BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

The glycosciences

Edited by Thisbe K. Lindhorst and Anja Hoffmann-Röder

Imprint

Beilstein Journal of Organic Chemistry
www.bjoc.org
ISSN 1860-5397
Email: journals-support@beilstein-institut.de

The *Beilstein Journal of Organic Chemistry* is published by the Beilstein-Institut zur Förderung der Chemischen Wissenschaften.

Beilstein-Institut zur Förderung der Chemischen Wissenschaften
Trakehner Straße 7–9
60487 Frankfurt am Main
Germany
www.beilstein-institut.de

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Silyl-protective groups influencing the reactivity and selectivity in glycosylations

Mikael Bols* and Christian Marcus Pedersen*

Review

Open Access

Address:
Department of Chemistry, University of Copenhagen, Universitetsparken 5, 2100 Copenhagen, Denmark

Email:
Mikael Bols* - bols@chem.ku.dk; Christian Marcus Pedersen* - cmp@chem.ku.dk

* Corresponding author

Keywords:
Carbohydrate; conformation; glycosylation; reactivity; selectivity

Beilstein J. Org. Chem. **2017,** *13,* 93–105.
doi:10.3762/bjoc.13.12

Received: 10 August 2016
Accepted: 23 December 2016
Published: 16 January 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Silyl groups such as TBDPS, TBDMS, TIPS or TMS are well-known and widely used alcohol protective groups in organic chemistry. Cyclic silylene protective groups are also becoming increasingly popular. In carbohydrate chemistry silyl protective groups have frequently been used primarily as an orthogonal protective group to the more commonly used acyl and benzyl protective groups. However, silyl protective groups have significantly different electronic and steric requirements than acyl and alkyl protective groups, which particularly becomes important when two or more neighboring alcohols are silyl protected. Within the last decade polysilylated glycosyl donors have been found to have unusual properties such as high (or low) reactivity or high stereoselectivity. This mini review will summarize these findings.

Introduction

Silicon-based protective groups of alcohols have a long history in organic chemistry [1-3]. The most popular and commercially available silyl-protective groups are trimethylsilyl (TMS), triethylsilyl (TES), *tert*-butyldimethylsilyl (TBS), *tert*-butyldiphenylsilyl (TBDPS), triisopropylsilyl (TIPS) as well as the diol-protective groups DTBS and TIPDS (Figure 1). Silyl groups have also early been used in the carbohydrate field to provide an alternative orthogonal protective group to the more conventional acetyl, benzoyl and benzyl groups. Particularly in oligosaccharide synthesis where many orthogonal hydroxy protective groups are required silicon protective groups have frequently been introduced in both glycosyl donors and acceptors. However, glycosylation with heavily silylated carbohydrate derivatives is comparatively new, and so is the significance that silyl groups have on the stereoselectivity and reactivity in glycosylation reactions [4]. These findings, which most have occurred in the last decade, will be reviewed here.

Review

One of the earliest glycosylations with a persilylated glycosyl donor was carried out by Kihlberg and Broddefalk who needed an acid-labile protective group [5]. They protected a thiocresyl

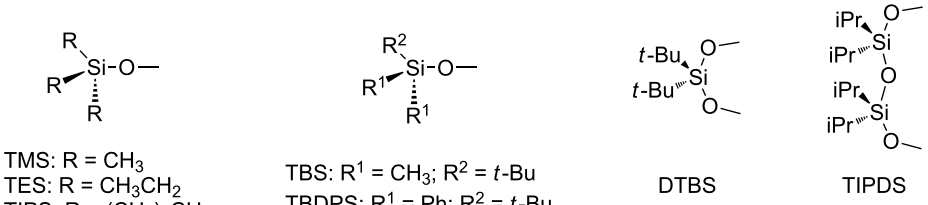
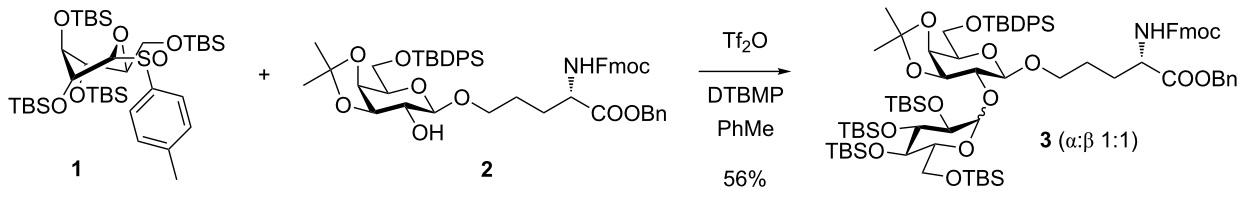


Figure 1: Silicon-protective groups typically used in carbohydrate chemistry.

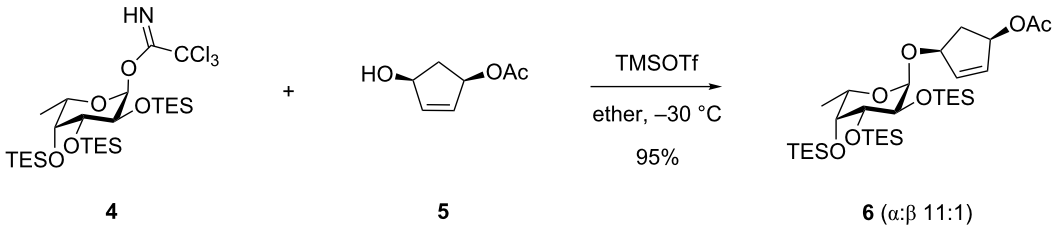
glucoside with TBS groups, oxidized the sulfur to sulfoxide **1** and used the latter to glucosylate the 2-OH of the galactose derivative **2** (Scheme 1). The reaction gave a 56% yield of **3** as a 1:1 mixture of α- and β-glucosides. Migration of a TBS group to the acceptor alcohol **2** was observed as a byproduct (10%). Attempts of glycosylating **2** with the thioglycoside or the corresponding glycosyl halides were unsuccessful. NMR studies of **1** revealed that the compound adopted a skew-boat conformation, based on the small ^3J coupling constants, as well as long range w-couplings. This conformational flip is induced by the presence of the bulky *trans*-vicinal silyl groups [6].

Also with the purpose of having acid-labile protective groups on the donor a TES-protected trichloroacetimidate of fucose, **4**, was employed by Myers et al. [7] in order to have protective groups compatible with their synthesis of neocarzinostatin. It was found that optimal glycosylation was performed with TMSOTf as a catalyst at low temperature and excess donor in diethyl ether since this gave the best α-selectivity (Scheme 2). Using other protective groups on the fucose part, such as 2,3-TIPDS and 4-*O*-TES led to glycosylation with only poor stereoselectivity [8]. The TES groups were also used successfully on the 2-methylamino analogue of **4**.

