RSCPublishing   Chemical Science

Realizing the Promise of Chemical Glycobiology

Journal:	*Chemical Science*
Manuscript ID:	SC-MRV-04-2013-050877
Article Type:	Minireview
Date Submitted by the Author:	01-Apr-2013
Complete List of Authors:	Wang, Lai-Xi; University of maryland School of Medicine, Biochemistry & Molecular Biology Davis, Benjamin; University of Oxford, Department of Chemistry

SCHOLARONE™
Manuscripts

Abstract

Chemical glycobiology is emerging as one of the most uniquely powerful sub-disciplines of chemical biology. The previous scarcity of chemical strategies and the unparalleled structural diversity have created a uniquely fertile ground that is both rich in challenges and potentially very profound in implications. Glycans (oligosaccharides, polysaccharides, and glycoconjugates) are everywhere in biological systems and yet remain disproportionately neglected – reviews highlighting this 'Cinderella status' abound. Yet, the last two decades have witnessed tremendous progress, notably in chemical and chemoenzymatic synthesis, 'sequencing' and arraying, metabolic engineering and imaging. These vital steps serve to highlight not only the great potential but just how much more remains to be done. The vast chemical (and functional) space of glycans remains to be truly explored. Top-down full-scale glycomic and glycoproteomic studies coupled with hypothesis-driven, bottom-up innovative chemical strategies will be required to properly realize the potential impact of glycoscience on human health, energy, and economy. In this review, we cherry-pick far-sighted advances and use these to identify possible challenges, opportunities and avenues in chemical glycobiology.

Introduction

25 years on from its first concise definition [*1*], the field of Glycobiology has continued to successfully highlight the myriad of complex processes in which carbohydrates play a vital role. It becomes clear that glycans, in the form of oligosaccharides, polysaccharides, glycoproteins, glycolipids, proteoglycans and other glycoconjugates can be key players in a number of important biological (often but not exclusively) recognition processes: intracellular trafficking, cell adhesion, development, cancer progression, host-pathogen interaction, and immune response, to name a few [*2-6*]. They may now be reasonably added to the 'central dogma molecules' (nucleic acids, proteins) as playing a distinct (if less immediate) role in information transfer [*7*] – they are an additional 'language of life' (**Fig. 1**), whose roles are constantly emerging [*8*]. Yet decoding this language has barely started. In comparison with proteins and nucleic acids that contain linear sequences of building blocks, glycans frequently form branching structures and the inter-residue linkages (glycosidic) can exist in different configurations (anomeric isomers). The biosynthetic assembly of glycans is also more complex – it is not template-driven, as is largely the case for proteins and nucleic acids – and involves complex sugar chain processing and trimming under the action of a series of competitive enzymes along secretory pathways. As a result, glycans are generally more diverse in structure and are often present as heterogeneous mixtures that are difficult to separate for detailed structure-activity relationship studies. However, this structural complexity, which has often been blamed for the slow development of hypothesis testing within the field, also offers powerful opportunities in the design of molecular experiments that will unpick the mechanism of this Biology. In this way Chemistry offers one unique, and as yet largely unrealized, strategy for dissecting this complexity where strictly biological approaches may fail.
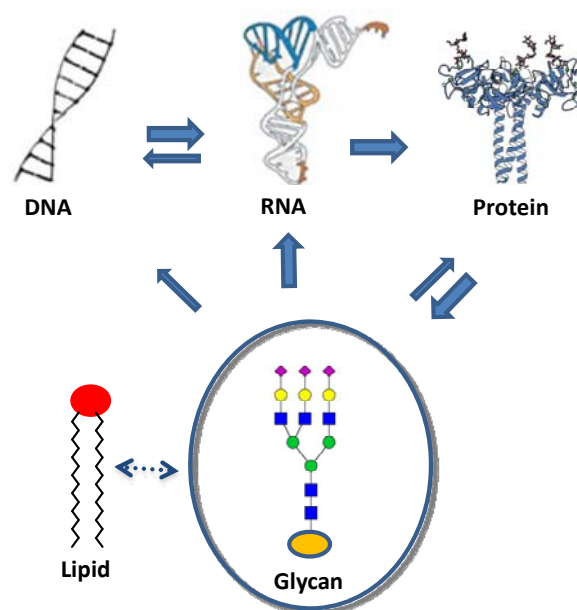
Fig. 1 A modified central dogma of biological information flow

The increasing importance and realization of Chemical Biology as a field have seen some successful analyses of its intellectual thrust. We use here a concept of the study of Biological questions and problems through molecular insight. The current move, perceived by many, of Biology as a field towards the cellular and organismal is vital but at the same time may have created an imbalance in terms of skill sets and lack of molecular knowledge. This creates a current opportunity for the Chemist. The aesthetic of molecular synthesis is a coherent and self-sufficient goal that does not need to justify itself in the context of other disciplines. Indeed, without genuine passion for the biological question, the synthesis of every permutation of structure ('collection of countless numbers of grapes' [Bacon]) is a strategy that we have learnt all-too-slowly over the last decades will not and does not work. This is yet more pertinent for the field of Glycoscience, where the number of permutations is simply too vast even for relatively simple structures [9]. The dazzling variety of potential configurational permutations means that chemists have an essential and powerful role in posing the 'right' molecular questions in Biology. To us this role sets Chemical Glycobiology above other domains as a near-perfect testing ground for how the chemical and biological sciences will evolve – the exciting depth and breadth of the challenges that we face as Glycoscience precludes the

arrogance of any one discipline or sub-discipline. The recently released US National Academy of Sciences report, entitled "Transforming the glycosciences: a roadmap for the future", offers one blueprint for future studies in Glycosciences. It too emphasizes the essence of a multi-disciplinary approach to problems in glycobiology (http://www.nap.edu/catalog.php?record_id=13446). Thus, all of the more-narrow, single-focus strategies that are (dumbly) optional in other fields (even other areas of Chemical Biology) cannot stand alone to answer questions in Glycoscience. This will forge superbly exciting strategies and here we guess at how these might develop. With technical and conceptual development, the time is now ripe to bring glycoscience to the mainstream of research.
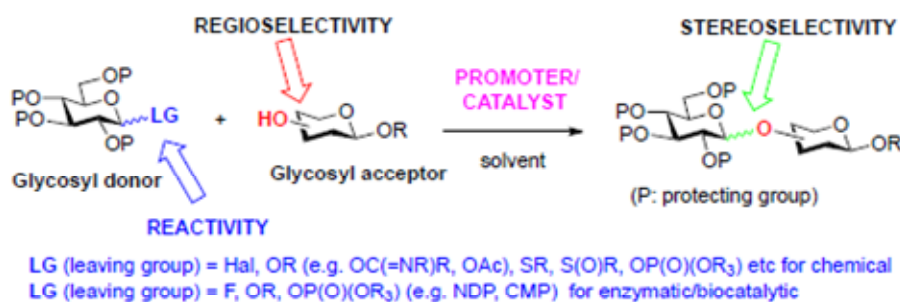
Chemical Glycobiology, as broadly defined, explores chemical approaches and concepts for deciphering and understanding the function of entire glycomes and related partners. Remarkable progress has been achieved within the last two decades, and this has been discussed in a number of excellent recent reviews [*10-16*]. While we should be proud of this, we must be also fully aware that the vast structural and functional space of glycans remains to be truly explored. The purpose of this speculative review is therefore not to recapitulate and list the many studies (and reviews) that comprehensively cover what has been tried but instead to speculate on challenges that remain.

The generalities of sugar synthesis and glycosylation

Synthesis of glycans and related chemical probes is essential if we are to precisely decipher structure, biosynthesis, metabolism and function in biological systems. The relative importance of synthetic glycans in glycobiology is exacerbated since homogeneous glycans (often found as part of glycoconjugates) are typically difficult to obtain from natural sources because of their structural micro-heterogeneity and, in many cases, their natural scarcity. Such synthetic glycans are also starting to play essential roles in developing carbohydrate-based therapeutics, such as vaccines [*17*]. However, the issues involved in the construction of complex oligosaccharides and glycoconjugates, including branching structures, anomeric stereochemistry, regiochemistry (even in linear oligomers), and coupling efficiency, pose many synthetic challenges. The complexity and diversity of oligosaccharides

5

and glycoconjugates demand the use of a range of the most efficient synthetic methods with sufficient flexibility and breadth of use, regardless of whether these are chemical, enzymatic, or chemoenzymatic.
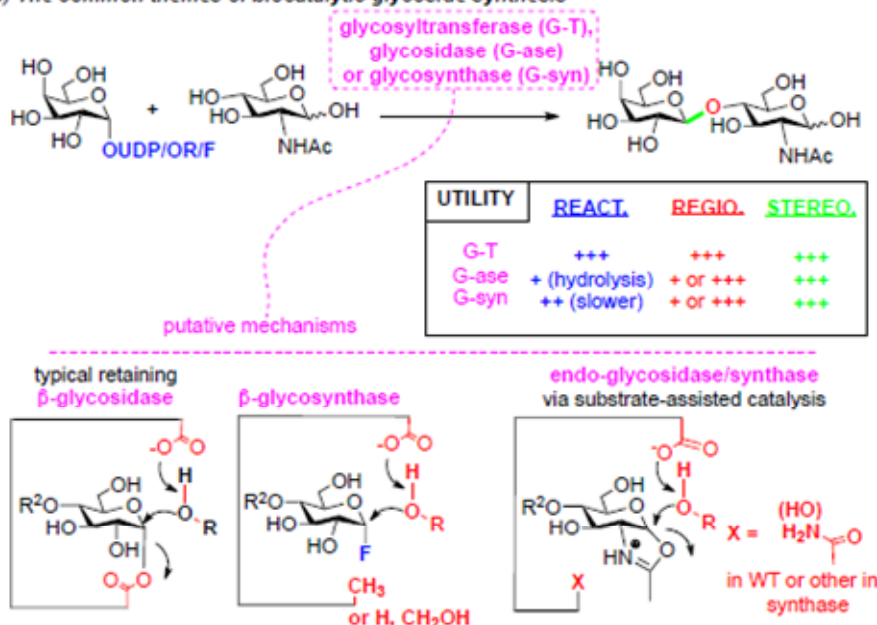

Fig. 2 Chemical and enzymatic synthesis of glycosides and oligosaccharides. a) General issues in glycoside synthesis; b) the common themes of biocatalytic glycoside construction.

