Generating orthogonal glycosyltransferase and nucleotide sugar pairs as next-generation glycobiology tools

Anna Cioce[1,2], Stacy A. Malaker[3,4] and Benjamin Schumann[1,2]

Abstract

Protein glycosylation fundamentally impacts biological processes. Nontemplated biosynthesis introduces unparalleled complexity into glycans that needs tools to understand their roles in physiology. The era of quantitative biology is a great opportunity to unravel these roles, especially by mass spectrometry glycoproteomics. However, with high sensitivity come stringent requirements on tool specificity. Bioorthogonal metabolic labeling reagents have been fundamental to studying the cell surface glycoproteome but typically enter a range of different glycans and are thus of limited specificity. Here, we discuss the generation of metabolic 'precision tools' to study particular subtypes of the glycome. A chemical biology tactic termed bump-and-hole engineering generates mutant glycosyltransferases that specifically accommodate bioorthogonal monosaccharides as an enabling technique of glycobiology. We review the groundbreaking discoveries that have led to applying the tactic in the living cell and the implications in the context of current developments in mass spectrometry glycoproteomics.

Addresses

[1] Chemical Glycobiology Laboratory, The Francis Crick Institute, 1 Midland Road, NW1 1AT, London, United Kingdom
[2] Department of Chemistry, Imperial College London, 80 Wood Lane, W12 0BZ, London, United Kingdom
[3] Department of Chemistry, Stanford University, 290 Jane Stanford Way, Stanford, CA, 94305, USA
[4] Department of Chemistry, Yale University, 275 Prospect Street, New Haven, CT, 06511, USA

Corresponding authors: Schumann, Benjamin (b.schumann@imperial.ac.uk); Malaker, Stacy A (stacy.malaker@yale.edu)

Introduction

Protein glycosylation is the most complex post-translational modification. Dysfunctions in the biosynthesis and turnover of glycan structures (the 'glycome'), as well as underlying glycoproteins (the 'glycoproteome') are associated with disease [1]. Attachment to proteins is primarily via Asn (N-linked) or Ser/Thr/Tyr (O-linked) side chains. Although N-linked glycosylation is found on consensus peptide sequons (N_X_S/T) and can thus be predicted, O-linked glycosylation often lacks such sequons (Figure 1a) [2,3]. Cell surface glycans are biosynthesized from 10 monosaccharide units by the combinatorial activity of more than 250 glycosyltransferases (GTs). Although the secretory pathway comprises an arsenal of GTs that influence each other through compensation and competition [4], deficiencies of individual GTs are related to congenital disorders of glycosylation [5]. Despite forays made into understanding the molecular details of GT activity, our insights are still limited by the complexity of the secretory pathway and the analytical challenges associated with studying glycans.

Glycoproteome analysis helps elucidate many structural and functional properties of glycoproteins as a readout of GT activity. These analyses have been fueled by tools in biology and chemistry. Standing out among the latter are the metabolic labeling (also called metabolic oligosaccharide engineering [MOE]) reagents that contain a bioorthogonal, reactive handle [6]. When fed to growing cells, the MOE reagent is incorporated into newly synthesized glycoproteins via biosynthetic machineries. Functionalized glycoproteins are then reacted with an enrichment probe allowing isolation and MS analysis [7], or a fluorescent probe for or imaging in living cells [8,9]. MOE reagent and probe must possess complementary functional groups that are chemically inert in a biological environment while reactive toward each other.

Figure 1
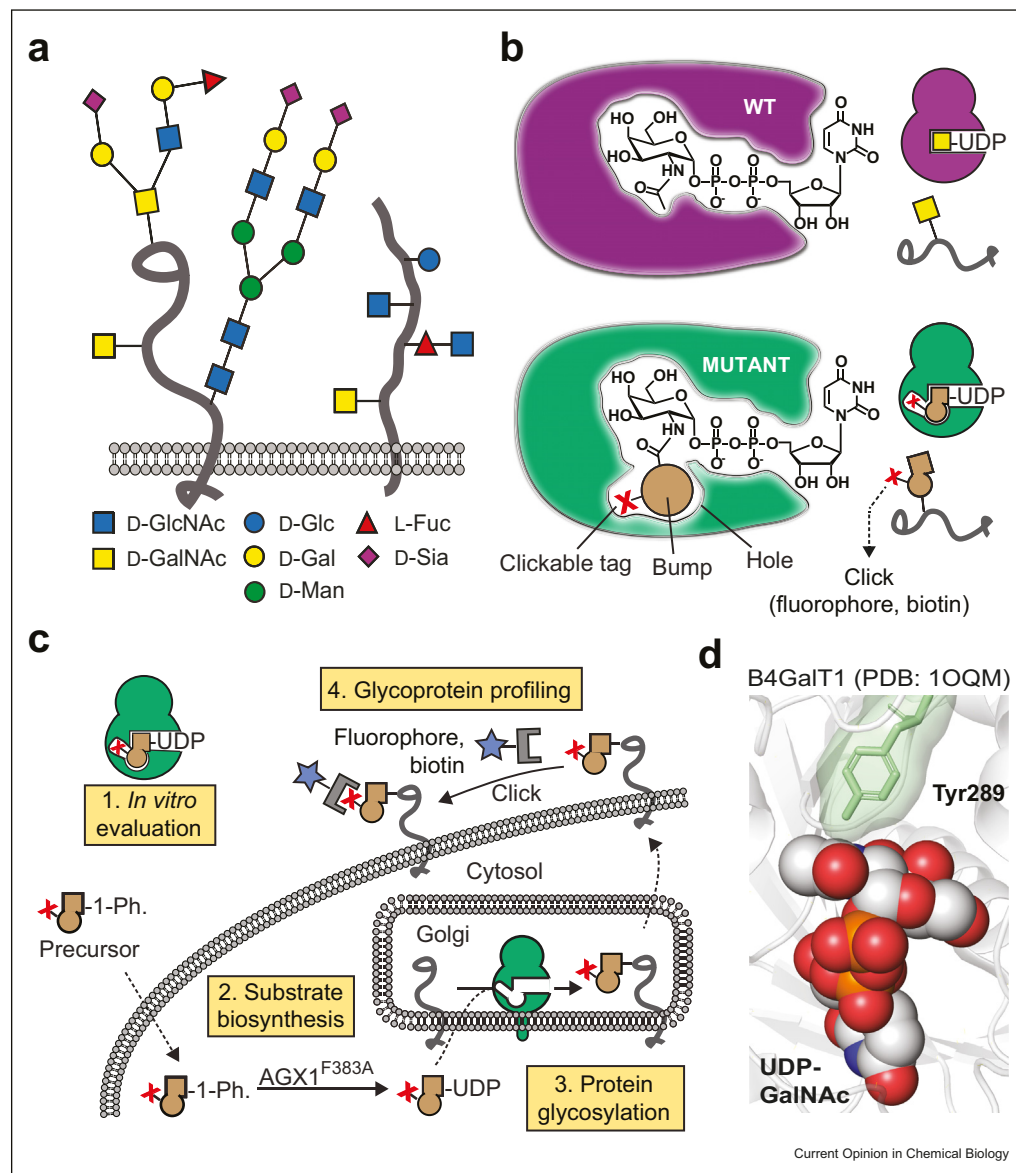
Using glycosyltransferase bump-and-hole engineering to understand cell surface glycosylation. (a) Diversity of cell surface glycans. **(b)** Bump-and-hole engineering enlarges the active site of a GT to accommodate a chemically modified nucleotide-sugar containing a bioorthogonal tag. **(c)** A blueprint of key steps to establish a cellular GT bump-and-hole system [11]. **(d)** Co-crystal structure of WT-B4GALT1 (PDB 1OQM) with UDP-GalNAc. B4GALT1 was subsequently engineered to accommodate bioorthogonal UDP-GalNAc analogs [25]. Adapted from *Molecular Cell,* Vol 78/5, B. Schumann et al., Bump-and-Hole Engineering Identifies Specific Substrates of Glycosyltransferases in Living Cells, 824–834.e15, Copyright (2020), with permission from Elsevier. GTs, glycosyltransferases.

Most bioorthogonal reactions involve copper-assisted (CuAAC) or strain-promoted (SPAAC) azide–alkyne cycloaddition, inverse electron demand Diels–Alder reactions between strained or terminal alkenes and tetrazine reporters [10], and the Staudinger ligation between azides and phosphines [6,8].

MOE reagents have produced valuable insights into various aspects of glycobiology [11–13]. Nevertheless, these reagents are of limited specificity by default for two reasons: (i) interconversion into other mono-saccharides with different biosynthetic fates and (ii) the substrate promiscuity of certain GTs. Thus, the bio-orthogonal label can be incorporated into undesired

substructures within the glycome. With the advent of sensitive methods of quantitative biology, more specific tools are needed to inform on the products of individual GTs [14].

Studying the products of individual members of a transferase family is the prime discipline of a chemical biology tactic called 'bump-and-hole (BH) engineering'. In a structure-guided process, an enzyme's catalytic pocket is enlarged by mutating bulky 'gatekeeper' amino acids into smaller ones, to create a BH mutant. This strategy creates an often hydrophobic 'hole' in the active site that is complementary to a synthetic substrate containing a bulky, 'bumped' functional group (Figure 1b). As the bumped substrate is only recognized by the BH mutant, this new enzyme—substrate pair is orthogonal to all other transferases in a complex system such as living cells. After inception in the field of kinases by Alaimo et al. Shokat and colleagues, [15], successful application to several enzyme families [16−24] has laid the foundation to the tactic being applied to GTs as an enabling method in the glycosciences (Figure 1c).