A glycosylation with a TES-protected glycosyl donor has also been performed in a case where the target contained a 6-*O*-acylglucoside and hence protective groups that could be removed under mild acidic conditions were needed [9]. This was for example used for the synthesis of the serine protease inhibitor banyaside. TES-protected glycosylimidates were also employed in the synthesis of antitumor saponins which contained partially acylated oligosaccharides. The TES groups could be removed by comparatively mild treatment with fluoride without hydrolysis or migration of *O*-acyl groups [10]. This strategy has also been applied to prepare partially acylated cholestan glycosides. In this case an imidate with a 2-*O*-acetate and 3,4-*O*-TES protection was used, which ensured stereoselectivity by neighboring-group participation [11]. For similar reasons the per-TES-protected thioglycoside **7** was employed to prepare the Lewis X trisaccharide: The reaction of **7** with disaccharide **8**



Scheme 1: Glycosylation with sulfoxide **1**.



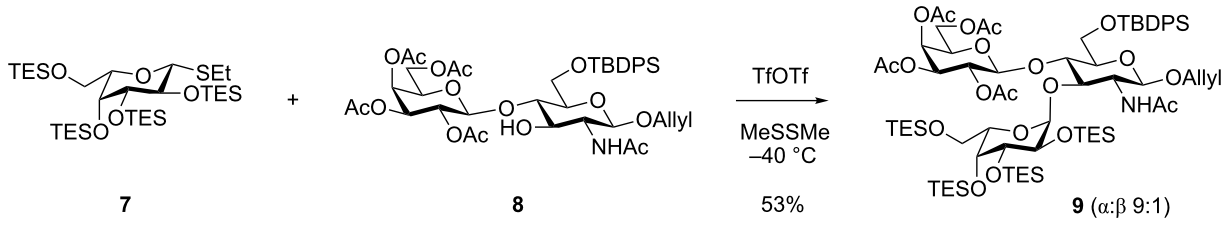
Scheme 2: Glycosylation with imidate **4**.

promoted by dimethyl disulfide and triflic anhydride gave trisaccharide **9** with high α-selectivity (Scheme 3) [12]. These conditions, using this promoter system, worked fine in a number of similar cases.

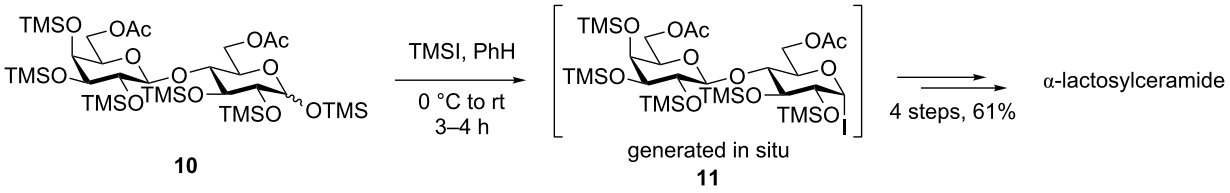
The less-stable trimethylsilyl group has been employed by Gervay–Hague and co-workers to protect glycosyl donors [13-17]. The reaction of a hexa-TMS-protected lactose derivative **10** with TMS iodide converted it to glycosyl iodide **11** that glycosylated alcohols in good yields (Scheme 4). The TMS protective groups are however rather unstable and they were exchanged to acetyl groups after the glycosylation step [13]. Nevertheless, the TMS-protected glycosyl iodides were useful intermediates because they were more reactive and less prone to elimination than the corresponding benzylated or acetylated glycosyl iodides.

Effect of silyl protective groups on the reactivity

Protective groups can profoundly influence the reactivity of carbohydrate derivatives and especially glycosyl donors [18]. This influence is due to the different electron-withdrawing capability of protective groups. During the glycosylation reaction the anomeric carbon becomes increasingly electron poor, with the formation of a glycosyl cation as the extreme. This development of a (partial) positive charge is less favorable with more EWD protective groups and the reaction becomes slower; i.e., the donor is less reactive (disarmed) [19]. Ester protective groups such as acetyl and benzoyl are among the most electron-withdrawing of the common protective groups, whereas benzyl (or methyl) groups are less so, which is reflected in the reactivity of glycosyl donors carrying these groups. As shown in Figure 2, the thioglycoside with benzyl ethers **13** is about



Scheme 3: Glycosylation with thioglycoside **7**.



Scheme 4: In situ formation of a silylated lactosyl iodide for the synthesis of α-lactosylceramide.

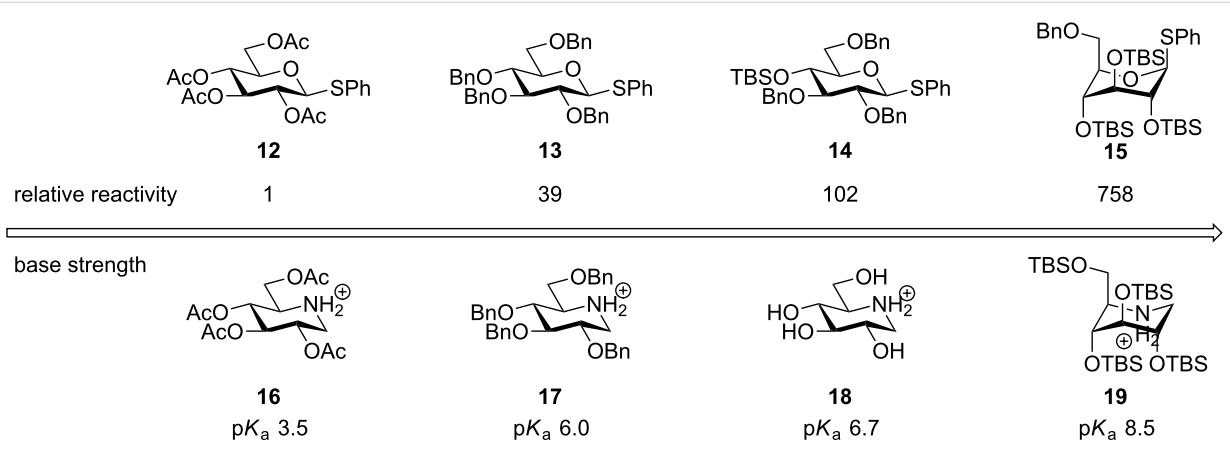


Figure 2: Comparison of the reactivity of glycosyl donors with the pK_a of the corresponding piperidinium ions.

