau, B. (2008). Enantioselective synthesis of tetrafluorinated glucose and galactose. Org. Lett. *10*, 3673–3676.

Tiruchinapally, G., Yin, Z., El-Dakdouki, M., Wang, Z., and Huang, X. (2011). Divergent heparin oligosaccharide synthesis with preinstalled sulfate esters. Chem. Eur. J. *17*, 10106–10112.

Van Straaten, K.E., Kuttiyatveetil, J.R., Sevrain, C.M., Villaume, S.A., Jimenez-Barbero, J., Linclau, B., Vincent, S.P., and Sanders, D.A. (2015). Structural basis of ligand binding to UDP-galactopyranose mutase from *Mycobacterium tuberculosis* using substrate and tetrafluorinated substrate analogues. J. Am. Chem. Soc. *137*, 1230–1244.

A. Varki, R.D. Cummings, J.D. Esko, P. Stanley, G.W. Hart, M. Aebi, A.G. Darvill, T. Kinoshita, N.H. Packer, and J.H. Prestegard, et al., eds. (2015-2017). Essentials of Glycobiology [Internet], 3rd edition (Cold Spring Harbor Laboratory Press).

Wang, J., Sánchez-Roselló, M., Aceña, J.L., del Pozo, C., Sorochinsky, A.E., Fustero, S., Soloshonok, V.A., and Liu, H. (2013). Fluorine in pharmaceutical industry: fluorine-containing drugs introduced to the market in the last decade (2001-2011). Chem. Rev. *114*, 2432–2506.

Wang, K.-C., Lyu, S.-Y., Liu, Y.-C., Chang, C.-Y., Wu, C.-J., and Li, T.-L. (2014). Insights into the binding specificity and catalytic mechanism of N-acetylhexosamine 1-phosphate kinases through multiple reaction complexes. Acta Cryst. D. *70*, 1401–1410.

Waterman, D.G., Winter, G., Gildea, R.J., Parkhurst, J.M., Brewster, A.S., Sauter, N.K., and Evans, G. (2016). Diffraction-geometry refinement in the DIALS framework. Acta Crystallogr. D Struct. Biol. *72*, 558–575.

White, A., Tull, D., Johns, K., Withers, S.G., and Rose, D.R. (1996). Crystallographic observation of a covalent catalytic intermediate in a β-glycosidase. Nat. Struct. Biol. *3*, 149–154.

Xia, J., Xue, J., Locke, R.D., Chandrasekaran, E., Srikrishnan, T., and Matta, K.L. (2006). Synthesis of fluorinated mucin core 2 branched oligosaccharides with the potential of novel substrates and enzyme inhibitors for glycosyltransferases and sulfotransferases. J. Org. Chem. *71*, 3696–3706.

Yan, J., Chen, X., Wang, F., and Cao, H. (2013). Chemoenzymatic synthesis of mono-and di-fluorinated Thomsen-Friedenreich (T) antigens and their sialylated derivatives. Org. Biomol. Chem. *11*, 842–848.

Yang, J., Fu, X., Jia, Q., Shen, J., Biggins, J.B., Jiang, J., Zhao, J., Schmidt, J.J., Wang, P.G., and Thorson, J.S. (2003). Studies on the substrate specificity of *Escherichia coli* galactokinase. Org. Lett. *5*, 2223–2226.

Yang, J., Fu, X., Liao, J., Liu, L., and Thorson, J.S. (2005). Structure-based engineering of *E. coli* galactokinase as a first step toward in vivo glycorandomization. Chem. Biol. *12*, 657–664.

Yang, J., Liu, L., and Thorson, J.S. (2004). Structure-based enhancement of the first anomeric glucokinase. ChemBioChem *5*, 992–996.

