JACS Au
AN OPEN ACCESS JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

pubs.acs.org/jacsau
Article

Rational Design of Dual-Domain Binding Inhibitors for *N*-Acetylgalactosamine Transferase 2 with Improved Selectivity over the T1 and T3 Isoforms

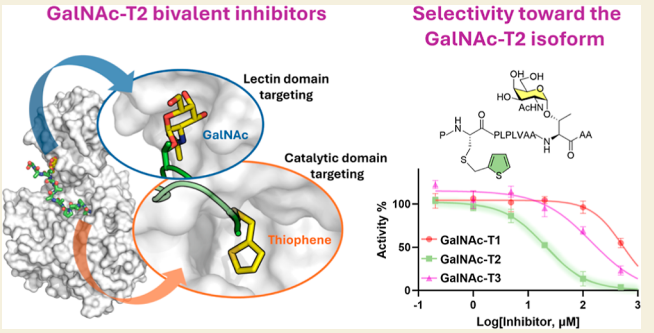
Ismael Compañón, Collin J. Ballard, Erandi Lira-Navarrete, Tanausú Santos, Serena Monaco, Juan C. Muñoz-García, Ignacio Delso, Jesus Angulo, Thomas A. Gerken, Katrine T. Schjoldager, Henrik Clausen, Tomás Tejero, Pedro Merino, Francisco Corzana, Ramon Hurtado-Guerrero, and Mattia Ghirardello*

Cite This: *JACS Au* 2024, 4, 3649−3656

Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: The GalNAc-transferase (GalNAc-T) family, consisting of 20 isoenzymes, regulates the *O*-glycosylation process of mucin glycopeptides by transferring GalNAc units to serine/threonine residues. Dysregulation of specific GalNAc-Ts is associated with various diseases, making these enzymes attractive targets for drug development. The development of inhibitors is key to understanding the implications of GalNAc-Ts in human diseases. However, developing selective inhibitors for individual GalNAc-Ts represents a major challenge due to shared structural similarities among the isoenzymes and some degree of redundancy among the natural substrates. Herein, we report the development of a GalNAc-T2 inhibitor with higher potency compared to those of the T1 and T3 isoforms. The most promising candidate features bivalent GalNAc and thiophene moieties on a peptide chain, enabling binding to both the lectin and catalytic domains of the enzyme. The binding mode was confirmed by competitive saturation transfer difference NMR experiments and validated through molecular dynamics simulations. The inhibitor demonstrated an IC_{50} of 21.4 μM for GalNAc-T2, with 8- and 32-fold higher selectivity over the T3 and T1 isoforms, respectively, representing a significant step forward in the synthesis of specific GalNAc-T inhibitors tailored to the unique structural features of the targeted isoform.



KEYWORDS: *N-acetylgalactosamine transferase, glycosyltransferase, inhibitor, glycopeptide, GalNAc, molecular dynamics, STD NMR*

INTRODUCTION

Glycosylation is one of the most complex and fundamental types of post-translational modification (PTM) occurring in cells,[1] and it plays a critical role in structural, metabolic, and regulatory cell functions.[2] This process is orchestrated by glycosyltransferase enzymes, and its dysregulation has emerged as a hallmark of numerous diseases such as cancer, where the aberrant glycosylation is implicated in tumor progression, immune evasion, and metastasis.[3−7] Mucin-type *O*-glycosylation stands out as one of the most abundant and diverse types of protein *O*-glycosylation. It normally targets serine (Ser) and threonine (Thr) residues and is characterized by the display of complex glycans on tandem repeats of mucin proteins.[8] The initial step in mucin *O*-glycosylation involves the transfer of an *N*-acetylgalactosamine (GalNAc) unit from uridine diphosphate-GalNAc (UDP-GalNAc) donors to Ser/Thr residues of acceptor substrates. In humans, this process is achieved by a family of up to 20 GalNAc-transferase (GalNAc-Ts) isoenzymes in an organized manner.[9] These GalNAc-Ts are localized in the Golgi apparatus and are distributed with varying expression levels across all organs and tissues.[10]

In the past 15 years, numerous studies have revealed specific correlations between the dysregulation of various GalNAc-Ts and the progression of human diseases. For instance, GalNAc-T3 has been implicated in tumoral calcinosis,[11] where it plays a coregulatory role in phosphate homeostasis.[12] GalNAc-T11 has been associated with chronic kidney decline due to its role in the glycosylation of low-density lipoprotein receptors and related receptors, including LRP2.[13−15] GalNAc-T6 overexpression showed to promote pancreatic[16] and colorectal cancer[17] development, while GalNAc-T7 upregulation corre-

Received: July 15, 2024
Revised: August 16, 2024
Accepted: September 4, 2024
Published: September 11, 2024