40 times more reactive towards glycosylation with methanol upon activation by NIS, than the acetylated counterpart **12**, but the thioglycosides with silyl ethers are even more reactive [20].

Thus the presence of a single *O*-TBS group (**14**) can more than double the reactivity while three (**15**) will increase the rate by 20 times as compared to benzyl. The increased reactivity of the silylated glycosyl donors is partially due to the *O*-silyl group being somewhat less electron withdrawing than the benzyl, but also due to the ability of bulky silyl groups to cause a change in the sugar ring conformation [21]. The influences of the various protective groups are also clearly reflected in their ability to alter the base strength of the transition state mimicking amine deoxynojirimycin (Figure 2) [22]. The acetylated amine **16** is vastly less basic than the benzylated analogue **17**, which is still less basic than the unprotected amine **18** which in many ways should be similar to an O-silylated compound **19** since the silyl group inductively is very comparable to the proton. Yet the silylated amine **19** is almost a 100 times more basic due to the conformational ring flipping induced by the bulky silyl groups. This extraordinary effect on the basicity and the donor reactivity stems from the conformational change in the sugar ring, which causes the OR groups in the 3 and 4 and occasionally the 2-position to adopt an (pseudo)axial orientation, which is less electron withdrawing [23]. This conformational change is induced when having *trans*-vicinal OR groups (Figure 3). Normally the bisequatorial orientation is preferable due to 1,3-diaxial interactions of axial substituents. This steric interaction can however be overridden when the R groups are sufficient bulky and hence the sugar ring changes the conformation. The electronegativity of the R group is probably also important;

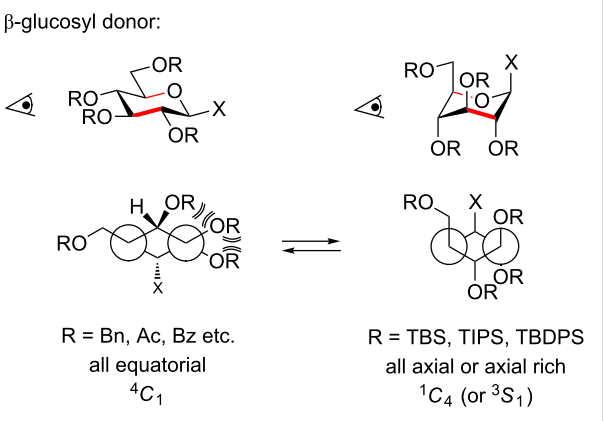


Figure 3: Conformational change induced by bulky vicinal protective groups such as TBS, TIPS and TBDPS. The vicinal clash overrules the 1,3-diaxial interaction, which is less influenced by bulky silyl ethers as these can rotate more freely in the axial-rich conformation. The projections are along the red bonds in the two models.

when more electropositive (as Si), the oxygen atoms become more electron rich and their repulsion becomes larger.

Changing the conformation of a heterocycle has, as mentioned, been studied using the piperidine model system. The pK_a of the corresponding piperidinium ion is a measure of the stereoelectronic effects and correlates with the glycosyl donor's reactivity observed. Forcing an OR group from an equatorial position into an axial position by, e.g., a bulky silyl group, increases the basicity of the piperidines, which is analogous to increasing the reactivity of the corresponding glycosyl donors.

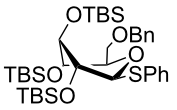
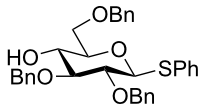
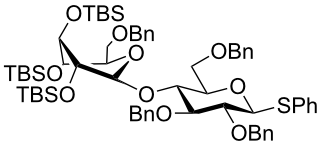
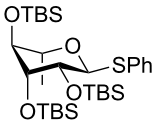
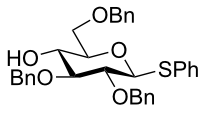
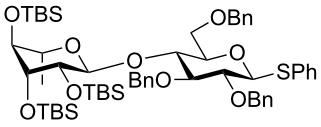
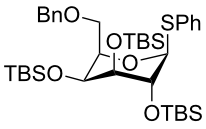
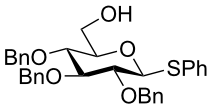
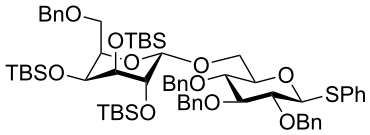
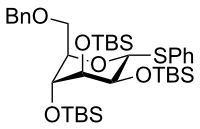
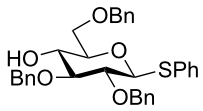
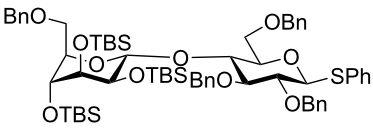
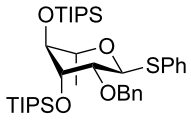
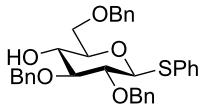
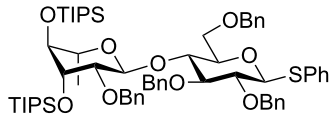
The increased reactivity is very clearly displayed when TBS or TIPS-protected thioglycosyl donors are mixed with benzylated thioglycoside acceptors under activating conditions (Table 1). The benzylated thioglycosides **21** and **26**, normally termed 'armed' due to their comparatively high reactivity, were selectively glycosylated by silylated thioglycosides (**20**, **23**, **25**, **28** and **30**) in high yield without any self-glycosylation of the armed donors [24,25]. Based on their extraordinary reactivity these silylated donors were termed 'superarmed'. The listed reactions (Table 1) were all highly stereoselective as well. The stereoselectivity is very dependent on the bulkiness of the protective group on C2 in the mannosyl (**28**), rhamnosyl (**23** and **30**) and glucosyl donors (**20**) (see also Scheme 11). In these systems the *trans* products are favored. In the galactosyl donor **25** the bulky C4 substituent shields the β-face of the donor and hence the glycosylation is very α-selective.

The remarkable difference in reactivity between disarmed, armed and superarmed donors **20**, **26** and **32**, respectively was used for "one-pot one addition" glycosylations having all 3 donors present together with all reagents from the start (Scheme 5). The activation of the individual donors was controlled by changing the temperature and the trisaccharide donor **33** could thereby be prepared in excellent yields [21].