The early days of GT BH engineering
In the early 2000s, Qasba and Ramakrishnan [25] engineered a GT to accommodate nonnatural, bio-orthogonal uridine diphosphate(UDP)-sugars for the first time. Bovine β-1,4-galactosyltransferase (B4GalT1) normally transfers galactose to *N*-acetylglucosamine (GlcNAc)-terminating structures. Mutation of the residue Tyr298 to Leu rendered BH-B4GalT1 reactive toward UDP-*N*-acetylgalactosamine (UDP-GalNAc) derivatives such as UDP-2-keto-galactose and UDP-azido-*N*-acetylgalactosamine **1** (UDP-GalNAz) (Figure 1d) [26−28]. These analogs are not bulky enough to be considered 'bumped', as they are used by many other WT-GTs [14,29]. Moreover, the cytosolic UDP-GalNAc/GlcNAc 4′-epimerase GALE interconverts UDP-GalNAz into the corresponding UDP-GlcNAc derivative, UDP-GlcNAz [14]. BH-B4GalT1/UDP-GalNAz would thus not be amenable for use in the living cell. Nevertheless, through the efforts of Hsieh-Wilson and others, the BH-B4GalT1/UDP-GalNAz system has been developed into a tool to profile proteins with the nucleocytoplasmic Ser/Thr-linked O-GlcNAc modification *in vitro*. [30−35].

Reprogramming metabolism to deliver UDP-GalNAc analogs into living cells
BH approachA cellular GT bump-and-hole approach requires the biosynthesis of bumped UDP-sugars in the cytosol. Piller and colleagueset al. [36] found that the enzymes of the GalNAc salvage pathway — the kinase GALK2 and the pyrophosphorylase AGX1 — exhibit low promiscuity toward chemical modifications of the GalNAc acetamide (Figure 2a). Bumped GalNAc analogs are thus not effective substrates of either enzyme

[11,12,37]. Because AGX1 and its close homolog AGX2 are also components of the GlcNAc salvage pathway, low substrate promiscuity impeded delivery of both bumped UDP-GalNAc and UDP-GlcNAc derivatives.

Again, protein engineering came to the rescue. Kohler et al. [37] mutated the gatekeeper residues Phe381 or Phe383 in human AGX1 to Gly and assessed turnover of the bumped GlcNDAz-1-phosphate (Figure 2b) [43]. The mutant AGX1^{F383G} was subsequently used to endow cells with the capacity to biosynthesize UDP-GlcNDAz **2** (Figure 2c). To circumvent the nonpermissive GALK2 step, a caged, membrane-permeable GlcNDAz-1-phosphate precursor was used for delivery. GlcNDAz was then used as a photo-crosslinkable reporter of O-GlcNAc, allowing for mapping of protein interaction partners by MS proteomics, eventually in conjunction with BH engineering of O-GlcNAc transferase OGT [38].

BH engineering of a human GT family: polypeptide GalNAc transferases (GalNAc-Ts)
Attempts to BHbump-and-hole engineer an entire GT family started in the early 2000s (personal communication by C. R. Bertozzi), but took almost two decades to be brought to fruition. The polypeptide GalNAc transferases (GalNAc-Ts) constitute one of the largest GT families in the human genome. Approximately 20 isoenzymes (called T1-T20) catalyze the first step in the biosynthesis of glycans primed by GalNAcα1-*O*-Ser/Thr, also called O-GalNAc or mucin-type protein glycosylation. Although GalNAc-Ts have been connected to a wide variety of diseases [2,39−42], it is still challenging to associate individual isoenzymes with a specific biological function due to their complex interplay in the secretory pathway. Uniquely suited to dissect GalNAc-T biology, the BH approach was not amenable to the GalNAc-T family before a series of key requirements were met by the groundbreaking work by Qasba, Hsieh-Wilson, Bertozzi, Kohler and many others.

The BH approach needs an enzyme/substrate co-crystal structure, the first of which (GalNAc-T10 and GalNAc) was published by Narimatsu and colleagueset al. [43]. The gatekeeper residues Ile and Leu could be identified and mapped for other isoenzymes based on structural and sequence homology (Figure 3a and b). Second, bioorthogonal, bumped UDP-GalNAc analogs were required to identify suitable BH-enzyme-substrate pairs *in vitro*. Choi, Wagner et al. [44] used a combination of chemical and chemoenzymatic syntheses to develop a collection of 20 UDP-GalNAc analogs. Mutation of both gatekeeper residues to Ala reprogrammed the nucleotide sugar specificity of WT-GalNAc-T1, T2, and T10 to chemically modified UDP-GalNAc analogs such as alkynes **3−5** *in vitro* (Figure 3a, c). We subsequently found that engineering preserved both the three-dimensional

Figure 2
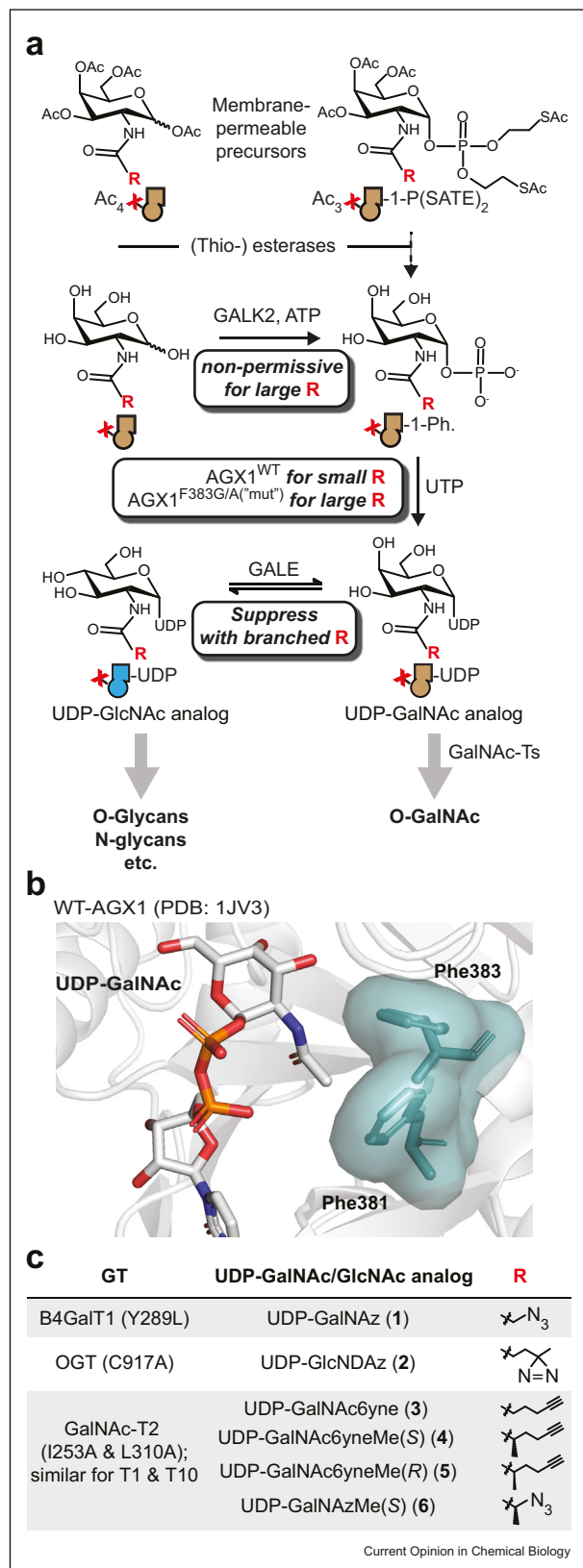
Reprogramming metabolism to deliver UDP-sugar analogs. (a) Schematic representation of the GalNAc salvage pathway applied to

structure of T2 and peptide substrate preference of T1 and T2 [11], ascertaining WT-like behavior of BH-GalNAc-Ts (Figure 3d).

A GalNAc-T BH system was then assembled in living cells. Our approach featured AGX1-mediated substrate delivery based on Kohler's strategy. Using the mutant AGX1[F383A], we delivered UDP-GalNAc analog **3** into living cells from the membrane-permeable, caged GalNAc-1-phosphate analog Ac₃GalNAc6yne-1-P(SATE)₂ that is deprotected in the cytosol, presumably by esterases and thioesterases (Figures 1c and 2a). Of note, the AGX1 F383A mutant had largely similar properties to the F383G mutant but accepted a broader GalNAc-1-phosphate substrate range in our hands and is called here 'mut-AGX1' (Figure 2a) [12]. Monitoring UDP-sugar biosynthesis by ion exchange chromatography of cell extracts was essential at this point to confirm substrate delivery. Immunofluorescence further revealed that BH-T1 and BH-T2 localize to the Golgi compartment, and expression can be controlled with a doxycycline (Dox)-inducible promoter (Figure 3e).

With delivery of bumped substrates and BH-GalNAc-T expression as crucial prerequisites, the stage was set to test incorporation of GalNAc analogs into cell surface glycans. Using the alkyne group as a bioorthogonal tag is not without caveats, as it relies on using CuAAC instead of SPAAC for reaction with reporter moieties such as fluorophores [6]. Wu and coworkers had developed biocompatible Cu(I) ligands to allow for CuAAC on the surface of living cells [45,46]. Picolyl azide derivatives developed by Ting and colleagues [47] that accelerate CuAAC by an order of magnitude served to characterize cell surface glycans introduced by individual GalNAc-Ts by both flow cytometry and in-gel fluorescence (Figure 3e). We found a several-fold increase of signal over background when a functional BH system was present, and a largely overlapping band pattern of labeled glycoproteins between GalNAc-T1 and T2. MS proteomics revealed that these were glycoproteins with highly O-glycosylated mucin domains, rendering them potential substrates of both isoenzymes T1 and T2 [11]. However, certain reproducible differences between T1- and T2-labeled band patterns indicated that substrates are specifically modified by individual isoenzymes.

Of note, the dependence on both AGX1[F383A] and BH-GalNAc-Ts for efficient labeling allowed us to assess

chemically modified GalNAc analogs. Suitable membrane-permeable precursors can be used to circumvent the GALK2 step if needed, but AGX1 engineering is necessary to deliver bumped UDP-GalNAc analogs. GALE-mediated epimerization to UDP-GlcNAc analogs can be suppressed by using branched acylamide side chains [12]. **(b)** Co-crystal structure of WT-AGX1 with UDP-GalNAc (PDB 1JV3) used to rationalize the F383G/A mutation that biosynthesizes bumped UDP-sugar analogs. **(c)** Structures of UDP-sugar analogs used in conjunction with GT engineering. GTs, glycosyltransferases.