The general aspects of chemical glycosylation and major glycosylation methods have been well reviewed and discussed elsewhere [13,18-21] but key aspects are worth re-iterating. Carbohydrate chemistry often takes selective protection and deprotection methods to their most comprehensive and elegant heights to achieve regio-selectivity in glycosylation [13,18-22]. This lack

6

of 'atom efficiency', of course, begs strategic questions but nonetheless it has proven a successful

reliable approach especially when coupled with novel glycoside-forming strategies. These include

manipulations of stereoelectronic effects, neighboring/remote protecting group participation, and

conformational constraints – all usefully explored to control stereoselectivity at the anomeric center. In

addition, there is a continuous effort to develop glycosyl donors with novel leaving groups and new

promoters/catalysts in order to enhance glycosylation efficiency, and to tune reactivity often with

useful knock on consequences for the better control of stereochemistry (**Fig. 2**). The progress in

chemical synthesis has been remarkable, as exemplified by the successful total chemical synthesis of

complex heparin oligosaccharide-based drugs (with well-defined sequence and substitution patterns)

[23], highly complex carbohydrate-containing natural products [24], and even homogeneous natural

glycoproteins [25-28]. From those looking-in from outside the field, such effort spent on the creation of

what is in essence just a mixed acetal or ketal linkage may seem a stunning self-indulgence for the

carbohydrate community to still be so heavily involved in more than 100 years after the first realization

of a real glycosylation reaction. Yet, we must be clear, there is no general synthetic method available

that can cover all of the diverse structures of glycans and glycoconjugates and, in this era of

oversimplified chemical strategy ("Click" and "automated synthesis"), it is pertinent to realize that each

glycan target may pose unique challenges of reactivity and selectivity. More 'programmable' (and

therefore automatable) strategies remain one option to create more general approaches, whilst on the

other hand (one-off) 'tricks' in regioselectivity and stereoselectivity remain attractive as bespoke

solutions to target structures. Even today, there remains an argument on what active intermediates

may be involved in fundamental glycosylation reactions and how the reactivity of the active species

contributes to the control of stereo-selectivity [29]. Innovations in achieving near perfect stereocontrol

in glycosylation remain as urgently sought-after goals [30] and the mechanistic ambiguities of

reactions at acetals ensures that there is still rich chemistry to be explored in glycosidic bond

formation.

Enzymatic and chemoenzymatic glycoside synthesis and manipulation

The need to simultaneously control reactivity, regio-selectivity, and stereo-selectivity, some might argue, has created a double-bind for carbohydrate synthesis: such complexity requires the greatest use of the available tools (e.g., extensive protecting group usage, arsenals of donors) and yet it is increasingly and strikingly atom inefficient. In addition, much of the role of the protecting group in glycan synthesis is to provide the 'grease' to solubilize the substrate in a manner suitable for our traditional organic solvents. In contrast, nature has devised many thousands of biocatalysts (enzymes) to achieve often perfect control of stereo- and regio-selectivity and usually high efficiency in transformations for glycan assembly in aqueous solution (arguably the ideal solvent for glycans), without the need of protecting groups. Thus, there seems no reason why we chemists should not embrace such special and efficient tools in our synthetic business and we would argue that synthetic carbohydrate scientists have been forward-thinking in their adoption of biocatalysis. In particular, combined processes (incorporating enzymatic transformations as key steps in a chemical synthesis), the so-called chemoenzymatic approach, can rapidly simplify synthetic schemes [*31*].

In theory, all types of enzymes involved in glycan biosynthesis, modification, and metabolism can be borrowed for synthesis. The opportunities for catalysis are huge since the numbers of enzymes involved in glycan assembly and metabolism far exceed those involved in the transformations of any other type of biomolecule. Glycosyltransferases are the natural enzymes responsible for constructing glycosidic bonds. Their high selectivity in glycosidic bond formation and the tolerance of certain modifications (within limits) on the donor and acceptor substrates make glycosyltransferases unique biocatalyst for synthesizing natural and selectively modified glycans [*32*]. Sugar nucleotide regeneration systems have been used now for some to by-pass the stoichiometric use of the expensive sugar nucleotides but this creation of 'catalytic donors' has not gained the prominence that it perhaps should in the synthetic community. The perceived intolerance of some enzymes can, in fact, sometimes be 'forced' *in vitro* using concentrations above the K_M; this has actually allowed broad use of quite different donors in many cases. Moreover, mutational studies

8

have led to mutant enzymes that demonstrate broader substrate tolerance, capable of accepting modified and altered substrates [31,33]. Furthermore, enzymatic synthesis should not be limited to purely *in vitro* enzymatic transformations. For example, functional transfer of glycan biosynthetic genes or even genes cluster into bacterium *Escherichia coli* (*E. coli*) has enabled a large-scale *in situ* production of heparin-like oligosaccharides, sialyloligosaccharides, ganglioside oligosaccharides and human milk oligosaccharides [34,35]. Thus, *E. coli*, biotechnology's "work horse", as well as other promising host systems (yeast, insect cells, and plant cells) could well be further explored as factories for producing diverse glycans on large scale (perhaps this form of Process Microbiology will become to be known as 'Synthetic Glycobiology' akin to the claiming of the term 'Synthetic Biology').

Glycosidases, the natural function of which is to hydrolyze glycosidic bonds in glycan metabolism, have strong potential in transglycosylation by virtue of their ability to manipulate bond formation at the anomeric center [36-38]. One of the most striking advance in this area is the invention of glycosynthases [39,40] : logically-designed glycosidase mutants that are devoid of product hydrolysis activity (an immediately obvious limitation in glycosidase-associated synthesis) but that are able to take activated donor substrates (e.g., glycosyl fluorides or sugar oxazolines) for transglycosylation [36-38]. Since the report of the first glycosynthase in 1998 [39,40] , various glycosynthases have been constructed from glycosidases belonging to more than a dozen of glycoside hydrolase families. Yet, this reflects only a small fraction of many thousands of existing glycosidases; often the principles of glycosidase design do not translate. The reasons for this likely lie in our (often necessary) assumptions regarding modes of such catalysis; more detailed mechanistic studies are undoubtedly required to unveil why some enzymes have more potent transglycosylation activities than others (**Fig. 2b**).

In addition to the enzymes exploited in glycosidic bond formation, a number of glycan-modifying enzymes, such as sulfotransferases, acetyltransferases, oxidases, and epimerases, to name a few, also bear high value for post-glycosylation modifications of oligosaccharides and glycoconjugates [31,41]. A very large space of enzyme capacity for synthesis remains to be explored.

9

Two notable examples showcase the power of the chemoenzymatic approach. One is the recently reported, *in vitro* chemoenzymatic synthesis of homogeneous ultralow molecular weight heparins (heparin hepta-saccharides) with well-defined sequence and sulfation patterns, which was fulfilled by a combined use of glycosyltransferases, sulfotransferases, and epimerase, starting from simple disaccharides [*42*]. The synthesis took only 10 to 12 steps and resulted in a 45% overall yield. This was achieved on the basis of an in depth understanding of the biosynthetic pathways, together with the efficient cloning and expression of those enzymes involved [*42*]. In comparison, the pure chemical synthesis of the related anticoagulant drug, Arixtra (a heparin pentasaccharide), a landmark in target chemical oligosaccharide synthesis, took more than 50 steps leading to a low (1%) overall yield [*23*]. The other example is the *ab initio* glycan remodeling of glycoproteins via endoglycosidase-catalyzed deglycosylation and block oligosaccharide transfer [*43,44*]. This highly convergent approach enables a quick assembly of well-defined glycoforms, such as the homogeneously Fc-glycosylated IgG antibodies, for structural and functional studies.
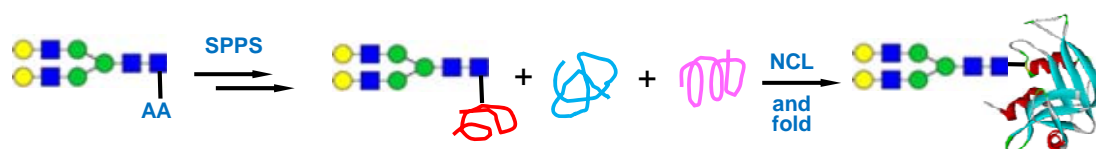
While often specificity is an advantage in enzymatic transformations, the great diversity in the types of glycosidic linkages also potentially means that large numbers of enzymes must be collected to cover this diversity. However, the once popular notion of "one linkage - one enzyme" might not be quite accurate, as some enzymes such as glycosyltransferases, particularly those from bacteria, can be, in fact, quite promiscuous in substrate recognition, enabling "glycorandomization" to diversify the glycosylation patterns of complex natural products [*45,46*]. Moreover, directed evolution, coupled with mechanistic studies and protein engineering, provides a powerful method to improve catalytic efficiency and plasticity of enzymes [*33,47-51*]. It is expected that studies along these lines will significantly expand the repertoire of enzymes useful for synthesis. In our opinion, the chemoenzymatic approach holds great promise, particularly in constructing those targets that are very large and contain complex carbohydrates, such as proteoglycans, glyco(sphingo)lipids, and glycoproteins.