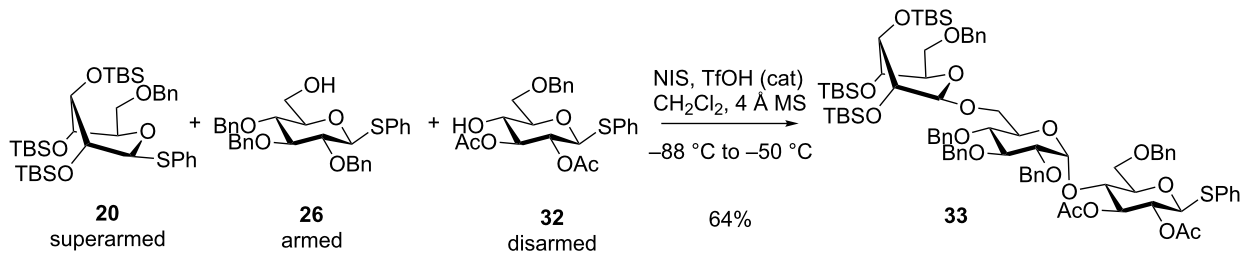
The reactivity of silylated donors have also been investigated by Hung, Wong and collaborators [26]. Investigating benzylated thioglucosides with a single or two TBS or TIPS groups in different positions they observed an increasing rate that were qualitatively similar to those described in Figure 2. Rate increases were however larger and TIPS protection had a greater rate-increasing effect than TBS protection.

The rate increases caused by a single silyl group in the 2,3 and 4-position are particular remarkable given that no obvious conformational change in the ground state is observed. Thus the increased rate must be caused by the group's ability to favor conformational inversion to the more reactive axial conformation in the transition state. This explains the comparatively large rate

Table 1: Reaction of silylated thioglycosides with benzylated thioglycoside acceptors.

Silylated donor	Benzylated donor	Product[a] (yield %)
 **20**	 **21**	 **22** (85%)
 **23**	 **21**	 **24** (90%)
 **25**	 **26**	 **27** (70%)
 **28**	 **21**	 **29** (81%)
 **30**	 **21**	 **31** (66%)

[a]Only the shown stereoisomer was obtained. Data taken from [24,25].

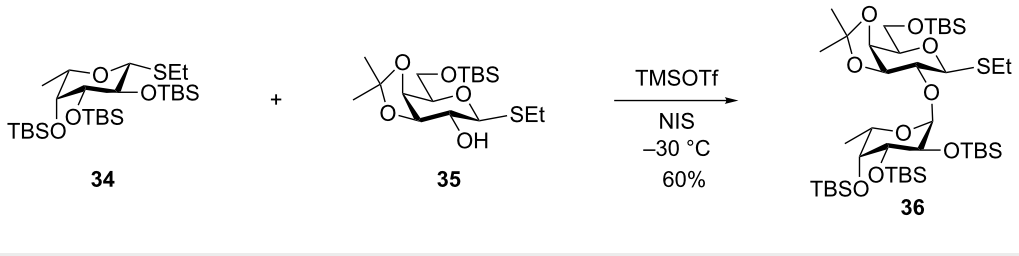


Scheme 5: An example of a "one pot one addition" glycosylation, where 3 glucosyl donors are mixed with 2.1 equiv NIS and a catalytic amount of TfOH. The individual donors are activated at different temperatures due to their reactivity and the trisaccharide donor is formed in an excellent yield.

enhancements observed by TBS and TIPS groups compared to unprotected OH and also that TIPS, which is more bulky than TBS, but essentially has the same inductive effect, causes a greater rate enhancement.

Gervay–Hague has reported that TMS-protected glycosyl iodides are remarkably more reactive than their benzyl-protected analogues [13,27]. While this rate enhancement is at least partially stemming from the change in the inductive effect, it is also possible that the comparatively more bulky TMS groups also cause an enhancing effect by favoring conformational inversion to the stereoelectronically more stable conformer in the transition state.

The reactivity of TBS-protected thioglycosides was further investigated by Scanlan and co-workers who made the fucosyl donor **34** (Scheme 6) [28]. Interestingly the NMR spectrum of this compound displayed line broadening indicating some con-



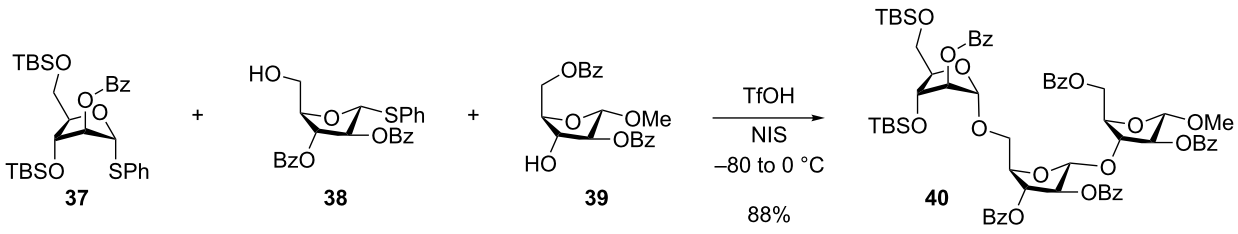
Scheme 6: Superarmed-armed glycosylation with thioglycoside **34**.

formational inversion, but the X-ray structure of the crystalline compound was in the conventional $^1C_4$ conformation. Yet the compound was clearly very reactive as it selectively could glycosylate the 2-OH of thioglycoside **35** giving **36** in a very good yield. Other acceptor alcohols were also glycosylated in a good yield and with high α-selectivity [28].

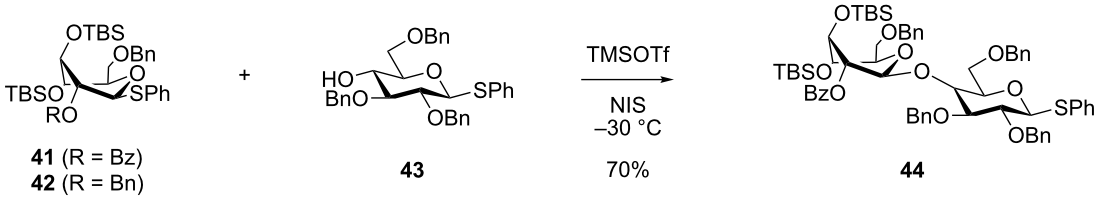
Yang and co-workers have extended the concept to the furanoside series [29]. They showed that the arabinofuranosyl donor **37** and its 2-*O*-TBS analogue were more reactive than the corresponding benzylated thioglycosides in competition reactions and used the reactivity differences in a one-pot glycosylation reaction between **37**, a disarmed donor/acceptor **38** and an acceptor **39**, which gave the trisaccharide **40** in a remarkable yield of 88% (Scheme 7). This reaction works so well because the more readily activated donor **37** reacts with the more reactive and accessible primary alcohol of **38** rather than with the secondary hydroxy group in **39**.