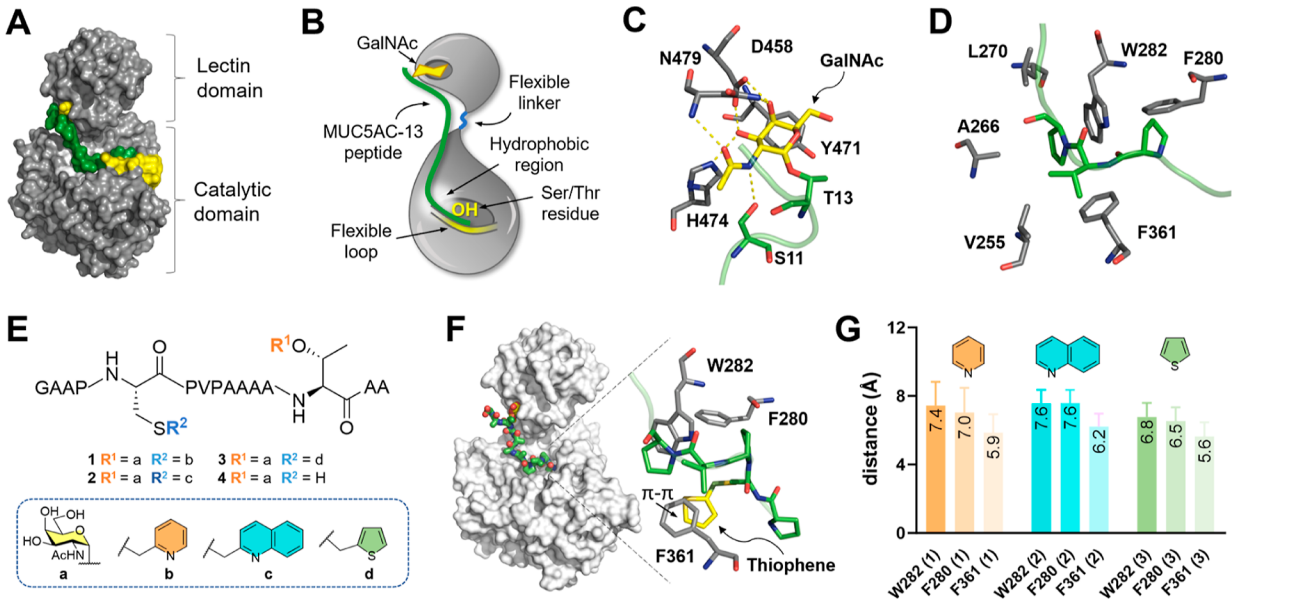


Figure 1. (A) Structure of GalNAc-T2 in complex with glycopeptide MUC5AC-13 and UDP. (B) Schematic representation of the GalNAc-T2 regions and of the binding mode with MUC5AC-13 glycopeptide. (C) Binding mode of the MUC5AC-13 GalNAc moiety (yellow) with GalNAc-T2 lectin pocket (gray). (D) Binding mode of MUC5AC-13 PVP residues (green) with the hydrophobic region of GalNAc-T2 (gray). PDB ID: 5AJP. (E) Chemical structure of inhibitors **1**−**4**. (F) Representative frame derived from 1.0 μs molecular dynamics (MD) simulation of compound **3** with GalNAc-T2 and expansion on the interactions between the thiophene moiety and residues W282, F280, and F361. (G) Average distance between the aromatic moieties of **1**, **2**, and **3** with the residues W282, F280, and F361 of GalNAc-Ts calculated through 1.0 μs MD simulations; data represent mean + SD (standard deviation).

lated with tumor growth progression in prostate cancer patients.[18]

Recently, through GWAS and animal knockout models, GalNAc-T2 dysfunction has been implicated in cholesterol and triglyceride metabolism.[19,20] Additionally, through its role in the GalNAcylation process of the insulin receptor, GalNAc-T2 has been identified as a key modulator of insulin signaling.[21] Furthermore, lack of GalNAc-T2 expression has been established as a congenital disorder of glycosylation (GALNT2-CDG) linked to neurodevelopmental delays, and to the insurgence of brain abnormalities.[20]

These findings underscore the high value of GalNAc-Ts as promising drug targets under severe human conditions. However, the development of novel drugs targeting specific GalNAc-Ts represents a formidable challenge due to the difficulty in achieving potent and selective inhibitor against the targeted isoform.[22−24] This is in contrast to protein phosphorylation, the most frequent type of PTM, which has seen significant success in drug development, with more than 70 approved drugs since 2001.[25] All GalNAc-T isoenzymes utilize UDP-GalNAc as an activated sugar-nucleotide donor, and most of the family members showed some degree of redundancy in their protein and glycoprotein acceptor substrates.[10,26] Therefore, inhibitors based on the sole structure of the donor or acceptor substrates are likely to exhibit promiscuous activity on a broad set of GalNAc-Ts. Several attempts into proving GalNAc-Ts druggability have been made, from the screening of libraries[24,27−29] to the specific design of uridine derivatives.[24,27,30,31] However, the high degree of similarity between GalNAc-T isoforms limited the development of specific inhibitors. The exception being T3Inh-1 that can act against GalNAc-T3 and showed no activity against the T2 and T6 variants, where it was used to lower the FGF23 levels in mice for chronic kidney disease treatment.[28] Bioengineering approaches also demonstrated to be successful for selective GalNAc-Ts inhibition. For instance, engineered GalNAc-Ts through the "bump-and-hole" strategy using sterically hindered UDP-GalNAc variants has been reported.[32,33] However, this strategy cannot be employed with nonengineered GalNAc-Ts in natural settings.

In an attempt to achieve a selective inhibition toward GalNAc-T2, we have prepared a series of glycopeptides rationally designed on the basis of structural knowledge gathered on this GalNAc-T isoform.[34−37] All GalNAc-Ts enzymes possess a lectin domain responsible for recognition and binding to a previously glycosylated position of the glycopeptide acceptor substrate (Figure 1A,B). This domain is connected through a flexible linker to the catalytic domain, which binds the UDP-GalNAc donor substrate and transfers the activated sugar unit to hydroxyl residues of the acceptor substrate. A flexible loop present nearby the catalytic pocket promotes the binding of the donor substrate, and its dynamics regulate the switch between active and inactive conformation of the enzyme.[34] GalNAc-Ts can glycosylate in a lectin-independent or lectin-dependent manner. The latter can be classified into two main subfamilies depending on the preference to glycosylate a position within a short-range (1 to 3 residues) or long-range (6 to 17 residues) to a previously glycosylated position.[10,38] The dynamics of the interdomain flexible linker and the enzyme electrostatic surface potential, dictate the preference to glycosylate long- or short-range positions toward the N- or C-terminus direction of the acceptor chain.[36,39] These structural features are likely used by the GalNAc-T family to promote mucin glycosylation in a hierarchical and organized manner. Previous studies reported the preference of GalNAc-T3/T4/T6/T12 to glycosylate long-range C-terminal positions from a prior N-terminal GalNAcylated residue. On the contrary, the GalNAc-T2/T14 isoforms