Figure 3
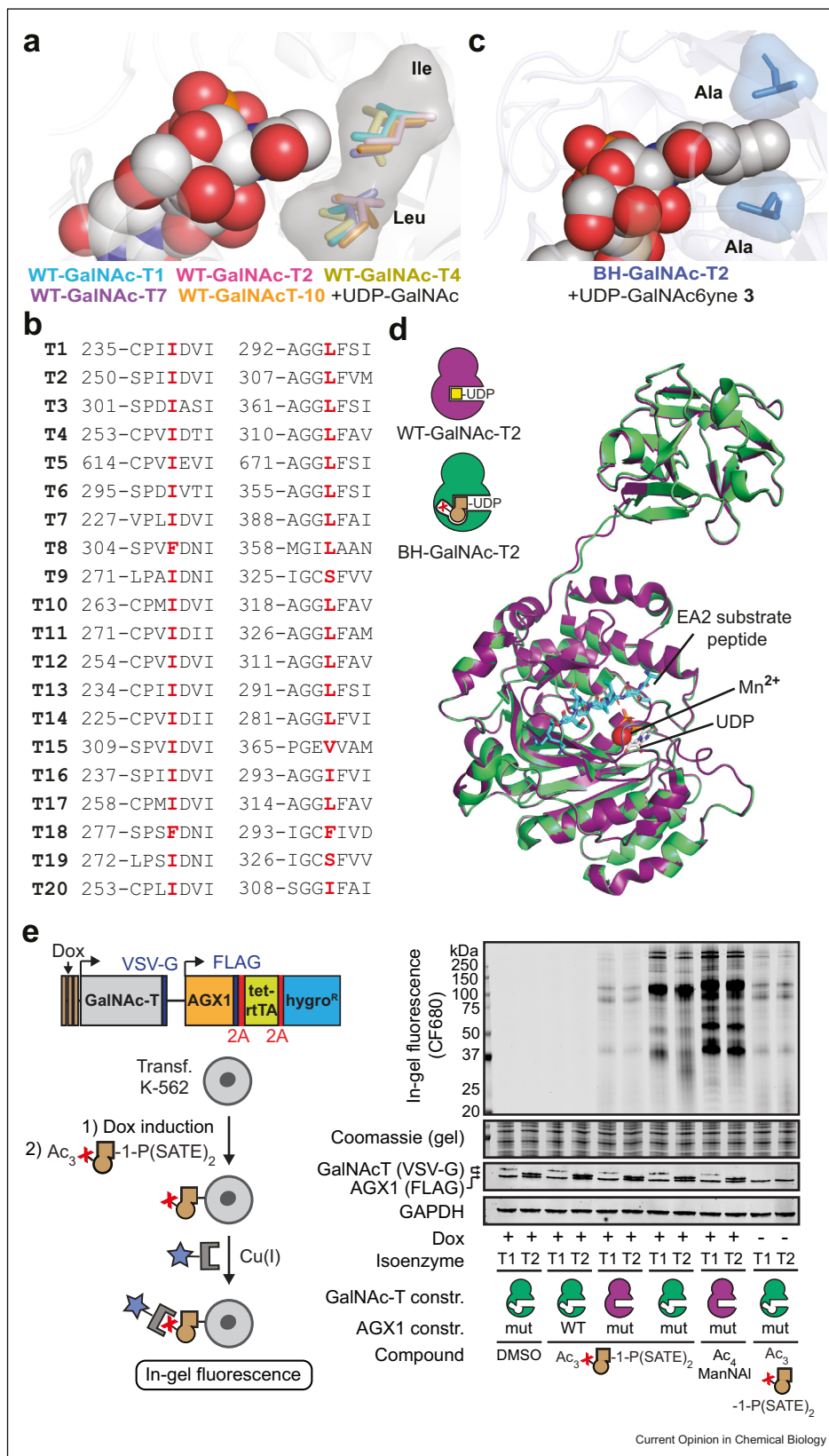
Structural basis for GalNAc-T bump-and-hole engineering. (a) Gatekeeper residues identified in the crystal structures of GalNAc-T1 (PDB 1XHB), T2 (PDB 4D0T), T4 (PDB 5NQA), T7 (PDB 6IWR), and T10 (PDB 2D7I). **(b)** Sequence alignment of gatekeeper residues in all 20 GalNAc-Ts. **(c)** Co-crystal

background labeling in the absence of either enzyme. We could thus rule out nonspecific fluorescence signal associated with an elimination—addition reaction on certain MOE reagents and confirm the validity of our strategy (Figure 3e) [48]. In contrast, we observed a notable background when bumped UDP-GalNAc analog **3** was biosynthesized but BH-GalNAc-Ts were not present. Nucleotide-sugar profiling revealed that a small portion of **3** is epimerized to the corresponding UDP-GlcNAc analog by GALE. Hypothesizing that N-linked glycans are the main destination of GlcNAc analogs on the cell surface, we substantially reduced background labeling by treating cell lysates with the de-N-glycanase PNGase F. Fluorescent labeling signal in the absence of Dox-induced BH-GalNAc-T expression resembled background signal, confirming the suitability of GalNAc-T BH engineering in K-562 cells (Figure 3e). We have since found that the use of branched acylamide side chains renders UDP-GalNAc derivatives resistant toward epimerization (Figure 2a), resulting in the probe UDP-GalNAzMe **6** (Figure 2c) that is specific for O-GalNAc glycans [12].

Creating gain-of-function reporter tools for the activity of individual GalNAc-Ts, the BH tactic provides a unique opportunity to directly analyze GalNAc-T-specific glycosylation sites when integrated into state-of-the art MS glycoproteomics workflows.

MS glycoproteomics

The field of MS glycoproteomics has seen a multitude of recent technical and conceptual advances [49—56]. However, the primary focus to map glycosylation sites still faces many challenges, especially when compared to traditional peptide-centered proteomics. One of the biggest challenges is the structural complexity of the glycome [57]. Glycans are made up of relatively few monosaccharide units, but the large number of possible branching sites, combined with the possible linkage stereochemistry, creates an enormous number of possible structures [58]. Further, these glycans can be found on several possible residues throughout the protein. In general, the complexity of glycans is compounded by issues with glycoprotein enrichment, instrumentation, and data analysis [59,60].

Enrichment

Glycoproteins are typically present at low abundance in complex samples, necessitating an enrichment step prior to MS analysis (Figure 4a) [61]. Types of enrichment broadly fall into three categories: affinity-based,

solid-phase extraction (SPE), and chemical methods. Affinity-based methods are most commonly used and involve lectins and/or antibodies to enrich glyco-conjugates. Lectins are glycan-binding proteins that enrich for particular glycan structures and have found use in numerous studies [62—65] but can be limited by low binding affinity and poor specificity. Generally, lectins are not well-suited for untargeted glyco-proteomics unless used in combination or succession.

To overcome these issues, SPE techniques have been developed, such as hydrophilic interaction chromatography (HILIC). Here, a hydrophilic stationary phase with a hydrophobic mobile phase preferentially retains hydrophilic glycopeptides [66—68]. When compared with lectins, HILIC has a much higher glycan promiscuity, enabling enrichment of a wide array of glycopeptides. In addition, several groups have investigated zwitterionic (ZIC-HILIC), [69—71], electrostatic repulsion liquid interaction chromatography (ERLIC) [72,73], and strong anion exchange (SAX-ERLIC) with HILIC to increase enrichment effectiveness and coverage of the glycoproteome. Other types of SPE include boronic acids that form covalent bonds with vicinal *cis*-diols on glycans [74—77], and titanium dioxide resins that are specific for negatively charged glycans [78—80].

Chemical enrichment methods involve derivatization or metabolic labeling of glycans, often followed by tagging the moiety with a secondary reporter, and then enriching for the glycopeptides [81]. One of the first examples was introduced by Aebersold et al. [82—85], where sialic acids were periodate-oxidized and enriched using hydrazide-functionalized beads. One drawback of this procedure is loss of glycan structural information. Enrichment of bioorthogonal MOE reagents is discussed below and in specialized reviews [6].

Fragmentation

Enriched glycopeptides are usually subjected to liquid chromatography (LC) followed by tandem MS analysis (Figure 4a). Various LC technologies have been reviewed elsewhere [86—89]. The most common workflows use reverse-phase high performance liquid chromatography (HPLC) coupled to electrospray ionization. Glycopeptides are then subjected to different types of fragmentation (i.e., tandem MS) to sequence the peptide, identify the glycan, and site-localize the glycosylation position. Collision-activated dissociation occurs when a peptide is subjected to

structure of BH-GalNAc-T2 (PDB 6NQT) with UDP-GalNAc6yne **3**. Gatekeeper residues are mutated to Ala to accommodate the aliphatic alkyne. (**d**) Superposition of WT-GalNAc-T2 (PDB 2FFU) and BH-GalNAc-T2 (PDB 6E7I) with EA2 substrate peptide (overlay of both structures), Mn^{2+} and UDP. (**e**) Metabolic labeling of cells transfected with AGX1 (WT or mut) and either GalNAc-T1 or T2 (WT or BH-mutant) constructs. Dox-inducible GalNAc-T expression was used in conjunction with feeding a caged precursor of UDP-GalNac6yne **3**. DMSO and the tagged sialic acid precursor Ac$_4$ManNAI served as negative and positive controls, respectively. Panel **E** reprinted from *Molecular Cell*, Vol 78/5, B. Schumann et al., Bump-and-Hole Engineering Identifies Specific Substrates of Glycosyltransferases in Living Cells, 824—834.e15, Copyright (2020), with permission from Elsevier.

collisions with helium atoms [90]. Beam-type collision-activated dissociation, called higher-energy collisional dissociation (HCD) on Orbitrap instruments, is similar but uses nitrogen instead of helium [91]. Each one of these collisions imparts vibrational energy to the peptide, eventually leading to cleavage of the most labile bond in the peptide. For glycopeptides, the most labile bonds are the glycosidic bonds (i) between peptide and glycan, and/or (ii) between individual monosaccharides (Figure 4c). Thus, the spectra generated are dominated by the losses of glycans from the peptide, protonated glycan species, and/or naked b-type and y-type peptide ions [86,92,93].