In the above study it was found that **37** was less reactive than the persilylated analogue [29], which was not obvious as the Demchenko group [30-35] has shown that a 2-*O*-ester can have an activating effect by the aid of anchimeric assistance [36]. This combination of conformational arming and anchimeric assistance was investigated by Heuckendorff et al., who studied the 2-*O*-benzoylated analogue of **20**, **41** (Scheme 8) [37]. They observed that though **41** was less reactive than the 2-*O*-benzyl derivative **42** it was nevertheless more reactive than the conventionally armed donor and could smoothly be coupled on the 4-OH group of the armed thioglycoside **43** without competing self-condensation of **43**.

The Yang group has also investigated superarmed galactothiofuranosides [38]. In line with the findings described above they found that the donor reactivity increased with the number of TBS protective groups in the molecule. However, the 3,5-di-*O*-TBS-2,6-di-*O*-benzoyl derivative was sufficiently reactive to



Scheme 7: One-pot double glycosylation with the conformationally armed thioglycoside **37**.



Scheme 8: Superarmed-armed glycosylation with thioglycoside **41**.

glycosylate partially benzoylated thioglycosides with high chemoselectivity and was therefore used in a range of high yielding oligosaccharide syntheses [38].

The bifunctional silicon protective group DTBS (Figure 1) has been used both to increase and decrease the reactivity of glycosyl donors. The 4,6-*O*-DTBS-protected thioglucoside **45** was found to be much less reactive than **20** and only couples to armed donor/acceptors in low yield (Figure 4) [24]. This is analogous to the effect of the very similar benzylidene group, which is deactivating the donor partially due to locking the structure in an unreactive conformation and due to the electronic effect of a *trans*-gauche conformation of the hydroxymethyl group [22,39].

Yang and collaborators found that **46** was less reactive than the fully benzoylated analogue, which is obviously also due to the DTBS group locking the molecule into an unreactive conformation [29]. In line with this, the analogue of **46** having a TIPDS group rather than a DTBS was not particularly unreactive, as it is more flexible due to the bigger ring. The Yang group used **46** in a one-pot synthesis of a trisaccharide, where they took advantage of **46** being less reactive than partially benzoylated arabinofuranosides [29]. The concept was extended to the galactofuranosyl series, but was less useful there [38]. A slightly lower reactivity of **47** was found relative to the fully benzoylated species.

DTBS groups can also be used to increase the reactivity of glycosyl donors [40]. A series of differently configured monosaccharide thioglycosides were subjected to linking the 3 and 6-OH group together with this silyl ether. This forces the glycosyl-donor conformation to change into an axial-rich conformation and hence into a superarmed donor (Table 2) making it possible to glycosylate an armed glycosyl donor selectively. This approach works for glucosides, mannosides, and galactosides and both, α- and β-thioglycosides [40]. It was shown by competition experiments that these tethered donors were even

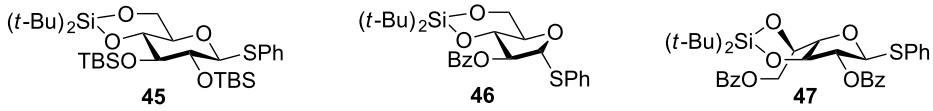
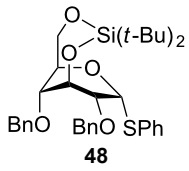
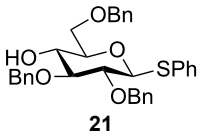
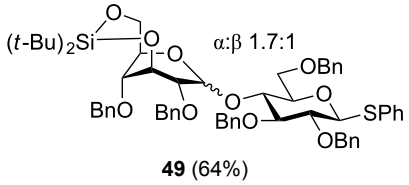
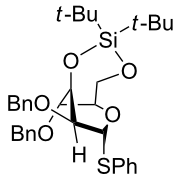
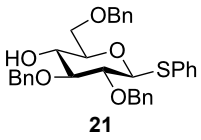
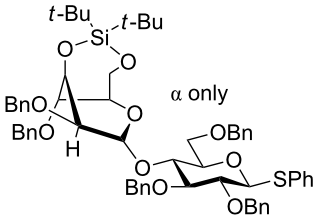
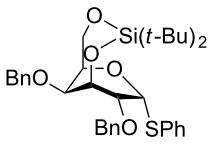
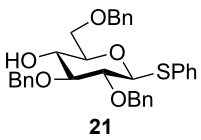
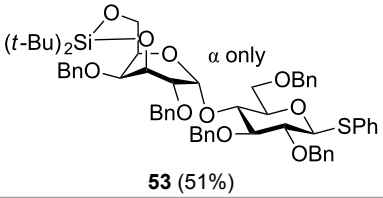


Figure 4: Donors disarmed by the di-*tert*-butylsilylene protective group.

Table 2: Reactions of 3,6-*O*-silyl-tethered thioglycosides.

Silylated donor	Benzylated donor	Product[a] (yield %)
 **48**	 **21**	 α:β 1.7:1 **49** (64%)
 **50**	 **21**	 α only **51** (70%)
 **52**	 **21**	 α only **53** (51%)

[a]Only the shown stereoisomer was obtained. Data taken from [40].

more reactive than the TBS-protected donors such as **20**. This was particularly the case for the α-anomers as a considerable reactivity difference between α- and β-thioglucosides was observed with the α-anomer consistently being more reactive. This suggested that the exact alignment of the leaving group is important for the reactivity, but a similar difference was not observed for other superarmed glycosyl donors (Figure 5).

Surprisingly, a 2,4-*O*-tethering of a glucosyl donor, giving the all axial conformation, did not increase the reactivity and the donor was found not to be superarmed. The explanation for this relates to the more strained conformation which counteracts a flattening of the conformation when approaching an sp^2-hybridized C1 in the TS [41].

Effect of silyl protective groups on the selectivity

The bulkiness of TBS groups in donors such as **20** can have a significant influence on the diastereoselectivity. Thus glycosylations with **20** (Table 1) gave exclusively the β-glucoside presumably due to steric hindrance for attack from the α-side [24,42]. The bulkiness of **20** was clearly seen in regioselective glycosylations performed by Felice et al. [43]. So, the glycosylation of the D-*allo*-configured acceptor **54** with **20** not only gave exclusively the β-glucoside, but resulted also in the glycosylation exclusively at the equatorial 4-OH group presumably due to the bulkiness of the silylated donor. Thus compound **55** (Scheme 9) was formed as the only product out of four possible isomers in 54% yield. When the D-*gluco*-configured acceptor analogue of **54** was used, a mixture of regioisomers was obtained.