10

The construction of homogeneous glycoproteins poses a special challenge (**Fig. 3**). One question to be asked is: Can we develop routine protocols to selectively alter the nature of glycans at specific sites in a multiply glycosylated glycoprotein, in analogy to site-directed mutagenesis in recombinant technology? Theoretically, total and semi-chemical synthesis, via native chemical ligation (NCL) and expressed protein ligation, offers a possibility to install different glycans at different sites in the context of a polypeptide (**Fig. 3a**) [*25-28*]. However, in reality, chemical technologies for total glycoprotein synthesis are far from maturation. The exploitation of various native chemical ligation strategies has significantly improved synthetic quality. Nevertheless, the choice of ligation strategy, the design of ligation components, and the efficiency of tandem ligations should all rightly be considered on a case-by-case basis for individual targets. Convergence and efficiency of ligation are a key to success and the installation of glycan prior to protein backbone assembly necessarily places a more linear emphasis on NCL routes.
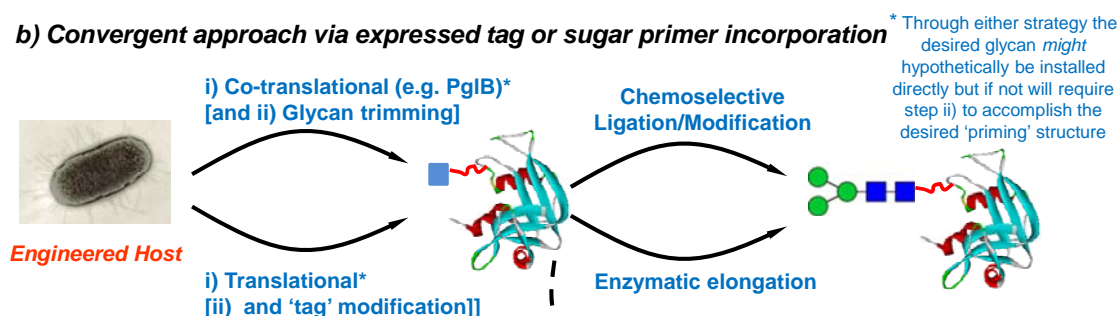
In this sense, enzymatic glycan remodeling of natural and recombinant glycoproteins (**Fig. 3c**) offers an exciting opportunity to convert heterogeneous glycoproteins to homogeneous glycoforms in a global and convergent glycosylation remodeling manner [*52*], as it combines some advantages of recombinant DNA technology for protein synthesis and the selectivity of *in vitro* enzymatic manipulation for attaching desired glycans. However, despite its power this method can struggle to control the site for convergent glycan attachment; methods are urgently needed to achieve differential site-selective glycan remodeling. One, yet to be demonstrated, solution would be novel endoglycosidases that can distinguish the nature of glycans for de-glycosylation and/or the location (site) of the monosaccharide primer needed for *en bloc* transfer of glycans. Co-translational glycosylation to determine the site selectivity provides another broad strategy (**Fig. 3b**). Although, in this regard, one of the 'killer apps' in the form of stop codon suppression by a glyco-amino acid has not yet been clearly demonstrated in a useable form [*53-55*]. The bacterial PglB oligosaccahryltransferase system has been nicely exploited *in vivo* to create valuable substrates using pathway recapitulation and repositioning of the appropriate DNXS/T consensus motif [*56*].

11

Another way to determine the site of glycosylation is to do this chemically by attaching a primer glycan through chemical site-selective protein chemistry. In this sense, the 'tag-and-modify strategy' has successfully allowed conversion of recombinant proteins into homogeneously glycosylated proteins, particularly when combined with the *in vitro* enzymatic sugar chain elongation [*57,58*]. Although these systems are tolerated by endoglycosidases to allow transglycosylation, one goal that remains to be addressed is how to introduce the tags with a fully native linkage found in natural glycoproteins.

a) Linear chemical glycoprotein synthesis

* Through either strategy the desired glycan *might* hypothetically be installed directly but if not will require step ii) to accomplish the desired 'priming' structure
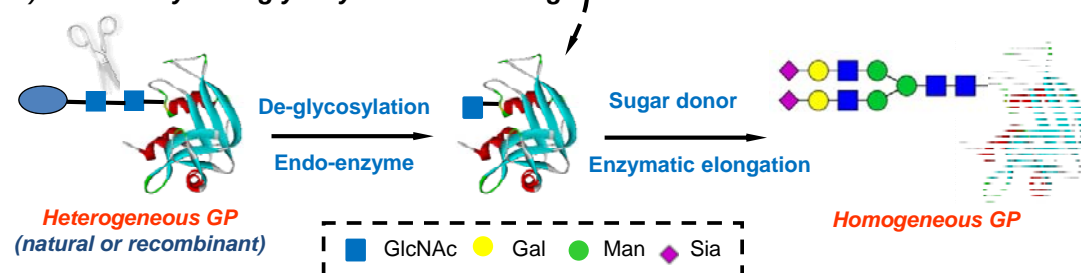
Fig. 3 Strategies for glycoprotein synthesis. a) Linear chemical glycoprotein synthesis; b) Convergent approach via expressed tag or sugar primer incorporation; c) Chemoenzymatic glycosylation remodeling

Automated and programmable synthesis

Conventional oligosaccharide synthesis often involves tedious time-consuming manipulations of a similar reaction type – protection and deprotection steps to define the regioselectivity and sometimes intensive chromatographic steps to purify the desired anomer(s) after each glycosylation. Automated synthesis is therefore a sensible and logical strategy to streamline glycan assembly by

minimizing the intermediate purification steps and by minimizing the manual nature of any repetition [59]. Major strategies that hold promise for automated and programmable oligosaccharide synthesis are summarized in **Fig. 4**. Solid-phase synthesis seems an appropriate adjunct and, indeed, di- or tri-saccharides were targets first addressed in the 1970s soon after the success of automated polymer-supported synthesis of polypeptides [60]. However, it wasn't until 2001 that an automated oligosaccharide synthesis strategy (using glycosyl phosphates on a solid support) captured general attention (**Fig. 4a**) [61]. This remarkable demonstrated success showed that an automated synthesis of complex oligosaccharides was conceptually and technologically feasible. Some have correctly highlighted that this disclosure did not provide a general solution but its power was to throw down a gauntlet to other synthetic methods – there is no doubt it has prompted a dramatic development in synthetic sophistication of both automated and non-automated methods.

Nonetheless, many exciting associated goals remain. Recent studies have started to effectively address initial drawbacks faced in this and other automated oligosaccharide syntheses, exemplified by the synthesis of increasingly more complex targets [62-64]. Nevertheless, a number of problems still remain when diverse glycan structures are concerned. In addition to the issues associated with choices of linkers and solid-support, and the accessibility of selectively protected monosaccharide building blocks, the greatest problems surround control of anomeric stereochemistry and the unpredictability of coupling efficiency in each glycosylation steps. Thus, more robust glycosylation method with perfect control of stereo-selectivity (particularly for the 1,2-cis glycosidic linkages) with excellent associated coupling yields (>99%) are welcome in this regard [30] and more are urgently needed, if compounded inefficiencies are not to continue to prove strategically unwieldy. Although more rarely considered (perhaps due to fears surrounding influence of support on biocatalytic efficiency), enzymatic transformations also hold great potential for automation; in a recent report, an automated enzymatic synthesis of sialyl Lewis X was fulfilled that applies sequential enzymatic glycosylations on a globular protein-like dendrimer as the solid support [65] albeit in a relatively low overall yield (mainly due to the low recovery efficiency of the polymer) (**Fig. 4b**). In

13

another study, a digital microfluidic technology was interestingly employed for enzymatic modifications of heparin sulfate chains immobilized on manipulable magnetic nanoparticles [66]. As more-and-more enzymes become available, biocatalytic automated platforms may allow high-throughput access of diverse naturally occurring glycans and a range of non-natural derivatives as well. Such concepts deserves further exploration in order to develop artificial systems that might, one day, effectively mimic the natural Golgi apparatus for oligosaccharide assembly and modification (interestingly both groups cited here viewed their very different technological approaches in this mimetic light).