Electron transfer dissociation (ETD) was developed, in part, to overcome the issue of predominant glycan fragmentation and allow for unambiguous glycosylation site localization [94,95]. Because this reaction does not rely on collision for fragmentation, the glycan remains attached to the peptide backbone and allows for site-localization of glycosylation (Figure 4c) [31,96]. As a downside, ETD relies on high charge density, which often is not the case in large (i.e. >1000 m/z) glycopeptide precursor ions. To overcome this challenge, several groups have implemented supplemental activation in ETD, including EThcD [92,97,98] and activated ion ETD (AI-ETD) [63,99].

The two most commonly used fragmentation techniques in glycoproteomics are (i) stepped-collision energy HCD, and (ii) HCD-triggered ETD or EThcD (HCD-pd-ETD) [93]. The former is most common and recommended for N-glycoproteomics, whereas the latter is necessary for O-glycoproteomics. In stepped-collision energy HCD, three collision energies are used on the same precursor and product ions are accumulated and analyzed together [100,101]. In HCD-pd-ETD, HCD is first performed on all precursor ions. HexNAc (GlcNAc or GalNAc) ions will generate a fingerprint consisting of 6 masses that can then be used to trigger ETD or EThcD, which often allows for unambiguous site localization of the glycan (Figure 4b) [102,103]. Chemical glycan modifications contribute additional trigger ions that can be beneficial (*see* in the following context).

Data analysis
A fundamental challenge of glycoproteomics is data analysis. Often, spectra are input into a search algorithm to assist data interpretation. While most search engines were developed for the identification of unmodified peptides and struggle with glycopeptides [104], some take glycosylation into account. Byonic is the current glycoproteomic 'gold standard' [105], although some reports suggest that Protein Prospector may be better at site-localization of O-glycans [106]. Recently, 'O-Pair

Search' and 'MSFragger' programs have been introduced and display great promise in confident localization and scoring of glycoproteomic spectra [107,108]. Several thorough reviews have assembled available glyco-bioinformatic tools [109–111]. Importantly, at this point, expert manual interpretation is still required to confidently assign glycan composition and site of attachment [60]. Glycoproteomics is in need of new search algorithms and databases to overcome this time consuming and low-throughput step. Chemical tools could be fundamental to this end, for instance, through introducing isotope-based glycoproteomics workflows [29,112,113].

Other developments in the field of glycoproteomics
We envision that several techniques currently in development will have a strong impact on glyco-proteomics. For instance, advances in glycopeptide ionization are sorely needed to better analyze glyco-conjugates [114–116]. Furthermore, at the moment, glycoproteomics cannot inherently ascertain glycan stereochemistry and linkage. New methods, such as ultraviolet photodissociation [117,118], infrared spectroscopy–MS and/or infrared multiphoton dissociation [119,120], or data independent acquisition [121,122] might allow for this information to be ascertained in a single glycoproteomics run. Furthermore, ion-mobility MS helps separate isomeric structures, which can lead to less convoluted glycopeptide spectra [58,123,124]. Finally, to catapult glycoproteomics to new heights, we need more straightforward methods to validate MS results in order to allow for facile investigation into the biological relevance of glycosylation [125].

The promise of chemical precision tools to simplify MS glycoproteomics
Ever since their advent in the late 1990s, MOE reagents have allowed for an alternative view into the glycome [135]. Bioorthogonal groups are ideal for modern chemical proteomics techniques [6]. Most approaches rely on biotin-containing clickable handles for glyco-peptide enrichment before MS glycoproteomics. To specifically enrich glycopeptides, a neutravidin-based on-resin proteolytic digestion protocol is often used to remove any nonglycosylated peptides. The biotin-picolyl azide reagent **7** has an additional acid-cleavable diphenyldisiloxane group developed by Tirrell et al. [112,113,126,127], and since then has been applied to glycoproteomics (Figure 4d). In our hands, the bio-orthogonal handle offered advantages during data acquisition as well as validation, as new fingerprint ions appeared during HCD and were used to trigger ETD. In the ETD spectrum, another fingerprint ion (194 m/z) was present, which served to validate spectra and could

Figure 4
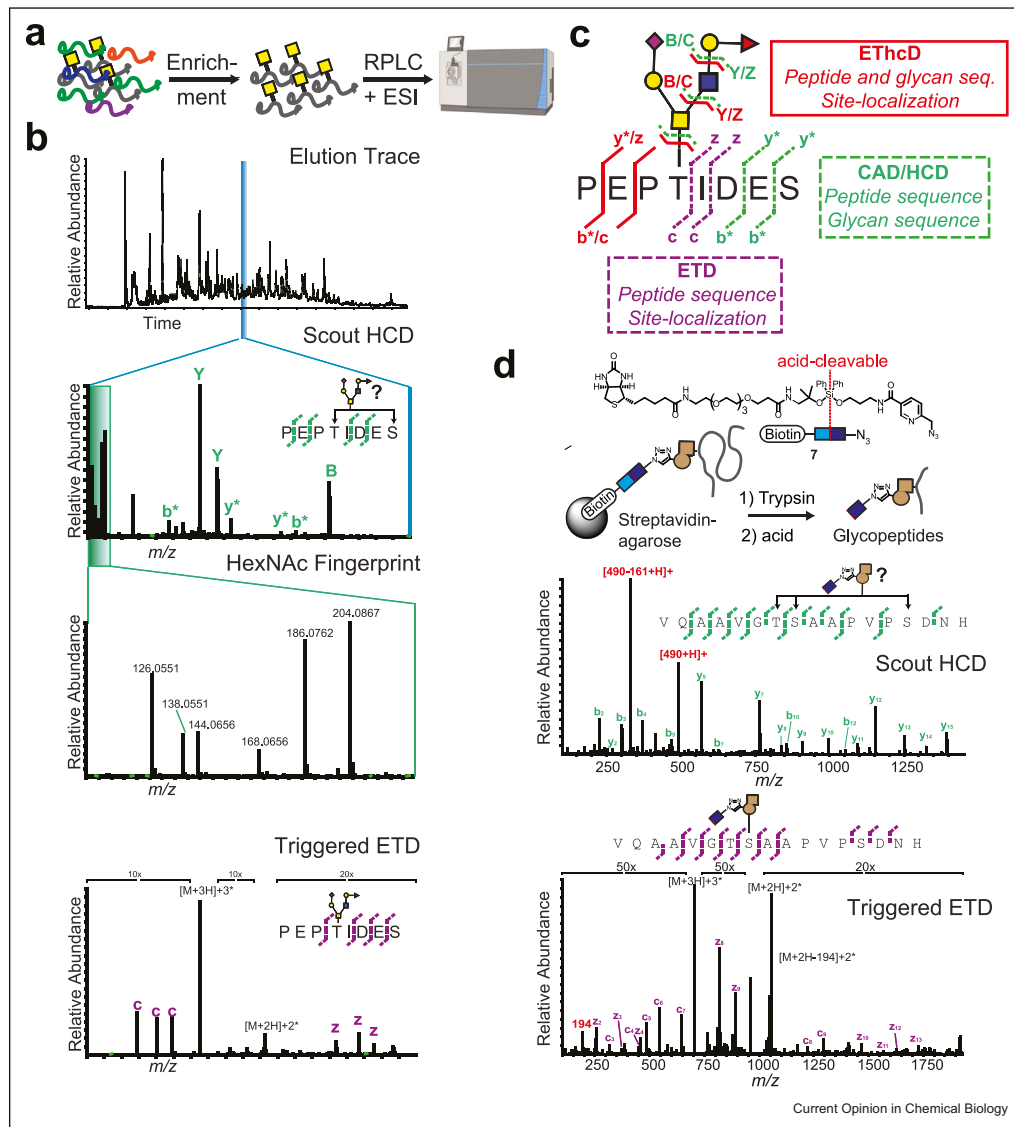
Glycoproteomic workflows enable analysis of the glycoproteome complexity. (a) Typical workflow before MS analysis. The complex sample contains unmodified peptides, differentially modified peptides, and glycopeptides. To enrich for the latter, samples are subjected to lectin columns, SPE extraction, or chemical enrichment procedures. The elution is then separated by RP-HPLC and electrospray ionized into the mass spectrometer. **(b)** Schematic of the instrumentation parameters often used in glycoproteomic analysis. The RP-HPLC elution trace is shown in the top panel and consists of a series of full mass spectra (MS1s). Typically, ions are selected in an abundance-dependent manner and subjected to HCD fragmentation (2nd panel). If a glycopeptide is present, a HexNAc fingerprint will be visible (3rd panel), which can then be used to trigger electron-based fragmentation (bottom panel). **(c)** Overview of glycopeptide fragmentation methods and the information they can provide. ETD (purple dashes) only fragments the peptide backbone, giving complementary c/z type ions with the glycan attached to the peptide. HCD (green dashes) fragments the peptide backbone as well as the glycosidic linkages, allowing for peptide and glycan sequencing, but often loses site-specificity. EThcD (red lines) combines the two techniques and allows for the most information to be gleaned from a single spectrum. Adapted from a study by Reiding et al. [92]. **(d)** Bump-and-hole chemical glycoproteomics provides a gain-of-function enrichment strategy (top). An HCD spectrum is shown under this, where 2 new fingerprint ions are present (491, 330) that can be used to trigger ETD (bottom). The ETD spectrum allows for site-localization of the modification and also demonstrates a new fingerprint ion (194) that can be used in search algorithms for more confident scoring. SPE, solid-phase extraction; RP, reverse-phase; HCD, higher-energy collisional dissociation; ETD, Electron transfer dissociation; HPLC, high performance liquid chromatography.

be used in search algorithms for more confident glycan assignment [11].