Yu, H., Thon, V., Lau, K., Cai, L., Chen, Y., Mu, S., Li, Y., Wang, P.G., and Chen, X. (2010). Highly efficient chemoenzymatic synthesis of β1-3-linked galactosides. Chem. Commun. *46*, 7507–7509.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and Virus Strains		
One Shot TOP10 chemically competent E. coli	Invitrogen	Cat# 10368022
BL21(DE3) competent E. coli	NEB	Cat# C2527I
Chemicals, Peptides, and Recombinant Proteins		
EcGalK	This paper, Prozomix Limited	N/A
PfGalK	This paper, Prozomix Limited	N/A
SpGalK	This paper, Prozomix Limited	N/A
LlGalK	This paper, Prozomix Limited	N/A
LgGalK	This paper, Prozomix Limited	N/A
BiGalK	This paper, Prozomix Limited	N/A
ScGalK	This paper, Fascione Lab	N/A
BliNahK	This paper, Prozomix Limited	N/A
BllNahK	This paper, Prozomix Limited	N/A
ApNahK	This paper, Prozomix Limited	N/A
CxNahK	This paper, Prozomix Limited	N/A
EtNahK	This paper, Prozomix Limited	N/A
TbNahK	This paper, Prozomix Limited	N/A
2-deoxy-2-fluoro-D-galactose (1a)	Carbosynth	Cat# MD04718
3-deoxy-3-fluoro-D-galactose (2a)	Carbosynth	Cat# MD05336
4-deoxy-4-fluoro-D-galactose (3a)	Carbosynth	Cat# MD05334
6-deoxy-6-fluoro galactose (4a)	Carbosynth	Cat# MD04719
2-deoxy-2-fluoro-D-glucose (5a)	Carbosynth	Cat# MD03509
3-deoxy-3-fluoro-D-glucose (6a)	Carbosynth	Cat# MD05769
4-deoxy-4-fluoro-D-glucose (7a)	Carbosynth	Cat# MD04720
6-deoxy-6-fluoro-D-glucose (8a)	Carbosynth	Cat# MD042558
2-deoxy-2-fluoro-D-mannose (9a)	Carbosynth	Cat# MD03540
D-Galactose	Sigma-Aldrich	Cat# G0750
β-Nicotinamide adenine dinucleotide, reduced dipotassium salt (β-NADH)	Sigma-Aldrich	Cat# N4505
Pyruvate kinase from rabbit muscle	Sigma-Aldrich	Cat# P1506
L-lactate dehydrogenase	Sigma-Aldrich	Cat# SAE0049
Phosphoenol pyruvate	Sigma-Aldrich	Cat# 10108294001
Deposited Data		
BiGalK (ADP complex)	This paper	PDB 6TEP
BiGalK (Gal complex)	This paper	PDB 6TER
BiGalK (2FGal complex)	This paper	PDB 6TEQ
Recombinant DNA		
EcGalK_pET28a	This paper, Prozomix Limited	N/A
PfGalK_pET28a	This paper, Prozomix Limited	N/A
SpGalK_pET28a	This paper, Prozomix Limited	N/A
LlGalK_pET28a	This paper, Prozomix Limited	N/A
LgGalK_pET28a	This paper, Prozomix Limited	N/A
BiGalK_pET22b	This paper, Fascione Lab	N/A
ScGalK_pET15b	Keenan et al., 2019	N/A

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
BliNahK_pET28a	This paper, Prozomix Limited	N/A
BllNahK_pET28a	This paper, Prozomix Limited	N/A
ApNahK_pET28a	This paper, Prozomix Limited	N/A
CxNahK_pET28a	This paper, Prozomix Limited	N/A
EtNahK_pET28a	This paper, Prozomix Limited	N/A
TbNahK_pET28a	This paper, Prozomix Limited	N/A
Software and Algorithms		
GraphPad Prism v. 7.05	GraphPad Software Inc.	https://www.graphpad.com/scientific-software/prism/
DIALS	Waterman et al., 2016	https://dials.github.io/
CCP4 suite	Potterton et al., 2018	http://www.ccp4.ac.uk/
Coot	Emsley et al., 2010	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/

RESOURCE AVAILABILITY

Lead Contact

Further information and requests for resources and reagents should be directed to the Lead Contact, Martin A. Fascione (martin.fascione@york.ac.uk).

Materials Availability

Plasmids used for the expression of BiGalK and ScGalK in this study can be requested from the Lead contact via a materials transfer agreement.