Fig. 4 Strategies for automated oligosaccharide synthesis

14
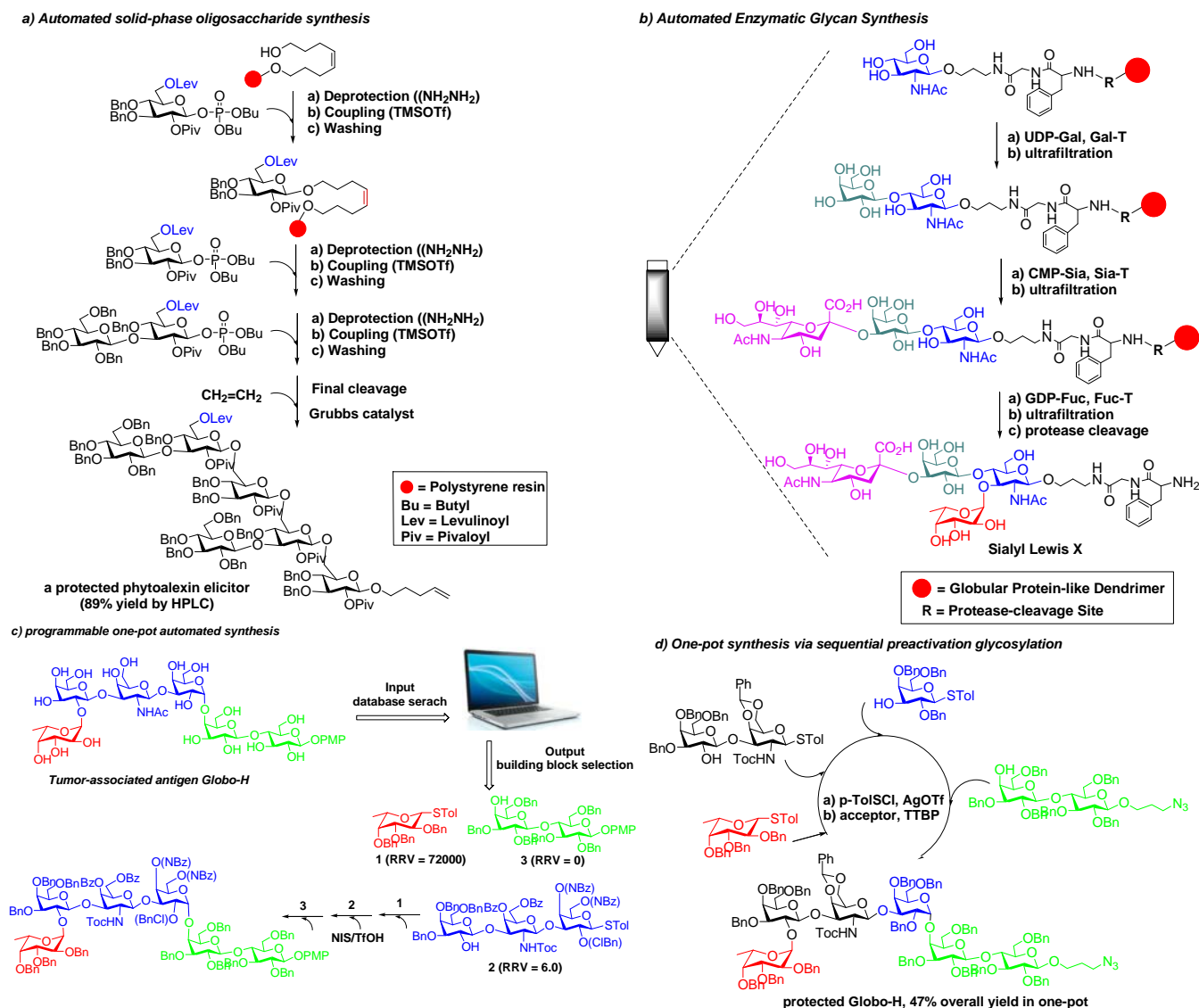
In parallel to solid-phase synthesis, there are several solution-based, one-pot, multi-step synthetic strategies that hold promise for automation. These include: programmable one-pot multi-step oligosaccharide synthesis based on relative reactivity of selected glycosyl donors (**Fig. 4c**) [*18,67*], one-pot orthogonal activation strategy [*68*], and one-pot pre-activation glycosylation strategy (**Fig. 4d**) [*69-71*]. Again, here, control is the name of the game – the success of these strategies for any given complex synthetic target relies on how one can efficiently fine tune the many factors involved, even the stoichiometry and mode-of-use of the reactants and activators, as well as their identity. As for all the strategies highlighted in this section, these methods serve to illustrate that far from there being *one* general solution, the answers will continue to also lie in robust and rapid access to rapidly optimized bespoke syntheses.

Glycan (and associated) arrays

Glycan (micro)arrays are emerging as a popular and important tool in deciphering glycan functions [*14,16,72*]. Microarray technology typically consists of immobilization on a surface (e.g., microplates / glass slides) and subsequent probing of interactions between the printed analyte and (glycan) binding partners (lectins, antibodies, enzymes, etc.) either in an isolated form or present in a biological context (sera, cell surface, or cell lysates). This format can offer several important advantages for glycomic studies, including high-throughput screening, sensitive and quantitative analysis, and the need for reduced amounts of glycans. A major application of glycan microarrays has been to detect and quantify specific interactions between glycans and diverse glycan binding proteins in mammalian, microbial and plant systems [*16*]. This has probed changes in affinity and specificity of glycan-protein interactions aimed at identifying specific biomarkers for the early diagnosis of diseases. Screening the sera from patients or vaccinated individuals in a glycan microarray format also allows quick assessment of immune responses, facilitating biomarker discovery and vaccine development [*73,74*]. In combination with sensitive mass spectrometric analysis, glycan microarray format has also

15

been successfully used for probing enzyme substrate specificity [75] and for discovering new glycosyltransferases [76].

Despite these significant advances, the technology is still far from reaching its full potential. Perhaps the most significant limitation of current glycan microarray format is a lack of sufficient number and structural diversity of glycans. For example, the most "comprehensive" glycan microarrays so far available from the Consortium of Functional Glycomics (CFG) (http://www.functionalglycomics.org) contain only 600 or so glycans. This covers only a very small fraction of basic natural glycan structures, which are estimated to be in the tens of thousands. Accelerated synthesis, including development of more efficient chemical and chemoenzymatic methods (likely employed for 'on array' synthesis), is urgently needed to fill the gap. But synthesis is not the only avenue to expanding the chemical space of glycans. In fact, a number of natural glycans can be adequately prepared directly from natural sources of glycoproteins, glycolipids, and glycosaminoglycans: over 200 natural glycans from several rich sources such as human milk, chicken ovalbumin, egg yolks, and glycosaminoglycans can be readily conjugated with a bifunctional fluorescent linker to facilitate their tagging, separation and printing in a microarray format [77,78]. Here, the fluorescent tag introduced in the glycans can also serve as a means for quantifying the immobilization density and efficiency (these particular glycan arrays were successfully used for identifying the natural ligands for galectins [78]). It should be noted that, in this pursuit, complete separation may not even be necessarily required for initial screening; once hits are narrowed to subgroups, they can be identified by further separation and characterization. Thus, development of more robust methods for isolation, characterization, and tagging of glycans from diverse natural sources should be a high priority for rapidly expanding the chemical space of glycans and as a way of accessing glycomes in a number of useful guises.
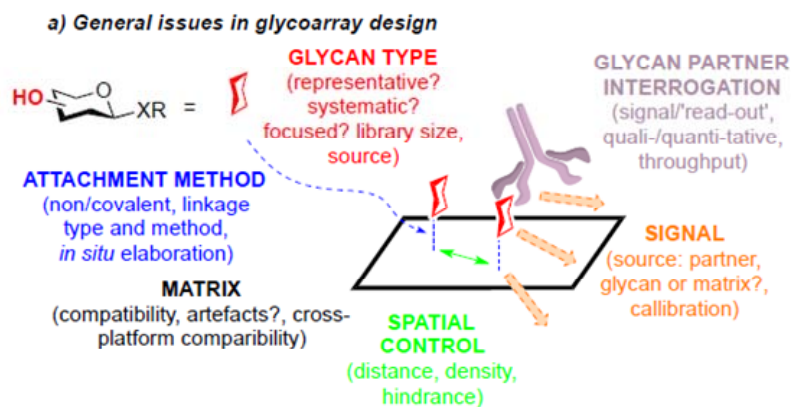
16

Fig. 5 General issues in glycan array design

Another major limitation of current glycan array technology is the lack of adequate control of glycan presentations (**Fig. 5**). Firstly, different array platforms clearly result in differences in the density and orientation of glycan presentations; these may dramatically change the affinity (avidity) and even specificity in glycan-protein interactions. Continuing cross-comparison among different array platforms is necessary in order to evaluate such factors in array fabrication [*79*]. Secondly, the presentations of glycans printed on microplates or glass slides could be quite different from those of glycans present in the context of glycoproteins and glycolipids. Synthesis of neoglycoproteins and neoglycolipids for array presentation may partially address this issue [*80-82*]. However, future generations of glycan microarrays should routinely include arrays of natural glycoproteins (varied glycoforms also) and glycolipids in order to understand the precise contributions of the lipid and protein portions to glycan-protein recognitions. This will be particularly important when micro- and macro-multivalency is involved (likely with variable mechanisms). Thirdly, future generations of glycan microarrays should also consider the dynamic nature of glycans present on cell surface. Fluidic glycan arrays offer an intriguing avenue to mimicking the mobility of glycoconjugates (rafts) imbedded in a lipid bilayer [*83*]. With efforts in the development of new approaches and tools, glycan microarray technology is expected to continue to play key roles in accelerating the decoding of the glycome. Nonetheless, we should not lose track of the vital importance of also delineating quantitative mechanistic information from glycan-protein interactions where we can – striking a balance between

rapid, qualitative and semi-quantitative methods on one hand, with a need to understand the

subtleties of modes of binding, and how these influence selectivity and kinetics (with consequences

for contextual function) on the other hand will prove vital. Moreover, whilst we perhaps naturally

gravitate to arrays that contain glycans (since they loosely mimic glycocalyx) we should also

recognize that arraying of carbohydrate-binding proteins as analytical tools for glycoconjugate solutes

can also prove powerful [*84*].

Glycan sequencing

Decoding the sequence and structure of glycans is of immediate importance in elucidating

their biological functions and associated structure-activity relationships. Whilst proteins and nucleic

acids can be routinely sequenced using well-established and robust methods, the sequencing of

glycans is more challenging. In contrast to the well-defined unmodified building blocks for proteins (20

amino acids) and DNA (4 basic nucleotides) and their single backbone linkages (amide and

phosphodiester, respectively), the possible natural monosaccharide building blocks for glycans far

exceed the number for amino acids or nucleotides. Moreover, these monosaccharide building blocks

display more diversity in their context of glycans (D- / L-, furanosyl / pyranosyl rings, regioisomerism

of linkage, linear / branching, α- or β- anomers) when forming glycosidic bonds. These structures can

be further complicated by diverse post-glycosylation modifications such as sulfation, acetylation,

deoxygenation, amination, and phosphorylation. Moreover, natural glycans in many cases appear in

the form of glycoconjugates (glycoproteins, glycolipids, proteoglycans) and are usually present as

mixtures associated also with heterogeneity of their context in those platforms.