https://doi.org/10.1021/jacsau.4c00633
JACS Au 2024, 4, 3649−3656

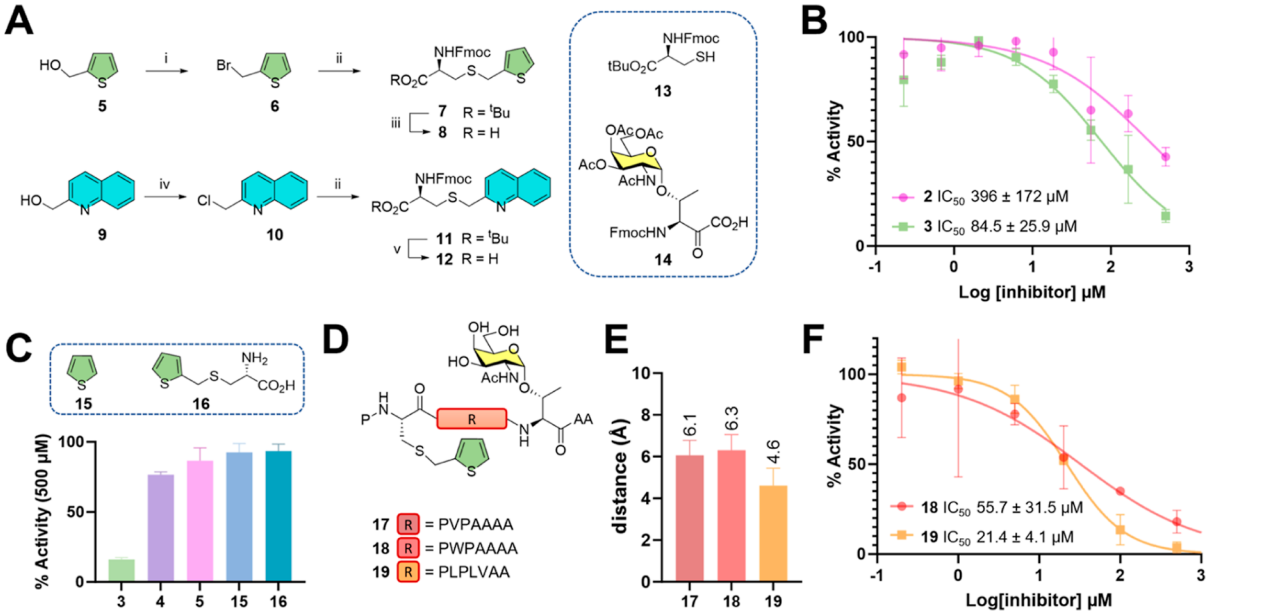


Figure 2. (A) Synthetic approach for the synthesis of building blocks **8** and **12**. Reagent and conditions: (i) HBr (33% in AcOH), Et_2O, 0 °C to rt, 16 h, 96%; (ii) **13**, $NaHCO_3$, Bu_4NBr, H_2O, EtOAc, 24 h, rt, 82%; (iii) EDT, TFA, DCM, 0 °C to rt, 1 h, 81%; (iv) $SOCl_2$, DCM, 0 °C to rt, 1 h, 85%; and (v) TFA, DCM, 0 °C to rt, 2 h, 87%. (B) Relative inhibitory activity % of compounds **2** and **3** against GalNAc-T2 using MUC1a as acceptor substrate. (C) Relative inhibitory activity % of compounds **3**, **4**, **5**, **15**, and **16** at a normalized concentration of 500 μM against GalNAc-T2 using MUC1a as acceptor substrate; data represent mean + SD. (D) Chemical structure of inhibitors **17**−**19**. (E) Average distance between the thiophene moiety of **17**, **18**, and **19** with residue F361 of GalNAc-Ts calculated through 1.0 μs MD simulations; data represent mean + SD (standard deviation). (F) Relative inhibitory activity % of compounds **18** and **19** against GalNAc-T2 using MUC1a as acceptor substrate.

preferably glycosylate toward the N-terminal direction,[26] and GalNAc-T1 is capable of glycosylating with both preferences.[40]

In order to achieve selectivity over the target T2 isoform, we devised the synthesis of chemically modified bivalent glycopeptides capable of targeting both the lectin and catalytic domain of the enzyme at the same time. This family of inhibitors features a GalNAc moiety at the C-terminus, which enables binding to the lectin domain as for natural substrates (Figure 1C), and an aromatic moiety placed 7 residues away toward the N-terminus position. We hypothesized that the aromatic moiety would interact with the hydrophobic region placed at the hinge between the catalytic domain and the flexible loop (Figure 1D), disrupting the dynamics of the loop during the substrate recognition process and inhibiting GalNAc-T2.

RESULTS AND DISCUSSION

MD techniques were used to screen different aromatic moiety candidates to improve the binding of a model glycopeptide to the hydrophobic region of GalNAc-T2.

The model peptide sequence GAAPĈPVPAAAAT*AA was used to screen different aromatic substituents, where Ĉ and T* represent an *S*-alkylated cysteine (Cys) and a *O*-GalNAcylated Thr, respectively. A simple alanine sequence was used as a spacer between T*, whereas the PV motif before was maintained as in the natural MUC1 acceptor substrate sequence to promote hydrophobic interaction with the aromatic region in proximity to the flexible loop hinge.

Three glycopeptide models bearing a pyridine (**1**), a quinoline (**2**), and a thiophene (**3**) aromatic substituent at Ĉ (Figure 1E) were studied through 1.0 μs MD simulations. The AMBER 22 software package was used and implemented with the appropriate force fields (see Supporting Information, Section 1).[41] All complexes were generated using the crystal structure of the active form of GalNAc-T2 in complex with UDP and a MUC5AC-like glycopeptide as the initial structure (PDB ID: 5AJP).[35]

According to these calculations, compound **3** (Figure 1F,G), with a thiophene moiety, can interact more tightly with the enzyme than compounds **1** and **2**, as inferred from the reduced distances calculated from the simulations for this aromatic moiety and the side chains of F280, W282, F280, and F361 of GalNAc-T2 (see Supporting Information, Figure S1 for representative frames of the MD simulation with compounds **1**−**3**). To validate our hypothesis, glycopeptides **2** and **3** were synthesized and tested against GalNAc-T2.

The thiophene derivative was prepared starting from commercially available 2-thiophenemethanol **5**, which was activated with HBr to provide bromide **6** (Figure 2A). Compound **6** was then reacted under mild basic conditions with the Cys derivative **13**[42] to give protected thiophene-Cys building block **7**. Finally, cleavage of the *t*-Bu protecting group with TFA furnished **8** in a good overall yield (Figure 2A). The quinoline derivative was prepared following a similar strategy, starting from commercially available 2-quinolinylmethanol **9**, which was converted into chloride **10** using $SOCl_2$, and then reacted with **13** providing derivative **11**. Finally, *t*-Bu cleavage with TFA afforded quinoline-Cys building block **12** in good overall yields (Figure 2A). The Fmoc-protected amino acid *α*-*O*-GalNAc-Thr **14** was prepared as described in previously reported procedures.[43] Glycopeptides **2** and **3** were then prepared through Fmoc-based solid phase peptide synthesis, and the target inhibitors were tested against GalNAc-T2.

The inhibition activity of glycopeptides **2** and **3** was determined using a radiometric assay using a labeled UDP-[^{3}H or ^{14}C]-GalNAc against a model peptide substrate MUC1a with sequence AHGVTSAPDTR, which was demonstrated to be a general substrate accepted by a wide range of

https://doi.org/10.1021/jacsau.4c00633
JACS Au 2024, 4, 3649−3656

GalNAc-Ts, including the T2 isoform.[38] The glycosylation products were isolated from the radiolabeled donors through chromatographic techniques and analyzed by scintillation counting. The analysis demonstrated that both glycopeptides **2** and **3** were able to inhibit GalNAc-T2, showing IC_{50} values of 396 ± 172 and 84.5 ± 25.9 μM for **2** and **3**, respectively (Figure 2B), in good agreement with the MD prediction.

To prove that the inhibition activity was the result of a bivalent interaction between the inhibitor and both the lectin and catalytic domain of GalNAc-T2, and was not the result of nonspecific interactions, we synthesized and tested peptide **4**, which retains the GalNAc and animo acid sequence of **3** but lacks the thiophene moiety (Figure 1E). Moreover, we measured the activity of a set of thiophene fragments, including compound **5**, thiophene **15**, and unprotected thiophene-Cys **16**, at a normalized concentration of 500 μM (Figure 2C). The assay demonstrated that only glycopeptide **3** could induce a relevant inhibition of the GalNAc-T2 activity while the whole set of thiophene fragments and peptide **4** were unable to inhibit the enzyme. This demonstrated that the aromatic moiety is necessary to exert the inhibition activity, likely due to its binding to the catalytic domain of the enzyme. However, the thiophene fragments alone did not act as inhibitors, suggesting that the two moieties of the bivalent glycopeptides were required for an optimal inhibition.