Data and Code Availability

The PDB files that support the findings of this study have been deposited in the Protein Data Bank under accession codes 6TEP, 6TEQ and 6TER. The Genbank accession numbers for the new NahKs investigated in this study are WP_091282512 (ApNahK), WP_045094128 (CxNahK), ODM08532 (EtNahK) and WP_062612735 (TbNahK).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Bacterial Strains

Chemically competent Escherichia coli (*E. coli*) Top10 cells (F- mcrA Δ(mrr-hsdRMS-mcrBC) ϕ80lacZΔM15 ΔlacX74 nupG recA1 araD139 Δ(ara-leu)7697 galE15 galK16 rpsL(Str^R) endA1 λ$^-$) were used as a cloning host in this study. For protein expression, chemically competent *E. coli* BL21(DE3) cells (B F$^-$ *ompT gal dcm lon hsdS$_B$*($r_B^-m_B^-$) λ(DE3 [*lacI lacUV5-T7p07 ind1 sam7 nin5*]) [*malB*$^+$]$_{K-12}$(λ^S)) were used. The bacterial cells were grown on LB agar at 37°C or in liquid LB medium at 37C with shaking at 180 rpm.

METHOD DETAILS

General Methods

All chemicals and solvents were purchased from Sigma-Aldrich, AlfaAesar or Fluorochem and used without further purification, unless otherwise stated. Compounds **1a-9a** were kindly supplied by Carbosynth (Compton, UK). All enzymes except for BiGalK and ScGalK were produced and provided by Prozomix (Haltwhistle, UK). Samples for HRMS were diluted in HPLC grade MeOH/H_2O and analysed either on a Bruker APEX III FT-ICR instrument or on a Bruker Daltonics microTOF instrument.

Chemical Synthesis

Compounds **10a** (Malassis et al., 2019), **12a** (Timofte and Linclau, 2008), **13a** (Malassis et al., 2019), **11a** (Mtashobya et al., 2015), **15a** (Golten et al., 2016) and **16a** (Fontenelle et al., 2017) were synthesized following procedures previously described and their spectroscopic data are in agreement with those previously reported. The synthesis of **14a**, **10b** and **17a** are detailed in Data S2.

Determination of Fluorosugar Substrate Specificity

Reactions were carried out in Tris buffer (50 μL, 100 mM, pH 8.0) at 37°C for 2 h (Gal) or 24 h (**1a-17a**). For all compounds, substrate specificity was determined in reactions containing monosaccharide (8 mM), ATP (10 mM), $MgCl_2$ (5 mM) and GalK or NahK (6 μg, final

kinase conc. 0.12 mg/mL). The presence of the sugar-1-phosphate product was determined by HRMS. The conversion of Gal was determined by the DNS assay. The conversions of the fluorosugars were measured by ^{19}F NMR.

Optimisation of Substrate Conversion

For optimised conversions of monosaccharide substrates (**2a-5a**, **7a**, **10a-11a**, **13a-14a**, **17a**) reactions were carried out with monosaccharide (5-10 mM), ATP (5-10 mM), $MgCl_2$ (5 mM), higher enzyme loading (90 μg BiGalK, 55 μg ScGalK, 15 μg EcGalK, 31 μg LgGalK, 31 μg LlGalK, 60 μg SpGalK, 32 μg PfGalK, 68 μg BliNahK, 37 μg BllNahK, 13 μg ApNahK, 12 μg CxNahK, 20 μg EtNahK, 10 μg TbNahK), and with prolonged incubation times (48-72 h). The highest conversions are reported in Tables S1.

^{19}F NMR Analysis

Samples (50 μL) were diluted with MeOH/H_2O (1:1, 450 μL) and centrifuged (10 000 x *g*, 5 min) to remove any insoluble matter. The solution was transferred to a 5 mm NMR tube, with either a 3 mm coaxial insert or a sealed glass capillary tube containing D_2O for locking and referencing. ^{19}F NMR spectra were recorded at 25°C on a Bruker Avance Neo 700 MHz spectrometer (operating at 658.78 MHz) equipped with a nitrogen-cooled TCI cryoprobe or on a Bruker Avance 500 MHz spectrometer (operating at 470 MHz) equipped with a QCI-F cryoprobe. Spectra were referenced through the D_2O lock (^{2}H) signal according to IUPAC recommended secondary referencing method, as implemented within TopSpin version 4.0.2 (Harris et al., 2002). A 30 degree pulse was used, with an acquisition time of 0.6 s and a recycle delay of 1.4 s. Baseline correction was applied to compensate for broad background signals from the probe. Representative spectra are provided in the dedicated section. Conversions were determined by taking the relative integration of the corresponding resonances of the starting material and product, which were typically baseline-separated.