The advantages of chemical glycoproteomics fueled by the gain-of-function nature of BH engineering enabled us to directly map the glycosylation sites primed by two GalNAc-T isoenzymes. We used HepG2 cells engineered to lack native expression of T1 or T2 and complemented with BH-T1 and -T2, respectively [128]. Upon biosynthesis of UDP-GalNAc analog **3**, differential glycosylation sites primed by both isoenzymes were mapped. Among these, several had been suggested to be isoenzyme-specific by Schjoldager et al. [128] using SimpleCells and thereby benchmarked our approach, while others were new. The added advantage of the BH tactic showed in two pieces of data that had been previously impossible to accrue: (i) our data resolved ambiguity about isoenzyme specificity of apolipoprotein AI glycosylation, which had been impaired by events of compensation and competition within the GalNAc-T family; and (ii) the chemically modified GalNAc analog was elaborated up to the tetrasaccharide $(Neu5Ac)_2$-Hex-GalNAc, in accordance with the so-called di-Sialyl-T antigen as the largest abundant O-GalNAc glycan made in HepG2 cells [129]. As the enrichment process based on biotinylation should be unbiased to glycan substructures, mapping glycan elaboration offers exciting opportunities to potentially combine site annotation and glycome analysis, which is currently challenging on a glycoproteome-wide scale.

Outlook

The success of the BH tactic suggests important implications for studying one of the most fascinating and obscure GT families. As the tactic relies on mutating gatekeeper residues that are conserved throughout the family, in principle, it should be applicable to other isoenzymes. Multiple GalNAc-Ts have been crystallized over the last years [43,130−133], and available structures seem to confirm this notion. The field of chemical glycoproteomics is rapidly evolving, constantly increasing the sensitivity of MS analysis. Current challenges associated with optimizing glycopeptide enrichment, purification, and ionizability will be tackled within the next few years to truly deliver the promise of chemical precision tools to quantitative biology [134].

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Varki A, Gagneux P: **Biological functions of glycans**. In *Essentials of glycobiology*. Edited by Varki A, Cummings RD, Esko JD, *et al*. 3rd ed., Cold Spring Harbor Laboratory Press; 2017.

2. Raman R, Raguram S, Venkataraman G, Paulson JC, Sasisekharan R: **Glycomics: an integrated systems approach to structure-function relationships of glycans**. *Nat Methods* 2005, **2**:817−824.

3. Taylor ME, Drickamer K: *Introduction to glycobiology*. 3rd ed. 2011.

4. Jaiman A, Thattai M: **Golgi compartments enable controlled biomolecular assembly using promiscuous enzymes**. *eLife* 2020, **9**, e49573.

5. Zilmer M, Edmondson AC, Khetarpal SA, Alesi V, Zaki MS, Rostasy K, Madsen CG, Lepri FR, Sinibaldi L, Cusmai R, *et al*.: **Novel congenital disorder of O-linked glycosylation caused by GALNT2 loss of function**. *Brain* 2020, **143**:1114−1126.

6. Parker CG, Pratt MR: **Primer click chemistry in proteomic investigations**. *Cell* 2020, **180**:605−632.
** This paper reports the important contribution of bioorthogonal chemistry, specifically azide−alkyne click chemistry, to traditional proteomic methods in elucidating protein modifications.

7. Hanson SR, Hsu T-L, Weerapana E, Kishikawa K, Simon GM, Cravatt BF, Wong C-H: **Tailored glycoproteomics and glycan site mapping using saccharide-selective bioorthogonal probes**. *J Am Chem Soc* 2007, **129**:7266−7267.

8. Saxon E, Bertozzi CR: **Cell surface engineering by a modified staudinger reaction**. *Science* 2007, **287**:2007−2011.

9. Laughlin ST, Bertozzi CR: **Imaging the glycome**. *Proc Natl Acad Sci* 2009, **106**:12−17.

10. Patterson DM, Nazarova LA, Xie B, Kamber DN, Prescher JA: **Functionalized cyclopropenes as bioorthogonal chemical reporters**. *J Am Chem Soc* 2012, **134**:18638−18643.

11. Schumann B, Malaker SA, Wisnovsky SP, Debets MF, Agbay AJ, Fernandez D, Wagner LJS, Lin L, Li Z, Choi J, *et al*.: **Bump-and-Hole engineering identifies specific substrates of glycosyltransferases in living cells**. *Mol Cell* 2020, **78**:824−834.
** Successful application of the bump-and-hole tactic to identify products of individual GalNAc-Ts, specifically GalNAc-T1 and T2, in the living cell.

12. Debets MF, Tastan OY, Wisnovsky SP, Malaker SA, Angelis N, Moeckl LKR, Choi J, Flynn H, Wagner LJS, Bineva-Todd G, *et al*.: **Metabolic precision labeling enables selective probing of O-linked N-acetylgalactosamine glycosylation**. *Proc Natl Acad Sci USA* 2020, https://doi.org/10.1073/pnas.2007297117.
** Development of the precision tool UDP-GalNAzMe to specifically elucidate cancer-relevant Ser/Thr- N-acetylgalactosamine (O-GalNAc) glycosylation in living cells.

13. Möckl L, Pedram K, Roy AR, Krishnan V, Gustavsson AK, Dorigo O, Bertozzi CR, Moerner WE: **Quantitative super-resolution microscopy of the mammalian glycocalyx**. *Dev Cell* 2019, **50**:57−72.

14. Boyce M, Carrico IS, Ganguli AS, Yu SH, Hangauer MJ, Hubbard SC, Kohler JJ, Bertozzi CR: **Metabolic cross-talk allows labeling of O-linked β-N- acetylglucosamine-modified proteins via the N-acetylgalactosamine salvage pathway**. *Proc Natl Acad Sci* 2011, **108**:3141−3146.
* An important contribution in understanding the biosynthesis of chemically tagged nucleotide-sugars, specifically UDP-GlcNAz, through epimerization.

15. Alaimo PJ, Shogren-knaak MA, Shokat KM: **Chemical genetic approaches for the elucidation of signaling pathways**. *Curr Opin Chem Biol* 2001, **5**:360−367.

16. Hertz NT, Wang BT, Allen JJ, Zhang C, Dar AC, Burlingame AL, Shokat KM: **Chemical genetic approach for kinase-substrate mapping by covalent capture of thiophosphopeptides and analysis by mass spectrometry**. *Curr Protoc Chem Biol* 2014, **2**:15−36.

17. Yang C, Mi J, Feng Y, Ngo L, Gao T, Yan L, Zheng YG: **Labeling lysine acetyltransferase substrates with engineered enzymes and functionalized cofactor surrogates**. *J Am Chem Soc* 2014, **135**:7791−7794.

18. Blum G, Bothwell IR, Islam K, Luo M: **Profiling protein methyl-ation with cofactor Analogue containing terminal alkyne functionality**. *Curr Protoc Chem Biol* 2015, **5**:67−88.

19. Carter-O'Connell I, Jin H, Morgan RK, David LL, Cohen MS: **Engineering the substrate specificity of ADP-ribosyltransferases for identifying direct protein targets**. *Physiol Behav* 2017, **176**:139−148.

20. Wang L, Brock A, Herberich B, Schultz PG: **Expanding the ge-netic code of Escherichia coli**. *Science* 2001, **292**:498−501.

21. Luther KB, Schindelin H, Haltiwanger RS: **Structural and mechanistic insights into lunatic fringe from a kinetic anal-ysis of enzyme mutants**. *J Biol Chem* 2009, **284**:3294−3305.

22. Seto NOL, Palcic MM, Compston CA, Li H, Bundle DR, Narang SA: **Sequential interchange of four amino acids from blood group B to blood group A glycosyltransferase boosts catalytic activity and progressively modifies substrate recognition in human recombinant enzymes**. *J Biol Chem* 1997, **272**:14133−14138.

23. Tumbale P, Jamaluddin H, Thiyagarajan N, Acharya KR, Brew K: **Screening a limited structure-based library identifies UDP-GalNAc-specific mutants of α-1,3-galactosyltransferase**. *Glycobiology* 2008, **18**:1036−1043.

24. Marcus SL, Polakowski R, Seto NOL, Leinala E, Borisova S, Blancher A, Roubinet F, Evans SV, Palcic MM: **A single point mutation reverses the donor specificity of human blood group B-synthesizing galactosyltransferase**. *J Biol Chem* 2003, **278**:12403−12405.

25. Ramakrishnan B, Qasba PK: **Structure-based design of β1,4-**
   ⁎ **galactosyltransferase I (β4Gal-T1) with equally efficient N-acetylgalactosaminyltransferase activity: point mutation broadens β4Gal-T1 donor specificity**. *J Biol Chem* 2002, **277**: 20833−20839.
   First application of the bump-and hole approach to a GT to introduce reactive chemical functional groups; the catalytic activity of the bovine β-1,4-galactosyltransferase was re-programmed toward UDP-GalNAc analogs such as UDP-2-keto-galactose and UDP-GalNAz.

26. Boeggeman E, Ramakrishnan B, Kilgore C, Khidekel N, Hsieh-Wilson LC, Simpson JT, Qasba PK: **Direct identification of nonreducing GlcNAc residues on N-glycans of glycoproteins using a novel chemoenzymatic method**. *Bioconjug Chem* 2007, **18**:806−814.

27. Qasba PK, Boeggeman E, Ramakrishnan B: **Site-specific link-ing of biomolecules via glycan residues using glycosyl-transferases**. *Biotechnol Prog* 2008, **24**:520−526.

28. Vocadlo DJ, Hang HC, Kim E, Hanover JA, Bertozzi CR: **A chemical approach for identifying O-GlcNAc- modified proteins in cells**. *Proc Natl Acad Sci* 2003, **100**:9116−9121.

29. Woo CM, Iavarone AT, Spiciarich DR, Palaniappan KK, Bertozzi CR: **Isotope-targeted glycoproteomics (IsoTaG): a mass-independent platform for intact N- and O-glycopeptide discovery and analysis**. *Nat Methods* 2015, **12**:561−567.

30. Alfaro JF, Gong CX, Monroe ME, Aldrich JT, Clauss TRW, Purvine SO, Wang Z, Camp DG, Shabanowitz J, Stanley P, *et al*.: **Tandem mass spectrometry identifies many mouse brain O-GlcNAcylated proteins including EGF domain-specific O-GlcNAc transferase targets**. *Proc Natl Acad Sci* 2012, **109**: 7280−7285.