Next, we investigated the possibility to achieving a stronger inhibition activity through the modification of the peptide sequence connecting the PĈP motif to the T* residue.

Using MD simulations, we screened a set of bivalent glycopeptides where the amino acid sequence was rationally modified to improve the $\pi-\pi$ interaction between the thiophene moiety of **3** and the GalNAc-T2 F361 residue. We modeled three different variants **17**, **18**, and **19** (Figure 2D) using amino acidic residues presenting nonpolar side chains next to the PĈP motif. This feature was chosen with the objective to improving the interactions between the nonpolar side chains and the hydrophobic region nearby the flexible loop of the enzyme.

The analysis of the distance between the thiophene moiety of glycopeptides **17**−**19** and the F361 residue of GalNAc-T2, derived from MD simulations, rendered compounds **18** and **19** as the weakest and the strongest ligands, respectively (Figure 2E). These findings aligned with the experimentally measured IC_{50} values. Compound **18** had an IC_{50} value of 55.7 ± 31.5 μM, while compound **19** had an IC_{50} value of 21.4 ± 4.1 μM (Figure 2F). These results demonstrate the potential of using MD simulations in the inhibitor design process.

To date, compound **19** represents one of the best inhibitors ever reported against GalNAc-T2, with a comparable potency to luteolin (IC_{50} 14.7 μM), which was discovered through the high-throughput screening of libraries of compounds[29] and a mimetic of the UDP-GalNAc donor substrate referred to as compound 1−68A (IC_{50} 15.0 μM).[27] However, these compounds were not tested against different GalNAc-Ts and are likely to have broad activity against various isoforms. This is expected, especially in the case of compounds 1−68A, which targets the UDP-GalNAc binding site that is a common glycoside donor shared by the whole GalNAc-T family.

To shed light on the binding mode of **19** with GalNAc-T2, we conducted a series of saturation transfer difference (STD) NMR experiments. This technique allows us to map the parts of the inhibitor in close contact with the enzyme, since the highest STD intensities correlate with the closest ligand−protein contacts in the bound state. As a result, STD NMR provides so-called ligand binding epitope mapping depicting the most important moieties responsible for binding. Additionally, to gather information on the location of binding, we used the STD NMR competition of **19** with UDP-GalNAc, which is known to bind in the catalytic domain of the protein (Figure 3A). Due to significant signal overlap, we could

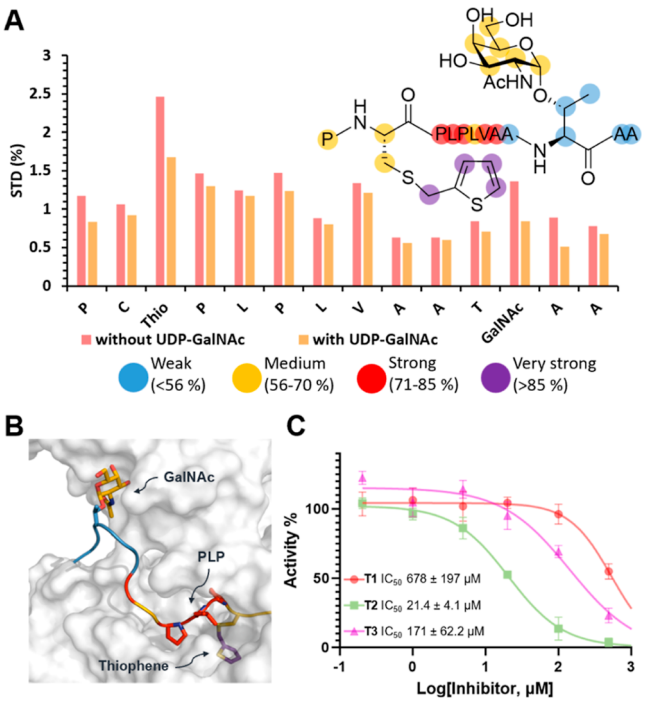


Figure 3. (A) STD NMR binding epitope map of **19** upon binding to GalNAc-T2, analyzed as an average per residue, due to significant proton signal overlapping. Colors represent different degrees of contact, from light blue to purple, as shown in the legend, and STD NMR competition experiments performed adding UDP-GalNAc to a sample containing **19** and GalNAc-T2. (B) Representative MD frame of **19** in complex with GalNAc-T2. Color code of **19** corresponds to the degree of contact with GalNAc-T2 as for Figure 3A. (C) Relative inhibitory activity % of compound **19** against GalNAc-T1, -T2, and -T3 using MUC1a as acceptor substrate.

provide a reliable analysis only based on averages per residue along the peptide, though the information gained is still relevant. Remarkably, the binding epitope map showed that the thiophene-bearing N-terminal end of the peptide received the largest amount of saturation from the enzyme. In fact, the thiophene residue is the moiety showing the closest contacts with the GalNAc-T2 surface in the bound state, suggesting that this region of thiophene glycopeptide **19** is fundamental to establishing favorable contacts with the enzyme. On the other hand, the GalNAc moiety also showed significant contacts with the surface of GalNAc-T2 in the bound state, despite the low STD intensities shown by the adjacent alanine residues. This result is compatible with a bidentate binding where the GalNAc sugar ring fits in the binding site of the lectin domain, whereas the thiophene and adjacent residues are located in the catalytic domain's binding site. The STD NMR data is correlated with the MD simulation, where the thiophene and the PLP motif are deeply inserted into the groove of the GalNAc-T2 catalytic site, establishing close contacts with the enzyme (Figure 3B). Similarly, the GalNAc moiety showed

https://doi.org/10.1021/jacsau.4c00633
JACS Au 2024, 4, 3649−3656

close contacts with the enzyme's lectin domain, while the C-terminal AA motif was exposed to the solvent in good agreement with the STD NMR results. The competition experiment with UDP-GalNAc showed a general decrease of the STD intensities with major changes occurring for the thiophene moiety, demonstrating that the inhibitors compete with the donor substrate too by interacting through the thiophene moiety with the catalytic site. These results corroborated that all of the key components of compound **19**, i.e., the GalNAc moiety, the nonpolar amino acid sequence, and the thiophene moiety, are involved in the binding to GalNAc-T2.