DNS Assay

The DNS assay was carried out as described previously (Yang et al., 2003). A sample aliquot (50 μL) was added to of DNS reagent (100 μL). Mixtures were heated at 100°C for 15 min, 4°C for 1 min and 20°C for 3 min in a PCR thermocycler, 15 μL of 40 % *w/v* potassium tartrate solution was added to each sample and the absorbance at 575 nm was recorded. Standard curves were prepared by using a series of sugar concentrations (0.5, 1.0, 2.0, 2.5, 3.0, 3.5, 4.0, 5.0, 6.0, 7.0, and 8.0 mM respectively) and ATP (10 mM) in Tris buffer (100 mM, pH 8.0) and submitting them to the DNS assay.

BiGalK Protein Production and Purification

A plasmid encoding the GalK gene from *Bifidobacterium longum* subsp. *infantis* (GenBank accession number: CEF10481.1) cloned into the NdeI and HindIII restriction sites of pET22b, was ordered from Genscript. The plasmid was introduced into chemically competent *E. coli* BL21(DE3) cells by heat shock and selected on LB agar with ampicillin (100 μg/mL) at 37°C for 16 h. Starter cultures were prepared by picking single clones into LB with ampicillin (100 μg/mL) and grown at 37°C for 8 h with shaking (180 rpm). 1 mL of starter culture was added to 1L of LB with ampicillin (100 μg/mL) and grown to an OD_{600} of 0.6 – 0.7 at 37°C with shaking (180 rpm). Cultures were placed in an ice bath for 15 min. IPTG was added to a final concentration of 0.2 mM and the cultures were grown at 16°C for 20 h with shaking (180 rpm). Cells were harvested by centrifugation (6000 x g, 6°C, 10 min) and the pellet was resuspended in lysis buffer (30 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mg/mL lysozyme, protease inhibitor, benzonase 1 U/mL). The cells were lysed by sonication on ice and the lysate was clarified by centrifugation (18 000 rpm, 4°C, 50 min). The lysate was loaded onto a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated with binding buffer (30 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole). After washing the column with 20 column volumes (CV) of binding buffer, recombinant BiGalK with a C-terminal His_6-tag was eluted with elution buffer (30 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole). The protein was desalted and exchanged into 30 mM Tris pH 8.0 buffer (substrate specificity assays) or 10 mM HEPES pH 6.5, 0.15 M NaCl buffer (crystallography) using a HiPrep™ 26/10 Desalting column (GE Healthcare). Further purification was carried out by using gel filtration chromatography (HiLoad 16/600 Superdex 200 pg) in either 30 mM Tris pH 8.0 buffer (substrate specificity assays) or 10 mM HEPES pH 6.5, 0.15 M NaCl buffer (crystallography). Typical yield was ~40 mg/L culture. Protein purity was assessed by SDS-PAGE (Figure S4A) and shown to be consistent with the predicted molecular mass (45.8 kDa) for the enzyme. Enzyme activity towards galactose was validated by the DNS assay.

ScGalK Protein Production and Purification

ScGalK protein production and purification was carried out as described previously (Keenan et al., 2019). Briefly, the SCO3136_-pET15b plasmid was introduced into chemically competent E. coli BL21(DE3) cells by heat shock and selected on LB agar with ampicillin (100 μg/mL) at 37 °C for 16 h. Starter cultures were prepared by picking single clones into LB with ampicillin (100 μg/mL) and grown at 37 °C for 6 h with shaking (180 rpm). Starter cultures were diluted 1/100 into LB with ampicillin (100 μg/mL) and grown to an OD_{600} of 0.8 at 37 °C with shaking (180 rpm). After incubation on ice (10 min) IPTG was added to a final concentration of 0.8 mM and the cultures were grown at 16 °C for 22 h with shaking (180 rpm). Cells were harvested by centrifugation (6000×g, 6 °C, 10 min) and resuspended in lysis buffer (30 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole, 1 mg/mL lysozyme protease inhibitor, benzonase 1 U/mL). The cells were lysed by sonication on ice and further centrifuged (18 000 rpm, 4 °C, 45 min). The lysate was loaded onto a 5 mL HisTrap FF column (GE Healthcare) pre-equilibrated with binding buffer (30 mM Tris pH 8.0, 300 mM NaCl, 20 mM imidazole). After washing the column (20 CV binding buffer) recombinant ScGalK with an N-terminal His_6-tag was eluted with elution buffer (30 mM Tris pH 8.0, 300 mM NaCl, 250 mM imidazole). The protein was desalted and exchanged into 30 mM Tris pH 8.0 buffer using

HiPrep™ 26/10 Desalting column (GE Healthcare). Purified ScGalK was stored at −80 °C in 10% (v/v) glycerol. Protein purity was assessed by SDS-PAGE (Figure S4B). Enzyme activity towards galactose was validated by the DNS assay.