This diversity and heterogeneity of glycans therefore presents a unique challenge in

sequencing and structural determination. Yet the level of difficulty and complexity in structural

analysis varies depending on the nature and source of the glycans, which also dictate the choice of

analytical strategies. Conventional and longstanding monosaccharide and permethylation analysis

can provide useful initial information on monosaccharide composition and site of linkages on

monosaccharide building blocks, but it has not typically been able to provide information on linkage types, linkage sequence, and anomeric configuration. MS and NMR are two most important tools for glycan sequencing and structural analysis and both technologies have now begun to take on greater challenges in structural characterization of complex oligosaccharides and glycoconjugates. For example, advances in mass spectrometry (nanoLC-ESI-MS, tandem MS, electron-capture dissociation and electron-transfer dissociation) have made it possible to reliably characterize N- and O-glycosylation of glycoproteins [85]. However, using MS it has been hitherto difficult to characterize regio- and stereo-isomers of glycans. Recent work on MS assignment of the stereochemistry and anomeric configuration of monosaccharides in oligosaccharides has shown promise in this direction [86,87]. MS coupled with specific enzymatic transformations also provides a useful tool for sequencing glycans, yet one current limitation is the availability of sufficiently diverse linkage-specific glycosidases.

Linear oligomers allow some simplification; for glycosaminoglycans (GAGs), recently, a top-down Fourier transform mass spectrometric sequencing technology was employed for sequencing bikunin, one of the simplest proteoglycans. Surprisingly, the GAG chain of proteoglycan bikunin was shown to have a "homogeneous", defined sequence for activity [88]. 'Top-down' strategy preserves whole sequences that conventional bottom-up strategy might miss. This example also emphasizes the importance of the choice of method – for samples consisting of mixtures of numerous different glycosaminoglycans, such a top-down strategy might prove tedious and difficult.

NMR has been particularly useful for polysaccharide sequencing through the application of various homo- and hetero-nuclear coupling techniques, which allows reliable assignment of the ring forms of monosaccharides, anomeric configurations, linkage types, and monosaccharide sequential connections [89]. In addition, it can also provide detailed three-dimensional structures of glycans in the context of glycoconjugates [90]. However, NMR usually requires a relatively large quantity of samples, and the sensitivity of natural isotope abundance is an issue. Thus, development of NMR techniques with greater sensitivity for polysaccharides and glycoconjugates of natural isotope

abundance is a high priority, whilst at the same time new methods that allow easy isotope labeling and enrichment of glycans with specific isotopes (^{13}C, ^{15}N, and/or ^{18}O, by chemical, enzymatic, or biosynthetic means) would prove powerful.

Novel approaches to glycan functions

For chemical biologists, function should be the striking goal. One criticism leveled at Chemical Glycobiology is that this aspect of our field is in its infancy as compared with methods for building and characterizing glycans. Key methods may offer incisive options as follows (**Fig. 6**).

a) Use of specific carbohydrate processing enzyme inibitors to modulate the glycome

b) In cellulo generation of biosynthetic inhbitors or probes from precursors

c) Glycan imaging through use of functional or chemical reporter groups (R)

d) Manipulating cell surface glycans (programming the glycocalyx)
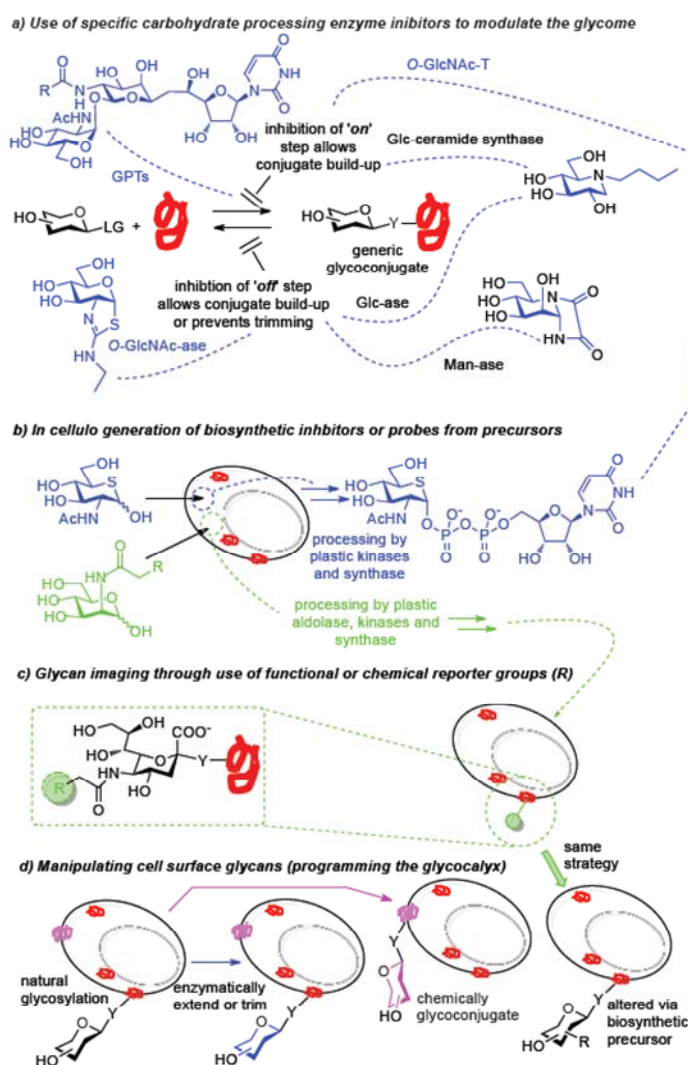
Fig. 6 Some chemical biology approaches to glycan functions

Glyco-chemical genetics. An important approach in Chemical Biology is to explore specific

exogenous (small molecule) ligands to probe gene-product functions in a cellular or organismal

context [91,92]. In this so-called 'Chemical Genetics' approach, small molecule ligands that bind

directly to proteins with high affinity and specificity are used to perturb protein functions in cellular

processes, resembling effects of genetic mutations. By analogy, novel small molecule probes that can

specifically perturb the functions of glycans should be particularly useful in functional glycomic

studies. In fact, glyco-chemical genetics is not new to our field, and many specific (natural or

synthetic) inhibitors of glycan processing enzymes, including glycosidases and glycosyltransferases,

have been discovered and implemented for manipulating functions in cells or *in vivo* [93].

Tunicamycin, a uridine analog isolated from *Streptomyces* that blocks the early steps of N-glycan

assembly thus deleting N-glycosylation globally, has played a key role in elucidating the functions of

protein N-glycosylation in protein folding, intracellular trafficking, and cellular communications.

Inhibitors against polypeptide N-acetylgalactosaminyltransferases (ppGalNAcTs), a class of enzymes

responsible for the attachment of the first sugar in mucin O-glycosylation have been discovered by

screening a synthetic uridine-based small molecule library [94]; these inhibitors were shown to

abrogate mucin-type O-linked glycosylation and to induce apoptosis in cultured cells. Recently,

potent inhibitors against the two key enzymes (OGA and OGT) responsible for a dynamic O-GlcNAc

glycosylation [95] were discovered [96,97]. As demonstrated in a mouse model of Alzheimer's

disease, direct use of a sugar thiazoline-based *O*-GlcNAcase (OGA) inhibitor led to an increase of O-

GlcNAc glycosylation level, resulting in the slowing of neurodegeneration without apparent adverse

effects [98]. Precursors/prodrugs may also prove powerful; the acetylated 5-thio analog of GlcNAc

could be taken into the biosynthetic pathway of UDP-GlcNAc in cells, leading to the *in situ* formation

of UDP-5SGlcNAc, which inhibited the OGT (O-GlcNAc transferase) activity [97] that uses UDP-

GlcNAc as a donor. Similarly, it was more recently demonstrated that selectively fluorinated sialic

acids and L-fucose analogues could be taken up and transformed into respective sugar nucleotides

that can also act as substrate-based inhibitors of enzymes that process them inside the cell [99]. Their

application resulted in global down-regulation of sialylation and fucosylation with clear consequent changes to glycome. The resulting functional consequence in myeloid cells was a loss of selectin binding and impaired leukocyte rolling [99] due to a reduced display of the associated sialylated and fucosylated glycoconjugates that act as selectin ligands.

In addition to global inhibition of glycosylation, individual glycan functions can be also dissected through the use of more specific inhibitors or molecular probes. These include the use of chemical primers serving as competing acceptor substrates for glycosyltransferases, the use of inhibitors against specific glycan-processing enzymes such as distinct α-mannosidase inhibitors, and the metabolic incorporation of sugar chain terminators to generate truncated glycoforms [11]. We know that structure-based design can lead to the discovery of potent inhibitors, as powerfully and now archetypally exemplified by the class of influenza sialidases – these were successfully developed into two widely used anti-influenza drugs, Relenza (zanamivir) and Tamiflu (oseltamivir) [100,101]. These inhibitors were designed to mimic the transition state of a sialoside during enzymatic hydrolysis, with the installation of a guanidinium or ammonium group at the C-4 position for ion pairing with the positively charged residues in the binding pocket, thus enhancing the affinity and specificity for viral sialidases over human enzymes. We are therefore aware as a field that mechanism can prove uniquely important in designs that generate potent small molecules as probes and drugs. More recently, novel mechanism-based influenza sialidase inhibitors were designed based on the structures of Relenza and Tamiflu in which fluorine atoms were introduced at the C-2 and C-3 positions [102]. These new inhibitors react specifically with influenza sialidases and inactivate the viral enzymes via the formation of a stable, covalently linked inhibitor-enzyme intermediate. Notably, these mechanism-based inhibitors demonstrate broad-spectrum antiviral activity against those drug-resistant influenza strains, making them attractive for further drug development.

As well as targeting catalytic function, synthetic glycan-based ligands are emerging as highly valuable tools to probe specific glycan-protein interactions in cellular processes. An example is the discovery of multivalent sialylated glycan ligands for CD22, a sialic acid–binding lectin from the Siglec

family that attenuates B cell receptor signaling, for suppressing B cell activation [*103,104*]. It was found that exposure of B cells to sialylated antigens led to inhibition of key steps in B cell receptor signaling and induced B cell tolerance. These studies showcase the value of the glycochemical genetic approach to dissecting cellular functions that would be otherwise difficult to unveil by traditional genetic methods. The power of chemistry therefore allows us to consider the creation of even quite large hybrid molecules in pure form to act as such perturbing probes or molecules. A so-called glycodendrinanoparticle (with a valency of up to 1620) was constructed in homogeneous form recently through the use of a nested multivalency in both glycodendrimer attachments and the self-assembly of many proteins into a virus-like particle [*105*]. The resulting constructs proved to be powerful binders of dendritic cell (DC) lectin DC-SIGN and picomolar inhibitors showing large multivalency effects in models of DC infection by Ebola virus.