Finally, we determined the potential selectivity of inhibitor **19**, which was designed based on structural features of GalNAc-T2, known to prefer glycosylating N-terminal residues from a prior C-terminal GalNAcylated position.[26] This potential selectivity was assessed over two enzymatic models using GalNAc-T3,[26] which exhibits the opposite preference, and GalNAc-T1, which can glycosylate with both preferences.[41] Moreover, these three enzymes hold high biological relevance, representing the most abundant GalNAc-Ts, which are widely expressed in most tissues.[44] We measured and compared the inhibitory activity of **19** against GalNAc-T1, T2, and T3 enzymes and demonstrated the high degree of selectivity toward the T2 isoform with IC_{50} values of 678 ± 197 μM, 21.4 ± 4.1 μM, and 170.8 ± 62.2 μM against GalNAc-T1, -T2, and -T3, respectively (Figure 3C). This corresponds to an 8x more potent inhibitory activity against GalNAc-T2 compared to T3, and a 32x higher selectivity with respect to the T1 isoform. Moreover, it was possible to calculate the kinetic parameters of the enzymatic inhibition for the three GalNAc-Ts (see Supporting Information in Table S1 in Section 4.4). The comparison of the K_I values calculated for both the acceptor and donor substrates showed 31× and 7× higher potencies for GalNAc-T2 inhibition with respect to GalNAc-T1 and T3, respectively. These findings demonstrate a strong correlation between the results, confirming the selectivity of glycopeptide **19** for inhibiting GalNAc-T2.

In summary, we identified and synthesized compound **19** as the first selective GalNAc-T2 inhibitor reported so far. It is worth noting that the GalNAc-T3/T4/T6/T12 isoforms prefer the glycosylate long-range C-terminal position; therefore, a certain degree of selectivity of **19** toward the T2 over T3 isoform was expected. Remarkably, **19** was, with unprecedent selectivity, able to inhibit the T2 over the T1 isoform, which are both capable to operate toward long-range N-terminal positions.[26] This demonstrates the great potential of our strategy toward the synthesis of specific GalNAc-T inhibitors.

CONCLUSIONS

In this study, a structure-guided rational design approach was used to synthesize a GalNAc-T2 inhibitor with improved potency over other isoforms by targeting both the lectin and catalytic domains of the enzyme. Compound **19** demonstrated how the dual targeting of the lectin and catalytic domain of GalNAc-T2 leads to an effective inhibition of the enzyme with an IC50 of 21 μM. Compared to the T1 and T3 isoforms compound **19** showed a 32- and 8-times greater selectivity toward the targeted T2 isoform, demonstrating how the MD-assisted rational design of bivalent inhibitors improves the selective inhibition of a single isoform. Further studies in cellular models to assess the effectiveness of compound **19** in the inhibition of GalNAc-T2 are required before its implementation in biological settings. However, this versatile strategy could be effectively exploited for the synthesis of selective inhibitors of other GalNAc-Ts, setting the stage for future drug development oriented toward glycan-based therapeutics.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/jacsau.4c00633.

MD simulation, chemical synthesis, STD NMR studies, and biochemical studies (PDF)

AUTHOR INFORMATION

Corresponding Author

Mattia Ghirardello − *Department of Chemistry and Instituto de Investigación en Química de la Universidad de La Rioja, Universidad de La Rioja, Logroño 26006, Spain;* orcid.org/0000-0002-2855-4801; Email: mattia.ghirardello@unirioja.es

Authors

Ismael Compañón − *Department of Chemistry and Instituto de Investigación en Química de la Universidad de La Rioja, Universidad de La Rioja, Logroño 26006, Spain*

Collin J. Ballard − *Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106, United States;* orcid.org/0000-0003-2869-6970

Erandi Lira-Navarrete − *Department of Cellular and Molecular Medicine, Faculty of Health Sciences, Copenhagen Center for Glycomics, University of Copenhagen, Copenhagen 2200, Denmark*

Tanausú Santos − *Department of Chemistry and Instituto de Investigación en Química de la Universidad de La Rioja, Universidad de La Rioja, Logroño 26006, Spain*

Serena Monaco − *School of Pharmacy, University of East Anglia, NR4 7TJ Norwich, U.K.;* orcid.org/0000-0001-9396-7568

Juan C. Muñoz-García − *School of Pharmacy, University of East Anglia, NR4 7TJ Norwich, U.K.; Instituto de Investigaciones Químicas, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, Sevilla 41092, Spain;* orcid.org/0000-0003-2246-3236

Ignacio Delso − *School of Pharmacy, University of East Anglia, NR4 7TJ Norwich, U.K.*

Jesus Angulo − *School of Pharmacy, University of East Anglia, NR4 7TJ Norwich, U.K.; Instituto de Investigaciones Químicas, Consejo Superior de Investigaciones Científicas and Universidad de Sevilla, Sevilla 41092, Spain;* orcid.org/0000-0001-7250-5639

Thomas A. Gerken − *Department of Biochemistry, Case Western Reserve University, Cleveland, Ohio 44106, United States; Departments of Biochemistry and Chemistry, Case Western Reserve University, Cleveland, Ohio 44106, United States*

Katrine T. Schjoldager − *Department of Cellular and Molecular Medicine, Faculty of Health Sciences, Copenhagen Center for Glycomics, University of Copenhagen, Copenhagen 2200, Denmark*

Henrik Clausen − *Department of Cellular and Molecular Medicine, Faculty of Health Sciences, Copenhagen Center for*

https://doi.org/10.1021/jacsau.4c00633
JACS Au 2024, 4, 3649−3656

Glycomics, University of Copenhagen, Copenhagen 2200, Denmark

Tomás Tejero − *Department of Organic Chemistry, Faculty of Sciences, University of Zaragoza, Zaragoza 50009, Spain; Institute of Chemical Synthesis and Homogeneous Catalysis, University of Zaragoza-CSIC, Zaragoza 50009, Spain;* orcid.org/0000-0003-3433-6701

Pedro Merino − *Department of Organic Chemistry, Faculty of Sciences, University of Zaragoza, Zaragoza 50009, Spain; Institute for Biocomputation and Physics of Complex Systems, University of Zaragoza, Zaragoza 50018, Spain;* orcid.org/0000-0002-2202-3460

Francisco Corzana − *Department of Chemistry and Instituto de Investigación en Química de la Universidad de La Rioja, Universidad de La Rioja, Logroño 26006, Spain;* orcid.org/0000-0001-5597-8127

Ramon Hurtado-Guerrero − *Department of Cellular and Molecular Medicine, Faculty of Health Sciences, Copenhagen Center for Glycomics, University of Copenhagen, Copenhagen 2200, Denmark; Institute for Biocomputation and Physics of Complex Systems, University of Zaragoza, Zaragoza 50018, Spain; Fundación ARAID, Zaragoza 50018, Spain;* orcid.org/0000-0002-3122-9401

Complete contact information is available at: https://pubs.acs.org/10.1021/jacsau.4c00633