Molecular cloning and protein production for EcGalK, LlGalK, PfGalK, SpGalK, LgGalK, BllNahK, BliNahK, ApNahK, CxNahK, EtNahK and TbNahK (performed by Prozomix)

DNA sequences encoding the GalKs/NahKs (codon optimised for expression in *E. coli*) with an N-terminal NdeI site and a C-terminal XhoI site (following the stop codon), were ordered from Twist Biosciences. After double digestion with NdeI and XhoI endonucleases to obtain sticky ends, the genes were ligated into an NdeI and XhoI double-digested pET28a vector. The resulting plasmid was introduced into chemically competent *E. coli* Top10 cells by heat shock and selected on LB agar with kanamycin (35 μg/mL) at 37°C for 16 h. Single colonies were picked single into LB with kanamycin (35 μg/mL) and grown at 37°C for 16 h with shaking (180 rpm). Plasmid DNA was isolated by using the QIAprep spin Miniprep Kit (QIAGEN) and insertion of the correct insert was validated by DNA sequencing. The plasmids were introduced into chemically competent *E. coli* BL21(DE3) cells (galactosidase free strain) by heat shock and selected and selected on LB agar with kanamycin (35 μg/mL) at 37°C for 16 h. Starter cultures were prepared by picking single clones into LB with kanamycin (35 μg/mL) and grown at 37°C for 16 h with shaking (180 rpm). Starter cultures were used to inoculated 750 mL of TBp autoinducing media and grown at 28°C for 24 h with shaking (150 rpm). Cells were harvested by centrifugation (8000 x g, 4°C, 5 min) and resuspended in 500 mM NaCl, 10 mM HEPES buffer (NaCl/HEPES buffer) containing 10 mM Imidazole at pH 7.5 (5 mL/g of cells). Cells were lysed at <5°C by sonication (4 x 1 min per 50mL of resuspended cells volume) and the lysate was clarified by centrifugation (11 000 rpm, 4°C, 30 min). The lysate was loaded onto a 7 mL nickel sepharose column pre-equilibrated with NaCl/HEPES buffer + 10 mM Imidazole. After washing the column with 5 column volumes (CV) of NaCl/HEPES buffer + 10 mM Imidazole and then 5 CV of NaCl/HEPES buffer + 50 mM Imidazole, the recombinant proteins were eluted using NaCl/HEPES buffer + 500 mM Imidazole at pH 6.75 with frequent checking of 10 mL aliquots of eluant using Bradford reagent until the blue colour response was no longer apparent. The recombinant protein was precipitated with 0.5g solid ammonium sulphate per mL of elution and stored at 4°C. A representative SDS-PAGE analysis of the samples is shown in Figure S4C. Before use, an aliquot was centrifuged, the supernatant discarded and the pellet redissolved in Tris buffer (100 mM, pH 8.0). Protein concentration was determined according to the Bradford assay.

Determination of BiGalK Kinetic Parameters

Galactokinase activity was determined by coupling the formation of ADP to the reactions catalysed by pyruvate kinase and lactate dehydrogenase and the consumption of NADH, as previously described (Platt et al., 2000). Reactions (150 μL) were assembled in microtitre plates containing 50 mM HEPES buffer (pH 7.5), 150 mM NaCl, 5 mM $MgCl_2$, 1 mM phosphoenol pyruvate, 1 mM NADH, 7.5 U pyruvate kinase and 10 U lactate dehydrogenase. Reactions were supplemented with various concentrations of Gal (0.125 – 6 mM) or 2FGal (1 - 8 mM) under saturating concentrations of ATP (10 mM). All reaction components were added to the microtitre plate except for the enzyme, mixed and preincubated at 37°C for 5 min. Reactions were initiated by the addition of BiGalK (5 nM) and the decrease in absorbance at 340 nm was measured every 30 s for 5 min. The initial velocity was calculated by determining the linear phase in the progress curve. Assays were assembled in triplicate. The data were fitted to the Michaelis-Menten equation and the apparent kinetic parameters were calculated by using GraphPad Prism v. 7.05. The error values are shown as standard error of the mean.