However, the ability of the bulky silyl groups to alter the conformation of the glycosyl-donor ring can be used to control the selectivity. Suzuki and collaborators showed that the *C*-arylation reactions with the 3,4-*O*-di(*tert*-butyldiphenylsilyl)-protected acetate **56** led to the α-glycoside **58** with high selectivity (Scheme 10). The reason for this selectivity is that the equatorial position is more accessible for attack [44]. However, if different protective groups and even the related TBS group were

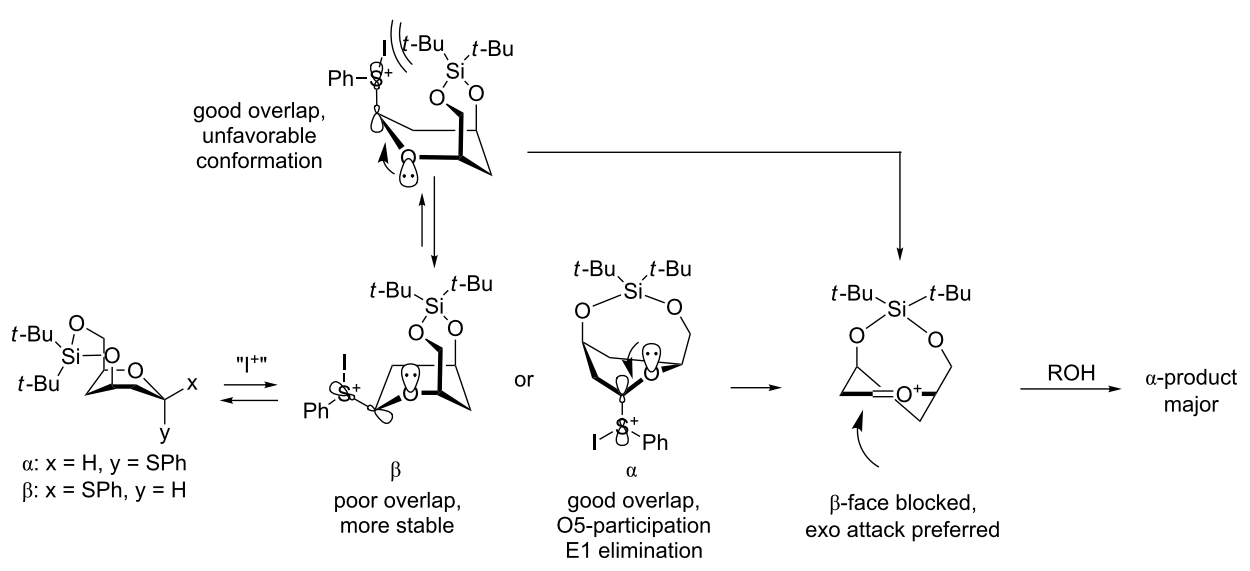
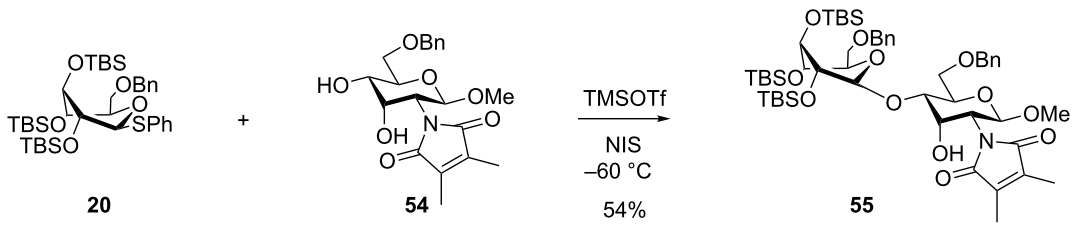
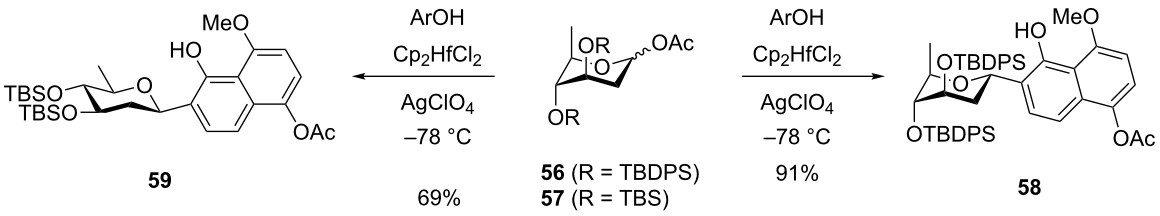


Figure 5: The influence of a 3,6-*O*-tethering on anomeric reactivity and glycosylation selectivity. The α-thioglycoside is more reactive as a conformational change is not needed to expel the sulfonium ion. This is not the case with the β-anomer. Selectivity is mainly controlled by sterics and hence the α-glycoside is kinetic product as the alcohol approach the oxocarbenium ion intermediate from the exo-side.



Scheme 9: Regio- and stereoselective glycosylation using the superarmed thioglycoside donor **20**.



Scheme 10: Superarmed donors used for *C*-arylation and the dependence of the size of the silylethers on the stereochemical outcome.

used, predominantly the β-glycoside **59** was obtained in a 14:1 α:β ratio.

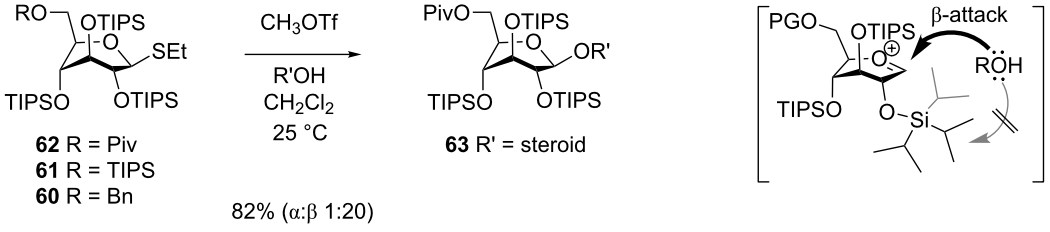
A similar conformation-controlled stereoselectivity has been demonstrated in radical reactions, however, with the twist that stereoselectivity here is opposite. The reduction of the seleno-glycosides (analogues to **20**) with tributyltin deuteride gave predominantly deuterium in the β-position for silylated derivatives in the $^1C_4$ conformation, because the reaction intermediate is a radical that prefers to be axial. On the other hand, with acetate protective groups, the addition of deuterium occurred predominantly from the β-side [45,46]. The principle of conformational stereocontrol was also used for the stereoselective addition of carbon radicals [46,47].