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank the Agencia Estatal de Investigación (AEI, PID2021-127622OB-I00 and PDC2022-133725-C21 to F.C., PID2022-136362NB-I00 to R.H.-G., PID2022137973NB-I00 to P.M.), Universidad de La Rioja (REGI22/16 Project), and Gobierno de Aragón (E34_R17 and LMP58_18) with FEDER (2014−2020) funds for 'Building Europe from Aragón' for financial support (to R.H.-G.). The support of the National Institutes of Health grant R01-GM113535 (to T.A.G.) is also acknowledged. M.G. and I.D. acknowledge the Marie Skłodowska-Curie fellowship program (grant agreement nos. 101034288 (M.G.) and 890779 (I.D.)) for financial support. J.A. and S.M. acknowledge support of BBSRC, grant BB/P010660/1. J.A. and J.C.M.-G. also acknowledge funding from the grant AEI/10.13039/501100011033/PID2022-142879NB-I00, cofunded by the European Regional Development Fund (ERDF), "A way of making Europe", and T.S. thanks MICIU/AEI/10.13039/501100011033 and the European Union NextGenerationEU/PRTR for his Juan de la Cierva contract, JDC2022-048607-I.

■ REFERENCES

(1) Schjoldager, K. T.; Narimatsu, Y.; Joshi, H. J.; Clausen, H. Global View of Human Protein Glycosylation Pathways and Functions. *Nat. Rev. Mol. Cell Biol.* **2020**, *21* (12), 729−749.

(2) Varki, A. Biological Roles of Glycans. *Glycobiology* **2017**, *27* (1), 3−49.

(3) Reily, C.; Stewart, T. J.; Renfrow, M. B.; Novak, J. Glycosylation in Health and Disease. *Nat. Rev. Nephrol.* **2019**, *15* (6), 346−366.

(4) Urata, Y.; Takeuchi, H. Effects of Notch Glycosylation on Health and Diseases. *Dev., Growth Differ.* **2020**, *62* (1), 35−48.

(5) Magalhães, A.; Duarte, H. O.; Reis, C. A. Aberrant Glycosylation in Cancer: A Novel Molecular Mechanism Controlling Metastasis. *Cancer Cell* **2017**, *31* (6), 733−735.

(6) Smith, B. A. H.; Bertozzi, C. R. The Clinical Impact of Glycobiology: Targeting Selectins, Siglecs and Mammalian Glycans. *Nat. Rev. Drug Discovery* **2021**, *20* (3), 217−243.

(7) Ghirardello, M.; Shyam, R.; Galan, M. C. Reengineering of Cancer Cell Surface Charges Can Modulate Cell Migration. *Chem. Commun.* **2022**, *58* (36), 5522−5525.

(8) *Essentials of Glycobiology*, 4th ed.; Varki, A., Cummings, R. D., Esko, J. D., Stanley, P., Hart, G. W., Aebi, M., Mohnen, D., Kinoshita, T., Packer, N. H., Prestegard, J. H., Schnaar, R. L., Seeberger, P. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor (NY), 2022.

(9) Malaker, S. A.; Riley, N. M.; Shon, D. J.; Pedram, K.; Krishnan, V.; Dorigo, O.; Bertozzi, C. R. Revealing the Human Mucinome. *Nat. Commun.* **2022**, *13* (1), 3542.

(10) de las Rivas, M.; Lira-Navarrete, E.; Gerken, T. A.; Hurtado-Guerrero, R. Polypeptide GalNAc-Ts: From Redundancy to Specificity. *Curr. Opin. Struct. Biol.* **2019**, *56*, 87−96.

(11) de las Rivas, M.; Paul Daniel, E. J.; Narimatsu, Y.; Compañón, I.; Kato, K.; Hermosilla, P.; Thureau, A.; Ceballos-Laita, L.; Coelho, H.; Bernadó, P.; Marcelo, F.; Hansen, L.; Maeda, R.; Lostao, A.; Corzana, F.; Clausen, H.; Gerken, T. A.; Hurtado-Guerrero, R. Molecular Basis for Fibroblast Growth Factor 23 O-Glycosylation by GalNAc-T3. *Nat. Chem. Biol.* **2020**, *16* (3), 351−360.

(12) Kato, K.; Jeanneau, C.; Tarp, M. A.; Benet-Pagès, A.; Lorenz-Depiereux, B.; Bennett, E. P.; Mandel, U.; Strom, T. M.; Clausen, H. Polypeptide GalNAc-transferase T3 and Familial Tumoral Calcinosis. *J. Biol. Chem.* **2006**, *281* (27), 18370−18377.

(13) Wang, S.; Mao, Y.; Narimatsu, Y.; Ye, Z.; Tian, W.; Goth, C. K.; Lira-Navarrete, E.; Pedersen, N. B.; Benito-Vicente, A.; Martin, C.; Uribe, K. B.; Hurtado-Guerrero, R.; Christoffersen, C.; Seidah, N. G.; Nielsen, R.; Christensen, E. I.; Hansen, L.; Bennett, E. P.; Vakhrushev, S. Y.; Schjoldager, K. T.; Clausen, H. Site-Specific O-Glycosylation of Members of the Low-Density Lipoprotein Receptor Superfamily Enhances Ligand Interactions. *J. Biol. Chem.* **2018**, *293* (19), 7408−7422.

(14) Pedersen, N. B.; Wang, S.; Narimatsu, Y.; Yang, Z.; Halim, A.; Schjoldager, K. T.-B. G.; Madsen, T. D.; Seidah, N. G.; Bennett, E. P.; Levery, S. B.; et al. Low Density Lipoprotein Receptor Class A Repeats Are O-Glycosylated in Linker Regions. *J. Biol. Chem.* **2014**, *289* (25), 17312−17324.

(15) Tian, E.; Wang, S.; Zhang, L.; Zhang, Y.; Malicdan, M. C.; Mao, Y.; Christoffersen, C.; Tabak, L. A.; Schjoldager, K. T.; Ten Hagen, K. G. Galnt11 Regulates Kidney Function by Glycosylating the Endocytosis Receptor Megalin to Modulate Ligand Binding. *Proc. Natl. Acad. Sci. U.S.A.* **2019**, *116* (50), 25196−25202.

(16) Li, Z.; Yamada, S.; Inenaga, S.; Imamura, T.; Wu, Y.; Wang, K.-Y.; Shimajiri, S.; Nakano, R.; Izumi, H.; Kohno, K.; Sasaguri, Y. Polypeptide N-Acetylgalactosaminyltransferase 6 Expression in Pancreatic Cancer Is an Independent Prognostic Factor Indicating Better Overall Survival. *Br. J. Cancer* **2011**, *104* (12), 1882−1889.

(17) Peng, X.; Chen, X.; Zhu, X.; Chen, L. GALNT6 Knockdown Inhibits the Proliferation and Migration of Colorectal Cancer Cells and Increases the Sensitivity of Cancer Cells to 5-FU. *J. Cancer* **2021**, *12* (24), 7413−7421.