The unique advantage of the glycochemical genetic approach using molecular ligands or inhibitors as probes is to perturb functions of a cellular process in a dynamic, reversible and tunable manner. This complements well the traditional genetic approaches that make "permanent" alternations of functions. However, the success of glycochemical genetic approach relies on the discovery of high-quality and specific carbohydrate-based or noncarbohydrate-based probes. Future studies should be directed not only to systematic, high-throughput screening of small-molecule compounds targeting specific sites or stages of glycan functions but also to the rational design of high-affinity ligands or inhibitors specific for glycan-recognition receptors and enzymes; and the design of novel probes that are suitable for *in vivo* (not just *in vitro*) perturbation of glycan functions.

Glycan imaging. The ability to directly detect the changes of glycans in biological systems offers an exciting opportunity to decipher their roles and functions in normal physiological and disease states. While lectins and glycan-specific antibodies have been widely used for many years for detecting cell-surface glycans and for glycan visualization on tissue sections, they are not widely suitable for all *in vivo* imaging (e.g., due to cytotoxicity and/or membrane impermeability) and the truly characterized selectivities of many 'anti-sugar' antibodies can be unclear or even absent. One

emerging technology for glycan imaging is the metabolic engineering of glycans with so-called

bioorthogonal chemical reporters [106,107]. In this approach, a specific chemical reporter is

incorporated into the glycan via biosynthetic incorporation of a sugar precursor modified with the

chemical reporter. The reporter group is then visualized by covalent reaction with an imaging probe

that can specifically detect and react with the chemical reporter. The currently most common chemical

reporters are azide or alkyne groups, since they are small and relatively biologically inert and they

have the potential to react specifically under physiological conditions. For example, an azide reporter

has been detected by reactions with a biotin-tagged phosphine via Staudinger reaction or with a

fluorescence-tagged alkyne via copper-catalyzed azide-alkyne cycloaddition reaction [108]. To avoid

perceived cytotoxicity of the copper catalyst in living system, strain-promoted azide-alkyne

cycloadditions has also been devised that use cyclooctyne derivatives as reactants, enabling reaction

without copper mediation [109-112]. Early focus was dominated by imaging of cell surface sialylated

glycans using azide- or alkyne-tagged ManNAc variants (ManNAz or ManNAl) as precursors. Since

then the metabolic glycan imaging approach has been expanded to the imaging of other types of

glycans including mucin-type O-glycans (using azide-tagged GalNAc as the precursor), O-GlcNAc

glycoproteins (using azide-tagged GlcNAc as the precursor), and fucosylated glycans (using 6-azido-

fucose or alkynyl fucose as the precursor). In addition, the imaging technology has also expanded

from its use in cultured cell systems to living organisms including zebrafish and mouse [108].

Glycan imaging provides an important tool for deciphering the localization and trafficking of

glycans in cellular processes. Moreover, noninvasive glycan imaging in living animals offers a unique

opportunity to detect, monitor, and quantify the dynamic changes of glycans associated with

development and disease states, which may unveil novel biomarkers for diagnosis and important

targets for drug discovery. However, a potential limitation of the metabolic glycan imaging is the lack

of cell type or tissue specificity, as the injected sugar precursors would be incorporated into all the

shared glycan biosynthetic pathways. A recently reported cell-specific metabolic labeling strategy that

uses ligand-targeted liposomes to encapsulate and selectively deliver azidosugar precursor to target

cells provides a promising solution to this circumvent this problem [*113*]. Future studies should be directed to the exploration of novel biosynthetic pathways to cover additional types of glycans in the glycome, and to understanding their limitations, when sugar precursors are biosynthetically used for more than one purpose [*114*]; new unnatural sugar precursors can be envisaged that possess better biocompatibility and favorable pharmacokinetic properties; and new orthogonal chemical reporters are needed that permit simultaneous imaging of different types of glycans in real time.

Selective cell-surface glycan remodeling. Diverse subtypes of glycans are present on cell surfaces. Site-specific remodeling of this glycocalyx provides an attractive approach to altering and understanding cell surface functions as well gaining novel ones. One notable example is the *ex vivo* glycan engineering of CD44 on human mesenchymal stem cells (hMSCs) that enabled redirection of cell migration to bone [*115*]. Thus, the native CD44 glycoform on MSCs that lacks the key α-1,3-fucosyl residue was converted into a hematopoietic cell E-selectin/L-selectin ligand via enzymatic α-1,3-fucosylation under mild physiological conditions; this conferred potent E-selectin binding and tropism to bone, without drastic effects on cell viability or multipotency. This and a subsequent study [*116*] unveils a great potential to program cellular trafficking via chemical engineering of glycans on a distinct membrane glycoprotein for directing cellular migration.

Another example is the development of universal red blood cells through enzymatic removal of blood group ABO antigens on the cell surface [*117*]. While this idea had been considered previously, a practical approach to remodeling the ABO antigens had not been fulfilled because of the lack of enzymes to efficiently remove the GalNAc and galactose residues present in the group A and group B antigens, respectively, without affecting red blood cells' integrity. The discovery of two specific bacterial glycosidases, an α-GalNAcase and an α-galactosidase, provided new hope; these are efficient and specific for cleavage of the A and B immunodominant monosaccharides under the mild conditions suitable for maintaining the integrity and functions of red blood cells [*118*]. This study suggests a practical approach to transforming group A, B, and AB red blood cells to the universal

group O type for transfusion. The study also emphasized the importance of discovering new, more efficient methods for glycan-specific cell-surface engineering for basic and clinical applications. Whilst many of these methods have relied to date on alterations using naturally-derived enzyme activities other approaches can be envisaged. The directed evolution of novel glycose oxidases offers tantalizing opportunities in this regard [*51*] . Catalytic chemical methods may also be considered; the Pd-catalyzed remodeling of cell surfaces through direct C–C bond formation [*119*] has recently been applied to the remodeling of bacterial glycocalyx leading to a 'chemical differentiation' that alters cellular interactions as determined by appropriate, different chemical glycosylating agents [*120*].

Conclusion

As a final general comment it is perhaps more important in this burgeoning discipline that we are very clear minded about our logic for discovery than in many other fields. Whilst an unnecessary (and often wrong) opposition is often created between "blue skies" research and applied science there is, nonetheless, a clear and useful distinction between the creation of knowledge (science) and its use (technology). This is particularly pertinent to Glycoscience since one tempting way that we might consider solving some problems that we face is through the use of 'bigger-and-better' and 'more-and-more'. In the field of arrays there is an argument that once we have a big enough array with enough sugars then we will be able to 'probe' every interaction. In the field of 'glycoproteomics' there is an argument that if we perform large enough, comprehensive 'top-down' proteomic studies then we will see sufficient examples to get a better handle on the patterns that are relevant to Glycobiology.

Both of us would like to finish by highlighting that whilst such 'powerful' resource-driven analyses will create great quantities of very useful data, the strength of many academic institutions (at least currently) is (still) best found in asking the right questions and creating the experiment to answer them. The size of the many challenges in our field means that this may end up being even more vital for our community than for others; it may be that Chemical Glycobiology is one of the first to be tested

by the temptations of such a 'post-academic' correlative approach. Sometimes we decide as a community that where we currently do not yet understand or see the direct causal links that we wish for, this might simply be because we haven't yet looked widely enough or deeply enough. This argument to 'do more' is one approach. Yet another, approach perhaps is to pause and think about new mechanisms (and hence roles for sugars) in Biology? There is the temptation for us to imbue sugars with 'meaning' that might not be there. Why does nature preserve such a diversity and complexity of glycan after so many years of evolution? Is the heterogeneity of glycoforms a requirement for functions or is it simply the results of a less perfect control of the biosynthetic endeavors? Why might an additional GlcNAc on a protein have function *cf* an additional hydroxyl? Although we know that hydroxylation and GlcNAc-ylation have specific functional effects it may also be that sometimes they provide simply a register or 'read-out' of a background organismal state. There will be a long list of questions and we have just begun to decode the function and dynamics of the glycome of any given organism. It becomes clear that chemistry, which provides unique molecular tools and approaches, will continue to play increasingly important roles in deciphering the diverse functions of glycans in biological systems.

Acknowledgements

References

(1)   T. W. Rademacher, R. B. Parekh, R. A. Dwek, *Annu. Rev. Biochem.,* **1988**, *57*, 785-838.

(2)   R. A. Dwek, *Chem. Rev.,* **1996**, *96*, 683-720.

(3)   A. Helenius, M. Aebi, *Science,* **2001**, *291*, 2364-2369.

(4)   G. W. Hart, R. J. Copeland, *Cell,* **2010**, *143*, 672-676.

(5)   J. D. Marth, P. K. Grewal, *Nat. Rev. Immunol.,* **2008**, *8*, 874-887.

(6)   R. Jefferis, *Nat. Rev. Drug Discov.,* **2009**, *8*, 226-234.

(7)   F. Crick, *Nature,* **1970**, *227*, 561-563.

(8)   Q. Chen, Y. Chen, C. Bian, R. Fujiki, X. Yu, *Nature,* **2013**, *493*, 561-564.

(9)   R. A. Laine, *Glycobiology,* **1994**, *4*, 759-767.

(10)   C. R. Bertozzi, L. L. Kiessling, *Science,* **2001**, *291*, 2357-2364.

(11)   J. A. Prescher, C. R. Bertozzi, *Cell,* **2006**, *126*, 851-854.

(12)   P. H. Seeberger, *Nat. Chem. Biol.,* **2009**, *5*, 368-372.