Crystallization, Data Collection and Structure Solution

Initial crystallization conditions were identified using commercially available screens from Hampton and Molecular Dimension in a 96-well sitting-drop screening format. Further optimization in a 24-well hanging-drop or 48-well sitting-drop format provided suitable conditions for reliable crystallization. Optimal crystals were reliably obtained by micro-seeding with previously obtained crystals. Crystal formation was only observed in the presence of ADP, despite the absence of this co-factor from the Gal- and 2FGal-bound complexes. For BiGalK in complex with ADP and $MgCl_2$, crystals were obtained at 20°C using the hanging drop vapour diffusion method by mixing 1 μL of a protein solution (10-17 mg/mL of BiGalK, 5 mM $MgCl_2$ and 10 mM ADP in 10 mM HEPES pH 6.5, 0.15 M NaCl buffer) with 1 μL of reservoir solution (0.1 M MES pH 6.0, 0.15 - 0.25 M NaCl and 18 - 22 % (w/v) polyethylene glycol 6000). For BiGalK in complex with Gal, crystals were obtained at 20°C using the sitting drop vapour diffusion method by mixing 1 μL of the same protein solution supplemented with 10 mM galactose with 1 μL of the same reservoir as used previously. For soaking, 1 μL of 100 mM Gal dissolved in reservoir solution was added to the drop containing the crystal and incubated at 20°C for 6 h. For BiGalK in complex with 2FGal, crystals were obtained at 20°C using the sitting drop vapour diffusion method by mixing 1 μL of a protein solution (17 or 10 mg/mL of BiGalK, 5 mM MgCl2, 10 mM ADP and 10 mM 2FGal in 10 mM HEPES pH 6.5, 0.15 M NaCl buffer) with 1 μL of reservoir solution (0.1 M MES pH 6.0, 0.25 M NaCl and 22 % (w/v) polyethylene glycol 6000). For soaking, 1 μL of 100 mM 2FGal dissolved in reservoir solution was added to the drop containing the crystal and incubated at 20°C for 6 h. All crystals were fished via a cryoprotectant, which comprised the reservoir solution supplemented with 10 % (v/v) glycerol, into liquid nitrogen. Data were collected at beamline IO4.1, Diamond Light Source (Didcot, UK). The data were processed using DIALS (Waterman et al., 2016) and scaled using AIMLESS (Evans and Murshudov, 2013). The structures were solved using the CCP4 suite of programs (Potterton et al., 2018). The structure of BiGalK in complex with ADP was solved by Molecular Replacement with *Phaser* (McCoy et al., 2007) using amino acids 30 – 359 from the crystal structure of the *L. lactis* GalK (Thoden and Holden, 2003) (PDB: 1PIE, 34.9 % identity) as the search model. For the structures of BiGalK in complex with Gal and 2FGal, solution was achieved by Molecular

Replacement with *Phaser* using the solved crystal structure of BiGalK in complex with ADP and Mg^{2+} as a starting model. The models were improved by iterative cycles of manual model building in *Coot* (Emsley et al., 2010), followed by refinement with *REFMAC* (Murshudov et al., 2011), including cycles using anisotropic B-factor refinement during the final cycles of refinement. Problems with diffraction anisotropy in the dataset limited the possibility of refining the structure to R/Rfree lower than 0.22/0.31 for the complex with Gal, however the collected dataset resulted in a good quality $2F_o-F_c$ map, with clearly defined amino acid sidechains and the ligand easily identifiable in the difference map. Data collection and refinement statistics are reported in Table S1.

QUANTIFICATION AND STATISTICAL ANALYSIS

All kinetic experiments were carried out in triplicate (n=3), and data fitting and statistical analysis was performed using GraphPad Prism version 7.05. For X-ray structures, we monitored the R_{free} during model building into electron density and used the Ramachandran plot to fix any outliers before submitting structures to the PDB/RCSB (https://www.rcsb.org/). All the crystallographic statistics for the structure and data are shown in Table S1. The ''Full wwPDB X-ray Structure Validation Report'' for each structure is available with each PDB code in the database. For the ^{19}F NMR analysis of kinase-mediate phosphorylation of fluorinated monosaccharides, reaction conversions were determined by taking the relative integration of the corresponding resonances of the starting material and product, which were typically baseline-separated.