31. Wang Z, Udeshi ND, Malley MO, Shabanowitz J, Hunt DF, Hart GW: **Enrichment and site mapping of O-linked N-acetylglucosamine by a combination of chemical/enzymatic tagging, photochemical cleavage, and electron transfer dissociation mass spectrometry**. *Mol Cell Proteomics* 2010, **9**: 153−160.

32. Burnham-Marusich AR, Snodgrass CJ, Johnson AM, Kiyoshi CM, Buzby SE, Gruner MR, Berninsone PM: **Metabolic labeling of Caenorhabditis elegans primary embryonic cells with azido-sugars as a tool for glycoprotein discovery**. *PloS One* 2012, **7**, e49020.

33. Khidekel N, Arndt S, Lamarre-vincent N, Lippert A, Poulin-
   ⁎ kerstien KG, Ramakrishnan B, Qasba PK, Hsieh-wilson LC: **A chemoenzymatic approach toward the rapid and sensitive detection of O-GlcNAc posttranslational modifications**. *J Am Chem Soc* 2003, **125**:16162−16163.
   In this paper the mutant Y289L B4GalT1, coupled with the bio-orthogonal probe UDP-GalNAz, was applied for the first time to a rapid and sensitive detection of O-GlcNAc glycosylated proteins.

34. Thompson JW, Griffin ME, Hsieh-Wilson LC: **Methods for the detection, study, and dynamic profiling of O-GlcNAc glyco-sylation**. *Methods Enzymol* 2019, **598**:101−135.

35. Darabedian N, Thompson JW, Chuh KN, Hsieh-Wilson LC, Pratt MR: **Optimization of chemoenzymatic mass tagging by strain-promoted cycloaddition (SPAAC) for the determination of O-GlcNAc stoichiometry by western blotting**. *Biochemistry* 2018, **57**:5769−5774.

36. Pouilly S, Bourgeaux V, Piller F: **Evaluation of analogues of**
   ⁎ **GalNAc as substrates for enzymes of the mammalian GalNAc salvage pathway**. *ACS Chem Biol* 2012, **7**:753−760.
   This paper shows the catalytic efficiency toward GalNAc analogs of enzymes involved in the GalNAc salvage pathway such as GALK2, AGX1 and GalNAc-T1.

37. Yu S, Boyce M, Wands AM, Bond MR, Bertozzi CR, Kohler JJ:
   ⁎⁎ **Metabolic labeling enables selective photocrosslinking of O-GlcNAc-modified proteins to their binding partners**. *Proc Natl Acad Sci* 2012, **109**:4834−4839.
   For the first time cells are endowed with a mutated enzyme (AGX1 F383G) that allows to biosynthesize UDP-derivatives of bulky GlcNAc analogs that otherwise cannot be processed.

38. Rodriguez AC, Yu S, Li B, Zegzouti H, Kohler JJ, Corporation P: **Enhanced transfer of a photocrosslinking GlcNAc analog by an O-GlcNAc transferase mutant with converted substrate specificity**. *J Biol Chem* 2015, **290**:22638−22648.

39. Kingsley PD, Ten Hagen KG, Maltby KM, Zara J, Tabak LA: **Diverse spatial expression patterns of UDP-GalNAc: poly-peptide N-acetylgalactosaminyl-transferase family member mRNAs during mouse development**. *Glycobiology* 2000, **10**: 1317−1323.

40. Peng RQ, Wan HY, Li HF, Liu M, Li X, Tang H: **MicroRNA-214 suppresses growth and invasiveness of cervical cancer cells by targeting UDP-N-acetyl-α-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 7**. *J Biol Chem* 2012, **287**: 14301−14309.

41. Lavrsen K, Dabelsteen S, Vakhrushev SY, Levann AMR, Haue AD, Dylander A, Mandel U, Hansen L, Frodin M, Bennett EP, *et al*.: **De novo expression of human polypeptide N-acetylgalactosaminyltransferase 6 (GalNAc-T6) in colon adenocarcinoma inhibits the differentiation of colonic epithelium**. *J Biol Chem* 2018, **293**:1298−1314.

42. Khetarpal SA, Schjoldager KT, Christoffersen C, Edmondson AC, Reutter HM, Ahmed B, Peloso GM, Vitali C, Zhao W, Hanasoge AV, *et al*.: **Loss of function of GALNT2 lowers high density lipoproteins in humans, nonhuman primates, and rodents**. *Cell Metab* 2017, **24**:234−245.

43. Kubota T, Shiba T, Sugioka S, Furukawa S, Sawaki H, Kato R,
   ⁎ Wakatsuki S, Narimatsu H: **Structural basis of carbohydrate transfer activity by human UDP-GalNAc: polypeptide α-N-acetylgalactosaminyltransferase (pp-GalNAc-T10)**. *J Mol Biol* 2006, **359**:708−727.
   This paper was the first to show the crystal structure of a GalNAc-T in complex with the hydrolyzed donor substrate UDP-GalNAc and was instrumental to identification of gatekeeper residues.

44. Choi J, Wagner LJS, Timmermans SBPE, Malaker SA,
   ⁎⁎ Schumann B, Gray MA, Debets MF, Takashima M, Gehring J,

Bertozzi CR: **Engineering orthogonal polypeptide GalNAc-transferase and UDP- sugar pairs**. *J Am Soc Chem* 2019, **141**: 13442−13453.

The bump-and-hole engineering tactic was applied to GalNAc-T1, T2 and T10 to perform activity assays *in vitro* with different UDP-GalNAc analogs. This work paves the way for assessing activity and substrate specificity of individual GalNAc-Ts in a biological context.

45. Wang W, Hong S, Tran A, Jiang H, Triano R, Yan A, Liu Y, Chen X, Wu P: **Sulfated ligands for the copper(I)-catalyzed azide-alkyne cycloaddition**. *Chem Asian J* 2011, **6**:2796−2802.

46. Besanceney-Webler C, Jiang H, Zheng T, Feng L, Soriano del Amo D, Wang W, Klivansky LM, Marlow FL, Liu Y, Wu P: **Increasing the efficacy of bioorthogonal click reactions for bioconjugation: a comparative study**. *Angew Chem Int Ed* 2011, **50**:8051−8056.

47. Uttamapinant C, Tangpeerachaikul A, Grecian S, Clarke S, Singh U, Slade P, Gee KR, Ting AY: **Fast, cell-compatible click chemistry with copper-chelating azides for biomolecular labeling**. *Angew Chem* 2012, **51**:5852−5856.

48. Qin K, Zhang H, Zhao Z, Chen X: **Protein S-Glyco-Modification through an elimination-addition mechanism protein S-Glyco-Modification through an elimination-addition mechanism**. *J Am Chem Soc* 2020, **142**:9382−9388.

49. Trinidad JC, Schoepfer R, Burlingame AL, Medzihradszky KF: **N- and O-Glycosylation in the murine synaptosome**. *Mol Cell Proteomics* 2013, **12**:3474−3488.

50. Shah P, Wang X, Yang W, Eshghi ST, Sun S, Hoti N, Chen L, Yang S, Pasay J, Rubin A, *et al.*: **Integrated proteomic and glycoproteomic analyses of prostate cancer cells reveal glycoprotein alteration in protein abundance and glycosylation**. *Mol Cell Proteomics* 2015, **14**:2753−2763.

51. Medzihradszky KF, Kaasik K, Chalkley RJ: **Tissue-specific glycosylation at the glycopeptide level**. *Mol Cell Proteomics* 2015, **14**:2103−2110.

52. Yu A, Zhao J, Peng W, Banazadeh A, Williamson SD, Goli M, Huang Yifan, Mechref Y: **Advances in mass spectrometry-based glycoproteomics**. *Electrophoresis* 2018, **39**:3104−3122.

53. Rudd P, Karlsson NG, KhooKH, Packer NH: **Glycomics and glycoproteomics**. In *Essentials of glycobiology*. Edited by Varki A, Cummings RD, Esko JD, *et al.* 3rd ed., Cold Spring Harbor Laboratory Press; 2017.

54. Shajahan A, Heiss C, Ishihara M, Azadi P: **Glycomic and glycoproteomic analysis of glycoproteins-a tutorial**. *Anal Bioanal Chem* 2017, **409**:4483−4505.

55. Alagesan K, Everest-Dass A, Kolarich D: **Isomeric separation and characterisation of glycoconjugates**. *Adv Exp Med Biol* 2018, **1104**:77−99.

56. Li Q, Xie Y, Wong M, Lebrilla C: **Characterization of cell glycocalyx with mass spectrometry methods**. *Cells* 2019, **8**: 882.

57.. ** Khoo K-H: **Advances toward mapping the full extent of protein site-specific O-GalNAc glycosylation that better reflects underlying glycomic complexity**. *Curr Opin Struct Biol* 2019, **56**:146−154.

This review focuses on O-glycoproteomics, specifically, providing commentary on the unique challenges posed by this type of glycosylation and current efforts to overcome them.

58. Mookherjee A, Guttman M: **Bridging the structural gap of glycoproteomics with ion mobility spectrometry**. *Curr Opin Chem Biol* 2018, **42**:86−92.

59. Chen Z, Huang J, Li L: **Recent advances in mass spectrometry (MS)-based glycoproteomics in complex biological samples**. *Trends Anal Chem* 2019, **118**:880−892.

60. * Thaysen-Andersen M, Packer NH, Schulz BL: **Maturing glycoproteomics technologies provide unique structural insights into the N -glycoproteome and its regulation in health and disease**. *Mol Cell Proteomics* 2016, **15**:1773−1790.

This review summarizes the qualitative and quantitative developments and applications in N-glycoproteomics.

61. Chandler KB, Costello CE: **Glycomics and glycoproteomics of membrane proteins and cell-surface receptors: present trends and future opportunities**. *Electrophoresis* 2016, **37**: 1407−1419.

62. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB, Schjoldager KT-BG, Lavrsen K, Dabelsteen S, Pedersen NB, Marcos-SilvaL, Gupta R, Bennet EP, Mandel U, Brunak S, Wandall HH, Levery SB, Clausen H: **Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology**. *EMBO J* 2013, **32**:1478−1488.