(13)   T. J. Boltje, T. Buskas, G. J. Boons, *Nat. Chem.,* **2009**, *1*, 611-622.

(14)   L. L. Kiessling, R. A. Splain, *Annu. Rev. Biochem.,* **2010**, *79*, 619-653.

(15)   C. Y. Wu, C. H. Wong, *Chem. Commun. (Camb),* **2011**, *47*, 6201-6207.

(16)   C. D. Rillahan, J. C. Paulson, *Annu. Rev. Biochem.,* **2011**, *80*, 797-823.

(17)   V. Verez-Bencomo, V. Fernandez-Santana, E. Hardy, M. E. Toledo, M. C. Rodriguez, L. Heynngnezz, A. Rodriguez, A. Baly, L. Herrera, M. Izquierdo, A. Villar, Y. Valdes, K. Cosme, M. L. Deler, M. Montane, E. Garcia, A. Ramos, A. Aguilar, E. Medina, G. Torano, I. Sosa, I. Hernandez, R. Martinez, A. Muzachio, A. Carmenates, L. Costa, F. Cardoso, C. Campa, M. Diaz, R. Roy, *Science,* **2004**, *305*, 522-525.

(18)   P. Sears, C. H. Wong, *Science,* **2001**, *291*, 2344-2350.

(19)   K. C. Nicolaou, H. J. Mitchell, *Angew. Chem. Int. Ed.,* **2001**, *40*, 1576-1624.

(20)   G. J. Boons, A. V. Demchenko, *Chem. Rev.,* **2000**, *100*, 4539-4566.

(21)   X. Zhu, R. R. Schmidt, *Angew. Chem. Int. Ed.,* **2009**, *48*, 1900-1934.

(22)   C. C. Wang, J. C. Lee, S. Y. Luo, S. S. Kulkarni, Y. W. Huang, C. C. Lee, K. L. Chang, S. C. Hung, *Nature,* **2007**, *446*, 896-899.

(23)   M. Petitou, C. A. van Boeckel, *Angew. Chem. Int. Ed.,* **2004**, *43*, 3118-3133.

(24)   D. P. Galonic, D. Y. Gin, *Nature,* **2007**, *446*, 1000-1007.

(25) Y. Yuan, J. Chen, Q. Wan, R. M. Wilson, S. J. Danishefsky, *Biopolymers,* **2010**, *94*, 373-384.

(26) R. J. Payne, C. H. Wong, *Chem. Commun. (Camb),* **2010**, *46*, 21-43.

(27) C. Piontek, D. Varon Silva, C. Heinlein, C. Pohner, S. Mezzato, P. Ring, A. Martin, F. X. Schmid, C. Unverzagt, *Angew. Chem. Int. Ed.,* **2009**, *48*, 1941-1945.

(28) I. Sakamoto, K. Tezuka, K. Fukae, K. Ishii, K. Taduru, M. Maeda, M. Ouchi, K. Yoshida, Y. Nambu, J. Igarashi, N. Hayashi, T. Tsuji, Y. Kajihara, *J. Am. Chem. Soc.,* **2012**, *134*, 5428-5431.

(29) M. Huang, P. Retailleau, L. Bohe, D. Crich, *J. Am. Chem. Soc.,* **2012**, *134*, 14746-14749.

(30) T. J. Boltje, J. H. Kim, J. Park, G. J. Boons, *Nat. Chem.,* **2010**, *2*, 552-557.

(31) R. M. Schmaltz, S. R. Hanson, C. H. Wong, *Chem. Rev.,* **2011**, *111*, 4259-4307.

(32) M. M. Palcic, *Curr. Opin. Chem. Biol.,* **2011**, *15*, 226-233.

(33) L. L. Lairson, B. Henrissat, G. J. Davies, S. G. Withers, *Annu. Rev. Biochem.,* **2008**, *77*, 521-555.

(34) H. Barreteau, E. Richard, S. Drouillard, E. Samain, B. Priem, *Carbohydr. Res.,* **2012**, *360*, 19-24.

(35) N. Fierfort, E. Samain, *J. Biotechnol.,* **2008**, *134*, 261-265.

(36) L. X. Wang, W. Huang, *Curr. Opin. Chem. Biol.,* **2009**, *13*, 592-600.

(37) G. Perugino, A. Trincone, M. Rossi, M. Moracci, *Trends Biotechnol.,* **2004**, *22*, 31-37.

(38) S. M. Hancock, M. D. Vaughan, S. G. Withers, *Curr. Opin. Chem. Biol.,* **2006**, *10*, 509-519.

(39) L. F. MacKenzie, Q. Wang, R. A. J. Warren, S. G. Withers, *J. Am. Chem. Soc.,* **1998**, *120*, 5583-5584.

(40) C. Malet, A. Planas, *FEBS Lett.,* **1998**, *440*, 208-212.

(41) H. Yu, X. Chen, *Org. Biomol. Chem.,* **2007**, *5*, 865-872.

(42) Y. Xu, S. Masuko, M. Takieddin, H. Xu, R. Liu, J. Jing, S. A. Mousa, R. J. Linhardt, J. Liu, *Science,* **2011**, *334*, 498-501.

(43)   J. J. Goodfellow, K. Baruah, K. Yamamoto, C. Bonomelli, B. Krishna, D. J. Harvey, M. Crispin,

C. N. Scanlan, B. G. Davis, *J. Am. Chem. Soc.,* **2012**, *134*, 8030-8033.

(44)   W. Huang, J. Giddens, S. Q. Fan, C. Toonstra, L. X. Wang, *J. Am. Chem. Soc.,* **2012**, *134*,

12308-12318.

(45)   C. Zhang, B. R. Griffith, Q. Fu, C. Albermann, X. Fu, I. K. Lee, L. Li, J. S. Thorson, *Science,*

2006, *313*, 1291-1294.

(46)   X. Fu, C. Albermann, J. Jiang, J. Liao, C. Zhang, J. S. Thorson, *Nat. Biotechnol.,* **2003**, *21*,

1467-1469.

(47)   F. A. Shaikh, S. G. Withers, *Biochem. Cell Biol.,* **2008**, *86*, 169-177.

(48)   G. J. Williams, R. D. Goff, C. Zhang, J. S. Thorson, *Chem. Biol.,* **2008**, *15*, 393-401.

(49)   G. J. Williams, C. Zhang, J. S. Thorson, *Nat. Chem. Biol.,* **2007**, *3*, 657-662.

(50)   S. M. Hancock, J. R. Rich, M. E. Caines, N. C. Strynadka, S. G. Withers, *Nat. Chem. Biol.,*

2009, *5*, 508-514.

(51)   J. B. Rannes, A. Ioannou, S. C. Willies, G. Grogan, C. Behrens, S. L. Flitsch, N. J. Turner, *J.

Am. Chem. Soc.,* **2011**, *133*, 8436-8439.

(52)   L. X. Wang, J. V. Lomino, *ACS Chem. Biol.,* **2012**, *7*, 110-122.

(53)   Z. Zhang, J. Gildersleeve, Y. Y. Yang, R. Xu, J. A. Loo, S. Uryu, C. H. Wong, P. G. Schultz,

Science, **2009**, *326*, 1187 (retraction).

(54)   Z. Zhang, J. Gildersleeve, Y. Y. Yang, R. Xu, J. A. Loo, S. Uryu, C. H. Wong, P. G. Schultz,

Science, **2004**, *303*, 371-373.

(55)   M. J. Schmidt, D. Summerer, *ChemBioChem,* **2012**, *13*, 1553-1557.

(56)   F. Schwarz, W. Huang, C. Li, B. L. Schulz, C. Lizak, A. Palumbo, S. Numao, D. Neri, M. Aebi,

L. X. Wang, *Nat. Chem. Biol.,* **2010**, *6*, 264-266.

(57)   J. M. Chalker, G. J. Bernardes, B. G. Davis, *Acc. Chem. Res.,* **2011**, *44*, 730-741.

(58)   M. Fernández-González, O. Boutureira, G. J. L. Bernardes, J. M. Chalker, M. A. Young, J. C.

Errey, B. G. Davis, *Chem. Sci.,* **2010**, *1*, 709-715.

(59)   P. H. Seeberger, *Carbohydr. Res.,* **2008**, *343*, 1889-1896.

(60)   C. Schuerch, J. M. Frechet, *J. Am. Chem. Soc.,* **1971**, *93*, 492-496.

(61)   O. J. Plante, E. R. Palmacci, P. H. Seeberger, *Science,* **2001**, *291*, 1523-1527.

(62)   P. H. Seeberger, *Chem. Soc. Rev.,* **2008**, *37*, 19-28.

(63)   F. A. Jaipuri, N. L. Pohl, *Org. Biomol. Chem.,* **2008**, *6*, 2686-2691.

(64)   C. H. Hsu, S. C. Hung, C. Y. Wu, C. H. Wong, *Angew. Chem. Int. Ed.,* **2011**, *50*, 11872-11923.

(65)   T. Matsushita, I. Nagashima, M. Fumoto, T. Ohta, K. Yamada, H. Shimizu, H. Hinou, K. Naruchi, T. Ito, H. Kondo, S. Nishimura, *J. Am. Chem. Soc.,* **2010**, *132*, 16651-16656.

(66)   J. G. Martin, M. Gupta, Y. Xu, S. Akella, J. Liu, J. S. Dordick, R. J. Linhardt, *J. Am. Chem. Soc.,* **2009**, *131*, 11041-11048.

(67)   T. K. Mong, H. K. Lee, S. G. Duron, C. H. Wong, *Proc. Natl. Acad. Sci. USA,* **2003**, *100*, 797-802.

(68)   S. Kaeothip, A. V. Demchenko, *Carbohydr. Res.,* **2011**, *346*, 1371-1388.

(69)   X. Huang, L. Huang, H. Wang, X. S. Ye, *Angew. Chem. Int. Ed.,* **2004**, *43*, 5221-5224.