This selectivity has also been demonstrated for electrophilic additions to the anomeric position. Shuto and collaborators showed that, while 2,3,4-tri-*O*-benzylxylopyranosyl fluoride reacted with allyltrimethylsilane and BF_3 to give a mixture of α- and β-1-*C*-allyl xylosides, the 2,3,4-*O*-TBS-protected fluoride which is in $^1C_4$ conformation, exclusively gave the β-xyloside. In contrast the xylosyl fluoride with a butane-2,3-bisacetal protective group, that keeps the conformation fixed in a $^4C_1$ conformation, only gave the α-xyloside [48]. This sort of behavior fits well with the reaction model proposed by Woerpel for these types of reactions [49].

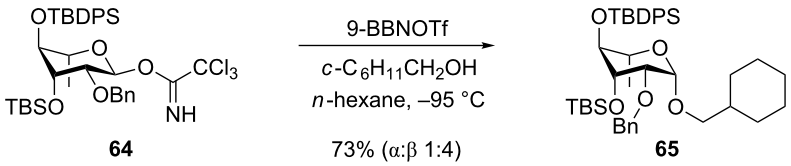
Yamada and collaborators were the first to show that this principle could be used for the stereoselective synthesis of *O*-glycosides [42]. They prepared thioglucosides **60**–**62** (Scheme 11) having 2,3,4-*O*-TIPS groups and either TIPS, benzyl or pivaloyl protective groups on the 6 position. These glucosyl donors were found to adopt the $^3S_1$ conformation and when they were reacted with methyl triflate and a glycosyl acceptor at room temperature they gave the β-glucosides in 45–92% yield and with 6:1 or better selectivity. The 6-*O* pivaloyl derivative **62** gave the best stereoselectivity (Scheme 11). The technique was later used in the synthesis of the natural product davidiin [50]. The 6-*O*-(3,5-diacetoxy-4-methoxy)benzoyl analogue of **62** was reacted with 3,5-diacetoxy-4-methoxybenzoic acid in the presence of methyl triflate, which gave the β-ester in 83% yield, showing that the principle works for ester synthesis, too.

The Yamada group also attempted to synthetize β-rhamnosides using this principle of conformational inversion [51]. The 3-*O*-TBS-4-*O*-TBDPS-protected trichloroacetimidate **64** was investigated and could give β-selectivity up to 4:1 (Scheme 12). The corresponding thioglycoside donor gave an almost fifty–fifty selectivity. Experiments performed with the 3,4-*O*-TIPS-protected thiorhamnoside donors (Table 1) were not more successful as the activation of this donor with NIS/TfOH also gave mixtures and often predominantly the α-rhamnoside [52]. This, together with the results with the α-selective TBS-protected mannosyl and galactosyl donors (Table 1) [24], shows that there is no general trend with respect to the selectivity of these donors.

On the other hand, the configurationally inverted fully TBS-protected phenyl thiorhamnoside was found to be highly α-selective (Table 1) presumably due to steric hindrance from



Scheme 11: β-Selective glucosylation with TIPS-protected glucosyl donors. The α-face is shielded by the bulky 2-*O*-TIPS protective group.



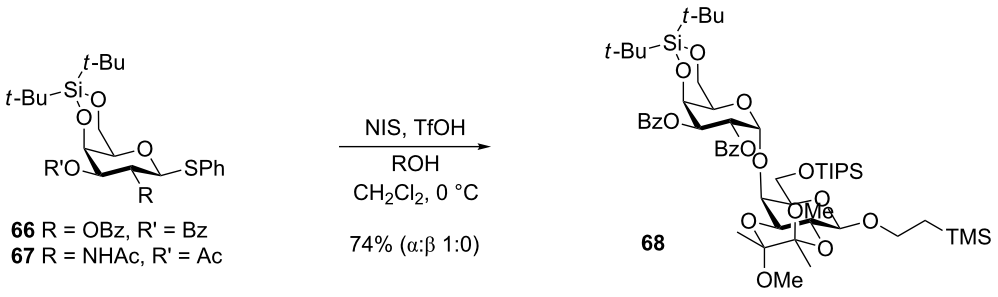
Scheme 12: β-Selective rhamnosylation with a conformationally inverted donor.

the 2-*O*-TBS on the β-face. This donor was recently used in the synthesis of glycosyltransferase acceptor substrates [53]. Yet the 2-*O*-TBS protection does not always have this effect. In a recent paper it was shown that in a 4,6-*O*-benzylidene-protected thioglucoside donor, which has been shown by Crich to be α-selective, the α-selectivity increased even more when a 2-*O*-benzyl was exchanged with 2-*O*-TBS or 2-*O*-TIPS [54]. The authors suggested that the silyl group had an inductive effect that favored α-formation.

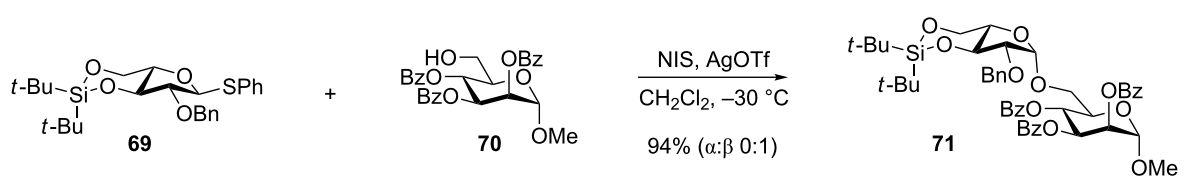
The 4,6-*O*-DTBS group has been shown to be an α-directing group in galactosylation reactions. Kiso and co-workers found that the galactosyl donor **66** (Scheme 13) reacted with several different acceptor alcohols giving exclusively the α-galactoside despite having a potentially β-directing benzoate group in the 2-position [55]. Thus the glycoside **68** was obtained in 74% yield as the only isolated product (Scheme 13). Equally remarkable is that the corresponding DTBS-protected galactosamine donors (such as **67**) displayed the same selectivity in the presence of the silyl group and thereby overriding the influence of a 2-phthalimido, *N*-Troc or *N*-Ac group. It was suggested that the bulky DTBS group is shielding the β-face and thereby preventing attack from that face of the oxocarbenium ion. This methodology has been applied to the synthesis of glycolipids and was shown to also work with 2-*O*-benzyl [56] or 2-*O*-TBS and with *N*-phenyltrifluoroacetimidate as the leaving group [57,58]. A somewhat similar influence has been observed with the much less steric demanding 4,6-*O*-benzylidene protective group [59].

A related stereoselectivity is induced by the DTBS group in arabinofuranosylations. Boons and collaborators found that the 3,5-DTBS-protected L-arabinosyl donor **69** upon reaction with acceptor **70** and activation with NIS/silver triflate gave exclusively the β-glycoside in a yield of 94% (Scheme 14).