63. Riley NM, Hebert AS, Westphall MS, Coon JJ: **Capturing site-specific heterogeneity with large-scale N-glycoproteome analysis**. *Nat Commun* 2019, **10**:1311.

64. Xiao H, Chen W, Smeekens JM, Wu R: **An enrichment method based on synergistic and reversible covalent interactions for large-scale analysis of glycoproteins**. *Nat Commun* 2018, **9**: 1692.

65. Trinidad JC, Barkan DT, Gulledge BF, Thalhammer A, Sali A, Schoepfer R, Burlingame AL: **Global identification and characterization of both O -GlcNAcylation and phosphorylation at the murine synapse**. *Mol Cell Proteomics* 2012, **11**:215−229.

66. Sun N, Wu H, Chen H, Shen X, Deng C: **Advances in hydrophilic nanomaterials for glycoproteomics**. *Chem Commun* 2019, **55**:10359−10375.

67. Jensen PH, Mysling S, Højrup P, Jensen ON: **Glycopeptide enrichment for MALDI-TOF mass spectrometry analysis by hydrophilic interaction liquid chromatography solid phase extraction (HILIC SPE)**. *Methods Mol Biol* 2013, **951**:131−144.

68. Scott NE, Parker BL, Connolly AM, Paulech J, Edwards AVG, Crossett B, Falconer L, Kolarich D, Djordjevic SP, Højrup P, Packer NH, Larsen MR, Cordwell SJ: **Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N -linked glycoproteome of campylobac**. *Mol Cell Proteomics* 2011, **10**:M000031−MCP201.

69. Calvano CD, Zambonin CG, Jensen ON: **Assessment of lectin and HILIC based enrichment protocols for characterization of serum glycoproteins by mass spectrometry**. *J Proteomics* 2008, **71**:304−317.

70. CaoW, Huang J, Jiang B, Gao X, Yang P: **Highly selective enrichment of glycopeptides based on zwitterionically functionalized soluble nanopolymers**. *Sci Rep* 2016, **6**:29776.

71. Alagesan K, Khilji SK, Kolarich D: **It is all about the solvent: on the importance of the mobile phase for ZIC-HILIC glycopeptide enrichment**. *Anal Bioanal Chem* 2017, **409**:529−538.

72. Sok Hwee Cheow E, Sim KH, de Kleijn D, Lee CN, Sorokin V, Sze SK: **Simultaneous enrichment of plasma soluble and extracellular vesicular glycoproteins using prolonged ultracentrifugation-electrostatic repulsion-hydrophilic interaction chromatography (PUC-ERLIC) approach**. *Mol Cell Proteomics* 2015, **14**:1657−1671.

73. Alpert AJ: **Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides**. *Anal Chem* 2008, **80**:62−76.

74. Chen W, Smeekens JM, Wu R: **A universal chemical enrichment method for mapping the yeast N -glycoproteome by mass spectrometry (MS)**. *Mol Cell Proteomics* 2014, **13**: 1563−1572.

75. Sparbier K, Wenzel T, Kostrzewa M: **Exploring the binding profiles of ConA, boronic acid and WGA by MALDI-TOF/TOF MS and magnetic particles**. *J Chromatogr B Analyt Technol Biomed Life Sci* 2006, **840**:29−36.

76. Zhang Q, Tang N, Brock JWC, Mottaz HM, Ames JM, Baynes JW, Smith RD, Metz TO: **Enrichment and analysis of nonenzymatically glycated peptides: boronate affinity chromatography coupled with electron-transfer dissociation mass spectrometry**. *J Proteome Res* 2007, **6**:2323−2330.

77. Xu Y, Wu Z, Zhang L, Lu H, Yang P, Webley PA, Zhao D: **Highly specific enrichment of glycopeptides using boronic acid-functionalized mesoporous silica**. *Anal Chem* 2009, **81**: 503–508.

78. Palmisano G, Eun Lendal S, Engholm-Keller K, Leth-Larsen R, Parker BL, Larsen MR: **Selective enrichment of sialic acid–containing glycopeptides using titanium dioxide chromatography with analysis by HILIC and mass spectrometry**. *Nat Protoc* 2010, **5**:1974–1982.

79. Larsen MR, Jensen SS, Jakobsen LA, Heegaard NHH: **Exploring the sialiome using titanium dioxide chromatography and mass spectrometry**. *Mol Cell Proteomics* 2007, **6**: 1778–1787.

80. Yan J, Li X, Yu L, Kin Y, Zhang X, Xue X, Ke Y, Liang X: **Selective enrichment of glycopeptides/phosphopeptides using porous titania microspheres**. *Chem Commun* 2010, **46**: 5488–5490.

81. •• Palaniappan KK, Bertozzi CR: **Chemical glycoproteomics**. *Chem Rev* 2016, **116**:14277–14306.
The authors provide an in-depth review on how chemical innovations combined with mass spectrometry have advanced our knowledge of the glycoproteome.

82. Zhang H, Li X, Martin DB, Aebersold R: **Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry**. *Nat Biotechnol* 2003, **21**:660–666.

83. Nilsson J, Rüetschi U, Halim A, Hesse C, Carlsohn E, Brinkmalm, Larson G: **Enrichment of glycopeptides for glycan structure and attachment site identification**. *Nat Methods* 2009, **6**: 809–811.

84. Halim A, Rüetschi U, Larson G, Nilsson J: **LC−MS/MS characterization of O-glycosylation sites and glycan structures of human cerebrospinal fluid glycoproteins**. *J Proteome Res* 2013, **12**:573–584.

85. Kurogochi M, Matsushista T, Amano M, Furukawa J, Shinohara Y, Aoshima M, Nishimura S-I: **Sialic acid-focused quantitative mouse serum glycoproteomics by multiple reaction monitoring assay**. *Mol Cell Proteomics* 2010, **9**:2354–2368.

86. •• Ruhaak LR, Xu G, Li Q, Goonatilleke E, Lebrilla CB: **Mass spectrometry approaches to glycomic and glycoproteomic analyses**. *Chem Rev* 2018, **118**:7886–7930.
The authors provide a thorough commentary on current glycomic and glycoproteomic techniques, describing the state of the art and efforts to advance the field.

87. Alagesan K, Everest-Dass A, Kolarich D: **Isomeric separation and characterisation of glycoconjugates**. In *Glycobiophysics*. Edited by Yamaguchi Y, Kato K, Springer; 2018.

88. Camperi J, Pichon V, Delaunay N: **Separation methods hyphenated to mass spectrometry for the characterization of the protein glycosylation at the intact level**. *J Pharm Biomed Anal* 2020, **178**:112921.

89. Huang B-Y, Yang C-K, Liu C-P, Liu C-Y: **Stationary phases for the enrichment of glycoproteins and glycopeptides: liquid Phase Separations**. *Electrophoresis* 2014, **35**:2091–2107.

90. Wells JM, McLuckey SA: **Collision-induced dissociation (CID) of peptides and proteins**. *Methods Enzymol* 2005, **402**:148–185.

91. Olsen JV, Macek B, Lange O, Makarov A, Horning S, Mann M: **Higher-energy C-trap dissociation for peptide modification analysis**. *Nat Methods* 2007, **4**:709–712.

92. • Reiding KR, Bondt A, Franc V, Heck AJR: **The benefits of hybrid fragmentation methods for glycoproteomics**. *Trends Anal Chem* 2018, **108**:260–268.
The authors provide excellent commentary on various glycopeptide fragmentation techniques, both alone and in combination.

93. •• Riley NM, Malaker SA, Driessen M, Bertozzi CR: **Optimal dissociation methods differ for N- and O-glycopeptides**. *J Proteome Res* 2020, https://doi.org/10.1021/acs.jproteome.0c00218.
A systematic and thorough investigation into which fragmentation technique is best for various glycopeptides. The study weighs the pros

and cons of each method and suggests optimal fragmentation techniques for both types of glycopeptide.

94. Coon JJ, Syka JEP, Shabanowitz J, Hunt DF: **Tandem mass spectrometry for peptide and protein sequence analysis**. *Biotechniques* 2005, **38**:519–523.

95. Syka JEP, Coon JJ, Schroeder MJ, Shabanowitz J, Hunt DF: **Peptide and protein sequence analysis by electron transfer dissociation mass spectrometry**. *Proc Natl Acad Sci* 2004, **101**: 9528–9533.

96. Myers SA, Daou S, Affar EB, Burlingame A: **Electron transfer dissociation (ETD): the mass spectrometric breakthrough essential for O -GlcNAc protein site assignments-a study of the O -GlcNAcylated protein Host Cell Factor C1**. *Proteomics* 2013, **13**:982–991.

97. Yang Y, Franc V, Heck AJR: **Glycoproteomics: a balance between high-throughput and in-depth analysis**. *Trends Biotechnol* 2017, **35**:598–609.

98. Yu Q, Canales A, Glover MS, Das R, Shi X, Liu Y, Keller MP, Attie AD, Li Lingiun: **Targeted mass spectrometry approach enabled discovery of O- glycosylated insulin and related signaling peptides in mouse and human pancreatic islets**. *Anal Chem* 2017, **89**:9184–9191.

99. Riley NM, Westphall MS, Coon JJ: **Activated ion-electron transfer dissociation enables comprehensive top-down protein fragmentation**. *J Proteome Res* 2017, **16**:2653–2659.

100. Cao L, Tolić N, Qu Y, Mend D, Zhao R, Zhang Q, Moore RJ, Zink EM, Lipton MS, Paša-Tolić L, Wu S: **Characterization of intact N- and O-linked glycopeptides using higher energy collisional dissociation**. *Anal Biochem* 2014, **452**:96–102.

101. Hinneburg H, Stavenhagen K, Schweiger-Hufnagel U, Pengelley S, Jabs W, Seeberger PH, Varón Silva D, Wuhrer M, Kolarich D: **The art of destruction: optimizing collision energies in quadrupole-time of flight (Q-TOF) instruments for glycopeptide-based glycoproteomics**. *J Am Soc Mass Spectrom* 2016, **207**:507–519.

102. Wu S-W, Pu T-H, Viner R, Khoo K-H: **Novel LC-MS2 product dependent parallel data acquisition function and data analysis workflow for sequencing and identification of intact glycopeptides**. *Anal Chem* 2014, **86**:5478–5486.