(18) Scott, E.; Hodgson, K.; Calle, B.; Turner, H.; Cheung, K.; Bermudez, A.; Marques, F. J. G.; Pye, H.; Yo, E. C.; Islam, K.; Oo, H. Z.; McClurg, U. L.; Wilson, L.; Thomas, H.; Frame, F. M.; Orozco-Moreno, M.; Bastian, K.; Arredondo, H. M.; Roustan, C.; Gray, M. A.; Kelly, L.; Tolson, A.; Mellor, E.; Hysenaj, G.; Goode, E. A.; Garnham, R.; Duxfield, A.; Heavey, S.; Stopka-Farooqui, U.; Haider, A.; Freeman, A.; Singh, S.; Johnston, E. W.; Punwani, S.; Knight, B.; McCullagh, P.; McGrath, J.; Crundwell, M.; Harries, L.; Bogdan, D.; Westaby, D.; Fowler, G.; Flohr, P.; Yuan, W.; Sharp, A.; de Bono, J.; Maitland, N. J.; Wisnovsky, S.; Bertozzi, C. R.; Heer, R.; Guerrero, R. H.; Daugaard, M.; Leivo, J.; Whitaker, H.; Pitteri, S.; Wang, N.; Elliott, D. J.; Schumann, B.; Munkley, J. Upregulation of GALNT7 in Prostate Cancer Modifies O-Glycosylation and Promotes Tumour Growth. *Oncogene* **2023**, *42* (12), 926−937.