(70)   B. Sun, B. Srinivasan, X. Huang, *Chem. Eur. J.,* **2008**, *14*, 7072-7081.

(71)   Z. Wang, Y. Xu, B. Yang, G. Tiruchinapally, B. Sun, R. Liu, S. Dulaney, J. Liu, X. Huang, *Chem. Eur. J.,* **2010**, *16*, 8365-8375.

(72)   O. Oyelaran, J. C. Gildersleeve, *Curr. Opin. Chem. Biol.,* **2009**, *13*, 406-413.

(73)   P. H. Liang, C. Y. Wu, W. A. Greenberg, C. H. Wong, *Curr. Opin. Chem. Biol.,* **2008**, *12*, 86-92.

(74)   O. Oyelaran, J. C. Gildersleeve, *Expert Rev. Vaccines,* **2007**, *6*, 957-969.

(75)   N. Laurent, R. Haddoub, J. Voglmeir, S. L. Flitsch, *Methods Mol. Biol.,* **2012**, *808*, 269-284.

(76)   L. Ban, N. Pettit, L. Li, A. D. Stuparu, L. Cai, W. Chen, W. Guan, W. Han, P. G. Wang, M. Mrksich, *Nat. Chem. Biol.,* **2012**, *8*, 769-773.

(77)   B. Xia, Z. S. Kawar, T. Ju, R. A. Alvarez, G. P. Sachdev, R. D. Cummings, *Nat. Methods,* **2005**, *2*, 845-850.

31

(78)   X. Song, B. Xia, S. R. Stowell, Y. Lasanajak, D. F. Smith, R. D. Cummings, *Chem. Biol.,* **2009**, *16*, 36-47.

(79)   V. Padler-Karavani, X. Song, H. Yu, N. Hurtado-Ziola, S. Huang, S. Muthana, H. A. Chokhawala, J. Cheng, A. Verhagen, M. A. Langereis, R. Kleene, M. Schachner, R. J. de Groot, Y. Lasanajak, H. Matsuda, R. Schwab, X. Chen, D. F. Smith, R. D. Cummings, A. Varki, *J. Biol. Chem.,* **2012**, *287*, 22593-22608.

(80)   O. Oyelaran, Q. Li, D. Farnsworth, J. C. Gildersleeve, *J. Proteome Res.,* **2009**, *8*, 3529-3538.

(81)   K. Godula, D. Rabuka, K. T. Nam, C. R. Bertozzi, *Angew. Chem. Int. Ed.,* **2009**, *48*, 4973-4976.

(82)   T. Feizi, W. Chai, *Nat. Rev. Mol. Cell Biol.,* **2004**, *5*, 582-588.

(83)   X. Y. Zhu, B. Holtz, Y. Wang, L. X. Wang, P. E. Orndorff, A. Guo, *J. Am. Chem. Soc.,* **2009**, *131*, 13646-13650.

(84)   D. C. Propheter, K. L. Hsu, L. K. Mahal, *Methods Mol. Biol.,* **2011**, *723*, 67-77.

(85)   S. J. North, P. G. Hitchen, S. M. Haslam, A. Dell, *Curr. Opin. Struct. Biol.,* **2009**, *19*, 498-506.

(86)   C. Konda, B. Bendiak, Y. Xia, *J. Am. Soc. Mass Spec.,* **2012**, *23*, 347-358.

(87)   T. T. Fang, B. Bendiak, *J. Am. Chem. Soc.,* **2007**, *129*, 9721-9736.

(88)   M. Ly, F. E. Leach, 3rd, T. N. Laremore, T. Toida, I. J. Amster, R. J. Linhardt, *Nat. Chem. Biol.,* **2011**, *7*, 827-833.

(89)   M. Martin-Pastor, A. Canales-Mayordomo, J. Jimenez-Barbero, *J. Biomol. NMR,* **2003**, *26*, 345-353.

(90)   A. W. Barb, J. H. Prestegard, *Nat. Chem. Biol.,* **2011**, *7*, 147-153.

(91)   B. R. Stockwell, *Nat. Rev. Genetics,* **2000**, *1*, 116-125.

(92)   S. L. Schreiber, *Nat. Chem. Biol.,* **2005**, *1*, 64-66.

(93)   T. M. Gloster, D. J. Vocadlo, *Nat. Chem. Biol.,* **2012**, *8*, 683-694.

(94)   H. C. Hang, C. Yu, K. G. Ten Hagen, E. Tian, K. A. Winans, L. A. Tabak, C. R. Bertozzi, *Chem. Biol.,* **2004**, *11*, 337-345.

(95)   G. W. Hart, C. Slawson, G. Ramirez-Correa, O. Lagerlof, *Annu. Rev. Biochem.,* **2011**, *80*, 825-858.

(96)   S. A. Yuzwa, M. S. Macauley, J. E. Heinonen, X. Shan, R. J. Dennis, Y. He, G. E. Whitworth, K. A. Stubbs, E. J. McEachern, G. J. Davies, D. J. Vocadlo, *Nat. Chem. Biol.,* **2008**, *4*, 483-490.

(97)   T. M. Gloster, W. F. Zandberg, J. E. Heinonen, D. L. Shen, L. Deng, D. J. Vocadlo, *Nat. Chem. Biol.,* **2011**, *7*, 174-181.

(98)   S. A. Yuzwa, X. Shan, M. S. Macauley, T. Clark, Y. Skorobogatko, K. Vosseller, D. J. Vocadlo, *Nat. Chem. Biol.,* **2012**, *8*, 393-399.

(99)   C. D. Rillahan, A. Antonopoulos, C. T. Lefort, R. Sonon, P. Azadi, K. Ley, A. Dell, S. M. Haslam, J. C. Paulson, *Nat. Chem. Biol.,* **2012**, *8*, 661-668.

(100) M. von Itzstein, W. Y. Wu, G. B. Kok, M. S. Pegg, J. C. Dyason, B. Jin, T. Van Phan, M. L. Smythe, H. F. White, S. W. Oliver, et al., *Nature,* **1993**, *363*, 418-423.

(101) M. von Itzstein, *Nat. Rev. Drug Discov.,* **2007**, *6*, 967-974.

(102) J. H. Kim, R. Resende, T. Wennekes, H. M. Chen, N. Bance, S. Buchini, A. G. Watts, P. Pilling, V. A. Streltsov, M. Petric, R. Liggins, S. Barrett, J. L. McKimm-Breschkin, M. Niikura, S. G. Withers, *Science,* **2013**, available online at DOI: 10.1126/science.1232552.

(103) A. H. Courtney, E. B. Puffer, J. K. Pontrello, Z. Q. Yang, L. L. Kiessling, *Proc. Natl. Acad. Sci. USA,* **2009**, *106*, 2500-2505.

(104) B. H. Duong, H. Tian, T. Ota, G. Completo, S. Han, J. L. Vela, M. Ota, M. Kubitz, N. Bovin, J. C. Paulson, D. Nemazee, *J. Exp. Med.,* **2010**, *207*, 173-187.

(105) R. Ribeiro-Viana, M. Sanchez-Navarro, J. Luczkowiak, J. R. Koeppe, R. Delgado, J. Rojo, B. G. Davis, *Nat. Commun.,* **2012**, *3*, 1303.

(106) J. A. Prescher, C. R. Bertozzi, *Nat. Chem. Biol.,* **2005**, *1*, 13-21.

(107) S. T. Laughlin, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA,* **2009**, *106*, 12-17.

(108) E. M. Sletten, C. R. Bertozzi, *Acc. Chem. Res.,* **2011**, *44*, 666-676.

(109)   P. V. Chang, J. A. Prescher, E. M. Sletten, J. M. Baskin, I. A. Miller, N. J. Agard, A. Lo, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA,* **2010**, *107*, 1821-1826.

(110)   N. J. Agard, J. A. Prescher, C. R. Bertozzi, *J. Am. Chem. Soc.,* **2004**, *126*, 15046-15047.

(111)   B. C. Sanders, F. Friscourt, P. A. Ledin, N. E. Mbua, S. Arumugam, J. Guo, T. J. Boltje, V. V. Popik, G. J. Boons, *J. Am. Chem. Soc.,* **2011**, *133*, 949-957.

(112)   A. A. Poloukhtine, N. E. Mbua, M. A. Wolfert, G. J. Boons, V. V. Popik, *J. Am. Chem. Soc.,* **2009**, *131*, 15769-15776.

(113)   R. Xie, S. Hong, L. Feng, J. Rong, X. Chen, *J. Am. Chem. Soc.,* **2012**, *134*, 9914-9917.

(114)   M. Boyce, I. S. Carrico, A. S. Ganguli, S. H. Yu, M. J. Hangauer, S. C. Hubbard, J. J. Kohler, C. R. Bertozzi, *Proc. Natl. Acad. Sci. USA,* **2011**.

(115)   R. Sackstein, J. S. Merzaban, D. W. Cain, N. M. Dagia, J. A. Spencer, C. P. Lin, R. Wohlgemuth, *Nat. Med.,* **2008**, *14*, 181-187.

(116)   S. P. Thankamony, R. Sackstein, *Proc. Natl. Acad. Sci. USA,* **2011**, *108*, 2258-2263.

(117)   G. Daniels, S. G. Withers, *Nat. Biotechnol.,* **2007**, *25*, 427-428.

(118)   Q. P. Liu, G. Sulzenbacher, H. Yuan, E. P. Bennett, G. Pietz, K. Saunders, J. Spence, E. Nudelman, S. B. Levery, T. White, J. M. Neveu, W. S. Lane, Y. Bourne, M. L. Olsson, B. Henrissat, H. Clausen, *Nat. Biotechnol.,* **2007**, *25*, 454-464.

(119)   C. D. Spicer, T. Triemer, B. G. Davis, *J. Am. Chem. Soc.,* **2012**, *134*, 800-803.

(120)   C. D. Spicer, B. G. Davis, *Chem. Commun. (Camb),* **2013**, *49*, 2747-2749.