103. Singh C, Zampronio CG, Creese AJ, Cooper HJ: **Higher energy collision dissociation (HCD) product ion-triggered electron transfer dissociation (ETD) mass spectrometry for the analysis of N -linked glycoproteins**. *J Proteome Res* 2012, **11**: 4517–4525.

104. Darula Z, Medzihradszky KF: **Analysis of mammalian O-Glycopeptides-We have made a good start, but there is a long way to go**. *Mol Cell Proteomics* 2018, **17**:2–17.

105. Bern M, Kil YJ, Becker C: **Byonic: advanced peptide and protein identification software**. *Curr Protoc Bioinformatics* 2012, **40**. 13.20.1-13.20.14.

106. Pap A, Klement E, Hunyadi-Gulyas E, Darula Z, Medzihradszky KF: **Status report on the high-throughput characterization of complex intact O-glycopeptide mixtures**. *J Am Soc Mass Spectrom* 2018, **29**:1210–1220.

107. Lu L, Riley NM, Shortreed MR, Bertozzi CR, Smith LM: **O-pair search with MetaMorpheus for O-glycopeptide characterization**. *bioRxiv* 2020:102327, https://doi.org/10.1101/2020.05.18.102327. 05.18.

108. Polasky DA, Yu F, Teo GC, Nesvizhskii AI: **Fast and comprehensive N- and O-glycoproteomics analysis with MSFragger-glyco**. *bioRXiv* 2020:102665, https://doi.org/10.1101/2020.05.18.102665. 05.18.

109. Hu H, Khatri K, Zaia J: **Algorithms and design strategies towards automated glycoproteomics analysis: algorithms and design strategies**. *Mass Spectrom Rev* 2017, **36**:475–498.

110. • Abrahams JL, Taherzadeh G, Jarvas G, Guttman A, Zhou Y, Campbell MP: **Recent advances in glycoinformatic platforms for glycomics and glycoproteomics**. *Curr Opin Struct Biol* 2020, **62**:56–69.

Excellent summary of the current state of bioinformatic platforms in glycobiology and mass spectrometry.

111. Bennun SV, Baycin Hizal D, Heffner K, Can O, Zhang H, Betenbaugh MJ: **Systems glycobiology: integrating glycogenomics, glycoproteomics, glycomics, and other 'omics data sets to characterize cellular glycosylation processes**. *J Mol Biol* 2016, **428**:3337–3352.

112. Woo CM, Felix A, Byrd WE, Zuegel DK, Ishihara M, Azadi P, Iavarone AT, Pitter SJ, Bertozzi CR: **Development of IsoTaG, a chemical glycoproteomics technique for profiling intact N- and O-glycopeptides from whole cell proteomes**. *J Proteome Res* 2017, **16**:1706–1718.

113. Qin K, Zhu Y, Qin W, Gao J, Shao X, Wang YL, Zhou W, Wang C, Chen X: **Quantitative profiling of protein O-GlcNAcylation sites by an isotope-tagged cleavable linker**. *ACS Chem Biol* 2018, **13**:1983–1989.

114. Plomp R, Bondt A, de Haan N, Rombouts Y, Wuhrer M: **Recent advances in clinical glycoproteomics of immunoglobulins (igs)**. *Mol Cell Proteomics* 2016, **15**:2217–2228.

115. Alagesan K, Kolarich D: To enrich or not to enrich: enhancing (glyco)peptide ionization using the CaptiveSpray nanoBoosterTM. bioRxiv 597922, doi:10.1101/597922.

116. Stavenhagen K, Kayili HM, Holst S, Koeleman CAM, Engel R,
* Wouters D, Zeerleder S, Sajih B, Wuhrer M: **N- and O -glycosylation analysis of human C1-inhibitor reveals extensive mucintype O -glycosylation**. *Mol Cell Proteomics* 2018, **17**:1225–1238.
This study is an excellent example of a targeted glycoproteomic effort. Here, the authors identify both N- and O-glycosylation on C1-INH using a combination of PGC separation, various enzymatic treatment, and thorough data analaysis.

117. Madsen JA, Joon Ko B, Xu H, Iwashkiw JA, Robotham SA, Shaw JB, Feldman MF, Brodbelt JS: **Concurrent automated sequencing of the glycan and peptide portions of O -linked glycopeptide anions by ultraviolet photodissociation mass spectrometry**. *Anal Chem* 2013, **85**:9253–9261.

118. Ko BJ, Brodbelt JS: **Comparison of glycopeptide fragmentation by collision induced dissociation and ultraviolet photodissociation**. *Int J Mass Spectrom* 2015, **377**:385–392.

119. Gray CJ, Compagnon I, Flitsch SL: **Mass spectrometry hybridized with gas-phase InfraRed spectroscopy for glycan sequencing**. *Curr Opin Struct Biol* 2020, **62**:121–131.

120. Schindler B, Barnes L, Gray CJ, Chambert S, Flitsch SL, Oomens J, Daniel R, Allouche AR, Compagnon I: **IRMPD spectroscopy sheds new (infrared) light on the sulfate pattern of carbohydrates**. *J Phys Chem* 2017, **121**:2114–2120.

121. Zhou C, Schulz BL: **Glycopeptide variable window SWATH for improved data independent acquisition glycoprotein analysis**. *Anal Biochem* 2020, **597**:113667.

122. Ye Z, Mao Y, Clausen H, Vakhrushev SY: **Glyco-DIA: a method for quantitative O-glycoproteomics with in silico-boosted glycopeptide libraries**. *Nat Methods* 2019, **16**:902–910.

123. Chen Z, Glover MS, Li L: **Recent advances in ion mobility−mass spectrometry for improved structural characterization of glycans and glycoconjugates**. *Curr Opin Chem Biol* 2018, **42**:1–8.

124. Izaham ARA, Ang C, Nie S, Bird LE, Williamson NA, Scott NE: **What are we missing by using hydrophilic enrichment? Improving bacterial glycoproteome coverage using total proteome and FAIMS analysis**. *bioRxiv* 2020:216903, https://doi.org/10.1101/2020.07.22.216903. 07.22.

125. Narimatsu H, Kaji H, Vakhrushev SY, Clausen H, Zhang H, Noro E, Togayachi A, Nagai-Okatani C, Kuno A, Zou X, Cheng L, Tao S-C, Sun Y: **Current technologies for complex glycoproteomics and their applications to biology/disease-driven glycoproteomics**. *J Proteome Res* 2018, **17**:4097–4112.

126. Szychowski J, Mahdavi A, Hodas JJL, Bagert JD, Ngo JT, Landgraf P, Dieterich DC, Schuman EM, Tirrell DA: **Cleavable biotin probes for labeling of biomolecules via azide-alkyne cycloaddition**. *J Am Chem Soc* 2010, **132**:18351–18360.

127. Miyamoto DK, Flaxman HA, Wu HY, Gao J, Woo CM: **Discovery of a Celecoxib Binding Site on Prostaglandin e Synthase (PTGES) with a Cleavable Chelation-Assisted Biotin Probe**. *ACS Chem Biol* 2019, **14**:2527–2532.

128. Schjoldager KT, Joshi HJ, Kong Y, Goth CK, King SL,
* Wandall HH, Bennett EP, Vakhrushev SY, Clausen H: **Deconstruction of O-glycosylation-GalNAc-T isoforms direct distinct subsets of the O-glycoproteome**. *EMBO Rep* 2015, **16**:1713–1722.
A seminal paper describing differential O-GlaNAc glycoproteomics of GalNAc-T KO cell lines. These data are a great reference for other glycoproteomics approaches.

129. Fujitani N, Furukawa J ichi, Araki K, Fujioka T, Takegawa Y, Piao J, Nishioka T, Tamura T, Nikaido T, Ito M, et al.: **Total cellular glycomics allows characterizing cells and streamlining the discovery process for cellular biomarkers**. *Proc Natl Acad Sci* 2013, **110**:2105–2110.

130. Rivas MD Las, Lira-Navarrete E, Daniel EJP, Companõn I,
* Coelho H, Diniz A, Jiménez-Barbero J, Peregrina JM, 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**:1959.
The first crystal structure of GalNAc-T4 to provide further evidence for conservation of the residues framing the catalytic site between different isoenzymes.

131. De Las Rivas M, Paul Daniel EJ, Coelho H, Lira-Navarrete E,
* Raich L, Companõn I, Diniz A, Lagartera L, Jiménez-Barbero J, Clausen H, et al.: **Structural and mechanistic insights into the catalytic-domain-mediated short-range glycosylation preferences of GalNAc-T4**. *ACS Cent Sci* 2018, **4**:1274–1290.
Important insight into the short- and long-range directing effects of substrate GalNAc glycosylation on GalNAc-T activity

132. de las Rivas M, Paul Daniel EJ, Narimatsu Y, Companõn I,
* Kato K, Hermosilla P, Thureau A, Ceballos-Laita L, Coelho H, Bernadó P, et al.: **Molecular basis for fibroblast growth factor 23 O-glycosylation by GalNAc-T3**. *Nat Chem Biol* 2020, **16**:351–360.
The first crystal structure of GalNAc-T3, further substantiating the structural homology between different isoenzymes.

133. Fernandez AJ, Daniel EJP, Mahajan SP, Gray JJ, Gerken TA,
* Tabak LA, Samara NL: **The structure of the colorectal cancerassociated enzyme GalNAc-T12 reveals how nonconserved residues dictate its function**. *Proc Natl Acad Sci* 2019, **116**:20404–20410.
The first crystal structure of GalNAc-T12 that is closely related to GalNAc-T4, providing a structural basis for isoenzyme similarities and differences.

134. Zol-Hanlon MI, Schumann B: **Open questions in chemical glycobiology**. *Commun Chem* 2020, **3**, https://doi.org/10.1038/s42004-020-00337-6.

135. Mahal LK, Yarema KJ, Bertozzi CR: **Engineering chemical reactivity on cell surfaces through oligosaccharide biosynthesis**. *Science* 1997, **276**:1125–1128, https://doi.org/10.1126/science.276.5315.1125.