(19) Khetarpal, S. A.; Schjoldager, K. T.; Christoffersen, C.; Raghavan, A.; Edmondson, A. C.; Reutter, H. M.; Ahmed, B.; Ouazzani, R.; Peloso, G. M.; Vitali, C.; Zhao, W.; Somasundara, A. V. H.; Millar, J. S.; Park, Y.; Fernando, G.; Livanov, V.; Choi, S.; Noé, E.; Patel, P.; Ho, S. P.; Kirchgessner, T. G.; Wandall, H. H.; Hansen, L.; Bennett, E. P.; Vakhrushev, S. Y.; Saleheen, D.; Kathiresan, S.; Brown, C. D.; Abou Jamra, R.; LeGuern, E.; Clausen, H.; Rader, D. J. Loss of Function of GALNT2 Lowers High-Density Lipoproteins in Humans, Nonhuman Primates, and Rodents. *Cell Metab.* **2016**, *24* (2), 234−245.
(20) Zilmer, M.; Edmondson, A. C.; Khetarpal, S. A.; Alesi, V.; Zaki, M. S.; Rostasy, K.; Madsen, C. G.; Lepri, F. R.; Sinibaldi, L.; Cusmai, R.; et al. Novel Congenital Disorder of O-Linked Glycosylation Caused by GALNT2 Loss of Function. *Brain* **2020**, *143* (4), 1114−1126.
(21) Verzijl, C. R. C.; Oldoni, F.; Loaiza, N.; Wolters, J. C.; Rimbert, A.; Tian, E.; Yang, W.; Struik, D.; Smit, M.; Kloosterhuis, N. J.; Fernandez, A. J.; Samara, N. L.; Ten Hagen, K. G.; Dalal, K.; Chernish, A.; McCluggage, P.; Tabak, L. A.; Jonker, J. W.; Kuivenhoven, J. A. A Novel Role for GalNAc-T2 Dependent Glycosylation in Energy Homeostasis. *Mol. Metab.* **2022**, *60*, 101472.
(22) Sun, X.-L. Glycosyltransferases as Potential Drug Targets. *Med. Chem.* **2013**, *3* (1), 1−2.
(23) Costa, A. F.; Campos, D.; Reis, C. A.; Gomes, C. Targeting Glycosylation: A New Road for Cancer Drug Discovery. *Trends Cancer* **2020**, *6* (9), 757−766.
(24) Hu, Y.; Feng, J.; Wu, F. The Multiplicity of Polypeptide GalNAc-Transferase: Assays, Inhibitors, and Structures. *ChemBioChem* **2018**, *19* (24), 2503−2521.
(25) Cohen, P.; Cross, D.; Jänne, P. A. Kinase Drug Discovery 20 Years after Imatinib: Progress and Future Directions. *Nat. Rev. Drug Discovery* **2021**, *20* (7), 551−569.
(26) Coelho, H.; Rivas, M. d.; Grosso, A. S.; Diniz, A.; Soares, C. O.; Francisco, R. A.; Dias, J. S.; Compañon, I.; Sun, L.; Narimatsu, Y.; Vakhrushev, S. Y.; Clausen, H.; Cabrita, E. J.; Jiménez-Barbero, J.; Corzana, F.; Hurtado-Guerrero, R.; Marcelo, F.; Corzana, F.; Hurtado-Guerrero, R.; Marcelo, F. Atomic and Specificity Details of Mucin 1 O-Glycosylation Process by Multiple Polypeptide GalNAc-Transferase Isoforms Unveiled by NMR and Molecular Modeling. *JACS Au* **2022**, *2* (3), 631−645.
(27) Hang, H. C.; Yu, C.; Ten Hagen, K. G.; Tian, E.; Winans, K. A.; Tabak, L. A.; Bertozzi, C. R. Small Molecule Inhibitors of Mucin-Type O-Linked Glycosylation from a Uridine-Based Library. *Chem. Biol.* **2004**, *11* (3), 337−345.
(28) Song, L.; Linstedt, A. D. Inhibitor of ppGalNAc-T3-Mediated O-Glycosylation Blocks Cancer Cell Invasiveness and Lowers FGF23 Levels. *eLife* **2017**, *6*, 1−15.
(29) Liu, F.; Xu, K.; Xu, Z.; de Las Rivas, M.; Wang, C.; Li, X.; Lu, J.; Zhou, Y.; Delso, I.; Merino, P.; et al. The Small Molecule Luteolin Inhibits N-Acetyl-α-Galactosaminyltransferases and Reduces Mucin-Type O-Glycosylation of Amyloid Precursor Protein. *J. Biol. Chem.* **2017**, *292* (52), 21304−21319.
(30) Busca, P.; Piller, V.; Piller, F.; Martin, O. R. Synthesis and Biological Evaluation of New UDP-GalNAc Analogues for the Study of Polypeptide-α-GalNAc-Transferases. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 1853−1856.
(31) Ghirardello, M.; de Las Rivas, M.; Lacetera, A.; Delso, I.; Lira-Navarrete, E.; Tejero, T.; Martín-Santamaría, S.; Hurtado-Guerrero, R.; Merino, P. Glycomimetics Targeting Glycosyltransferases: Synthetic, Computational and Structural Studies of Less-Polar Conjugates. *Chem.—Eur. J.* **2016**, *22* (21), 7215−7224.
(32) Choi, J.; Wagner, L. J. S.; Timmermans, S. B. P. E.; Malaker, S. A.; Schumann, B.; Gray, M. A.; Debets, M. F.; Takashima, M.; Gehring, J.; Bertozzi, C. R. Engineering Orthogonal Polypeptide GalNAc-Transferase and UDP-Sugar Pairs. *J. Am. Chem. Soc.* **2019**, *141* (34), 13442−13453.
(33) Calle, B.; Gonzalez-Rodriguez, E.; Mahoney, K. E.; Cioce, A.; Bineva-Todd, G.; Tastan, O. Y.; Roustan, C.; Flynn, H.; Malaker, S. A.; Schumann, B. Bump-and-Hole Engineering of Human Polypeptide N-Acetylgalactosamine Transferases to Dissect Their Protein Substrates and Glycosylation Sites in Cells. *STAR Protoc.* **2023**, *4* (1), 101974.
(34) Lira-Navarrete, E.; Iglesias-Fernández, J.; Zandberg, W. F.; Compañón, I.; Kong, Y.; Corzana, F.; Pinto, B. M.; Clausen, H.; Peregrina, J. M.; Vocadlo, D. J.; Rovira, C.; Hurtado-Guerrero, R. Substrate-Guided Front-Face Reaction Revealed by Combined Structural Snapshots and Metadynamics for the Polypeptide N-Acetylgalactosaminyltransferase 2. *Angew. Chem., Int. Ed.* **2014**, *53* (31), 8206−8210.
(35) Lira-Navarrete, E.; Rivas, M. D. L.; Compañón, I.; Pallarés, M. C.; Kong, Y.; Iglesias-Fernández, J.; Bernardes, G. J. L.; Peregrina, J. M.; Rovira, C.; Bernadó, P.; Bruscolini, P.; Clausen, H.; Lostao, A.; Corzana, F.; Hurtado-Guerrero, R. Dynamic Interplay between Catalytic and Lectin Domains of GalNAc-Transferases Modulates Protein O-Glycosylation. *Nat. Commun.* **2015**, *6* (1), 6937.
(36) Rivas, M. d. L.; Lira-Navarrete, E.; Daniel, E. J. P.; Compañón, I.; Coelho, H.; Diniz, A.; Jiménez-Barbero, J.; Peregrina, J. M.; Clausen, H.; Corzana, F.; et al. The Interdomain Flexible Linker of the Polypeptide GalNAc Transferases Dictates Their Long-Range Glycosylation Preferences. *Nat. Commun.* **2017**, *8* (1), 1959.
(37) Mahajan, S. P.; Srinivasan, Y.; Labonte, J. W.; DeLisa, M. P.; Gray, J. J. Structural Basis for Peptide Substrate Specificities of Glycosyltransferase GalNAc-T2. *ACS Catal.* **2021**, *11* (5), 2977−2991.
(38) Pedersen, J. W.; Bennett, E. P.; Schjoldager, K. T.-B. G.; Meldal, M.; Holmér, A. P.; Blixt, O.; Cló, E.; Levery, S. B.; Clausen, H.; Wandall, H. H. Lectin Domains of Polypeptide GalNAc Transferases Exhibit Glycopeptide Binding Specificity. *J. Biol. Chem.* **2011**, *286* (37), 32684−32696.
(39) Ballard, C. J.; Paserba, M. R.; Daniel, E. J. P.; Hurtado-Guerrero, R.; Gerken, T. A. Polypeptide N-acetylgalactosaminyltransferase (GalNAc-T) isozyme surface charge governs charge substrate preferences to modulate mucin type O-glycosylation. *Glycobiology* **2023**, *33* (10), 817.
(40) Collette, A. M.; Hassan, S. A.; Schmidt, S. I.; Lara, A. J.; Yang, W.; Samara, N. L. An Unusual Dual Sugar-Binding Lectin Domain Controls the Substrate Specificity of a Mucin-Type O-Glycosyltransferase. *Sci. Adv.* **2024**, *10* (9), No. eadj8829.
(41) Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J. T.; Brozell, S. R.; Cerutti, D. S.; Cheatham, T. E., III; Cisneros, G. A.; Cruzeiro, V. W. D.; Darden, T. A.; Duke, R. E.; Giambasu, G.; Gilson, M. K.; Gohlke, H.; Goetz, A. W.; Harris, R.; Izadi, S.; Izmailov, S. A.; Kasavajhala, K.; Kaymak, M. C.; King, E.; Ko-valenko, A.; Kurtzman, T.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Liu, J.; Luchko, T.; Luo, R.; Machado, M.; Man, V.; Manathunga, M.; Merz, K. M.; Miao, Y.; Mikhailovskii, O.; Monard, G.; Nguyen, H.; O'Hearn, K. A.; Onufriev, A.; Pan, F.; Pantano, S.; Qi, R.; Rahnamoun, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Schott-Verdugo, S.; Shajan, A.; Shen, J.; Simmerling, C. L.; Skrynnikov, N. R.; Smith, J.; Swails, J.; Walker, R. C.; Wang, J.; Wang, J.; Wei, H.; Wolf, R. M.; Wu, X.; Xiong, Y.; Xue, Y.; York, D. M.; Zhao, S.; Kollman, P. A. *Amber 2022*; University of California: San Francisco, 2022; .
(42) Ghirardello, M.; Delso, I.; Tejero, T.; Merino, P. Synthesis of Amino-Acid−Nucleoside Conjugates. *Asian J. Org. Chem.* **2016**, *5* (12), 1525−1534.
(43) Compañón, I.; Guerreiro, A.; Mangini, V.; Castro-López, J.; Escudero-Casao, M.; Avenoza, A.; Busto, J. H.; Castillón, S.; Jiménez-Barbero, J.; Asensio, J. L.; Jiménez-Osés, G.; Boutureira, O.; Peregrina, J. M.; Hurtado-Guerrero, R.; Fiammengo, R.; Bernardes, G. J. L.; Corzana, F. Structure-Based Design of Potent Tumor-Associated Antigens: Modulation of Peptide Presentation by Single-Atom O/S or O/Se Substitutions at the Glycosidic Linkage. *J. Am. Chem. Soc.* **2019**, *141* (9), 4063−4072.
(44) Yang, W.; Tian, E.; Chernish, A.; McCluggage, P.; Dalal, K.; Lara, A.; Ten Hagen, K. G.; Tabak, L. A. Quantitative Mapping of the in Vivo O-GalNAc Glycoproteome in Mouse Tissues Identifies

https://doi.org/10.1021/jacsau.4c00633
JACS Au 2024, 4, 3649−3656

GalNAc-T2 O-Glycosites in Metabolic Disorder. *Proc. Natl. Acad. Sci. U.S.A.* **2023**, *120* (43), No. e2303703120.

https://doi.org/10.1021/jacsau.4c00633
JACS Au 2024, 4, 3649−3656