Similarly, the reaction of **70** with the corresponding perbenzylated donor only gave a 2:1 β:α-ratio of **71** [60]. It was proposed that the selectivity was caused by a favored β-attack on the oxocarbenium ion in an E_3 conformation as the corresponding α-attack would lead to an unfavorable eclipsed conformation. The exchange of the 2-*O*-benzyl with a 2-*O*-TIPS leads to some



Scheme 13: α-Selective galactosylation with DTBS-protected galactosyl donors.

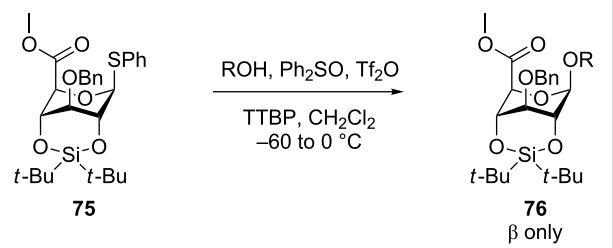


Scheme 14: β-Selective arabinofuranosylation with a DTBS-protected donor.

erosion of stereoselectivity though the donor was still highly β-selective [61]. Independently of the work by the Boons group, Crich et al. showed that using preactivation conditions on the equivalent D-arabinofuranosyl donor resulted in rupture of the β-selectivity [62]. Ito and co-workers studied the influence of tethering the 3- and 5-OH by a 3,5-*O*-(tetraisopropyldisilylene)acetal and also found the arabinofuranolysations to be β-selective, despite the more flexible system [63]. Interestingly it was recently found that exchanging the 3,5-DTBS group with trifluoroacetates retained a high β-selectivity, which suggests that the stereoselectivity is also related to the deactivating properties of the protective group [64].

Cyclic silyl protective groups were also recently found to have a beneficial influence on the α-selectivity obtained in glycosylations using glucals [65]. The reaction of 3,4-*O*-TIPDS-protected glucal **72** with acceptor alcohols such as **73**, catalyzed by *p*-TsOH, gave exclusively α-glucoside **74** (Scheme 15). When the same glycosylation was performed with the fully benzylated or TBS-protected glucal the reaction gave a lower yield and was accompanied by some formation of the β-anomer and some Ferrier rearrangement product. With donor **72** the reaction was however high yielding and exclusively α-selective for a range of alcohols. Surprisingly the 6-deoxy version of **72** gave a lower α-selectivity. The observations were explained with the assistance of DFT calculations as being due to the TS structure (formed from **72**) being in an α-selective $^4H_3$ conformation with the 6-TIPS group in an electronically favored gauche–gauche conformation [66], that causes additional shielding from the β-face [65].

The influence of having a 2,4-*O*-di-*tert*-butylsilylene (DTBS) in a glucosyl donor was, as earlier mentioned, not increasing the reactivity of the donor, but it influences the selectivity in the glycosylation. The α-site of the donor becomes the *endo* face, which results in an attack from the β-site. In a conventional glucosyl donor this leads to a 1:10 β-selectivity [41]. Recently this behavior has been used by Furukawa et al. in a β-controlled glucuronylation, where the bulky silylene in **75** ensures high selectivity without neighboring group participation (Scheme 16) [67].



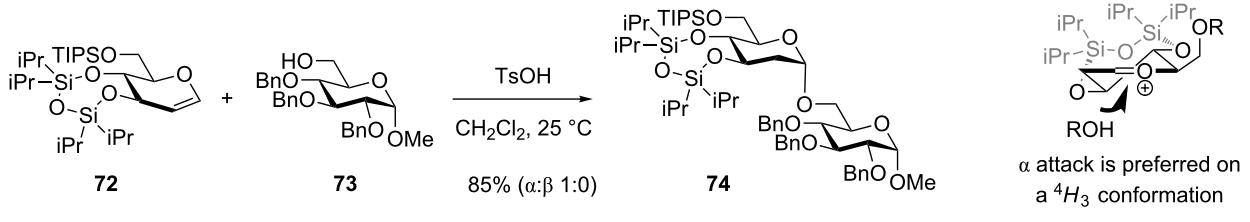
Scheme 16: Highly β-selective glucuronylation using a 2,4-DTBS-tethered donor.

Conclusion

Much indicates that glycosyl donors with silyl protective groups generally are more reactive than their alkylated counterparts presumably due to the *O*-silyl group being slightly less electron withdrawing than, e.g., a benzyl group. However, the reactivity increase is further augmented when bulky silyl groups, that cause a conformational change to an axial-rich conformation, are present. Such "superarmed" donors have a reactivity beyond what is obtained conventionally because the axial or pseudo-axial OR groups are less electron withdrawing. On the other hand, when the conformation is restricted by cyclic silyl protective groups (i.e., DTBS and TIPDS) and equatorial rich, comparatively unreactive donors result. Similarly, DTBS groups can be used to create superarmed donors by locking the conformation in an axial-rich state.

The silyl groups can also profoundly influence the stereoselectivity but in less obvious ways. Many TBS-protected donors are stereoselective – in some cases selectivities appear to be caused by steric hindrance from the 2-*O*-TBS group. For *C*-glycosides it has been possible to obtain conformationally derived stereocontrol so that persilylated donors adopting a $^1C_4$ conformation give the β-products. However, for *O*-glycosylation, this type of selectivity has been difficult to achieve.

Some very useful stereoselectivities are obtained with DTBS and TIPDS-protected galactosyl, mannosyl and arabinosyl donors. Here the selectivity is very much related to the conformational restriction and face-discrimination imposed by the cyclic silyl group upon the system.



Scheme 15: α-Selective glycosylation with a TIPDS-protected glucal donor.

References

1. Sauer, R. O. *J. Am. Chem. Soc.* **1944,** *66,* 1707–1710. doi:10.1021/ja01238a030
   See for the first alkyl silyl ethers prepared.
2. Lalonde, M.; Chan, T. H. *Synthesis* **1985,** 817–845. doi:10.1055/s-1985-31361
   See for early use as a protective group.
3. Muzart, J. *Synthesis* **1993,** 11–27. doi:10.1055/s-1993-25779
4. Saito, S.; Hirohara, Y.; Narahara, O.; Moriwake, T. *J. Am. Chem. Soc.* **1989,** *111,* 4533–4535. doi:10.1021/ja00194a078
   See for an early example of conformational restriction of non-carbohydrate ring systems mediated by bulky silyl ethers.
5. Broddefalk, J.; Bergquist, K.-E.; Kihlberg, J. *Tetrahedron Lett.* **1996,** *37,* 3011–3014. doi:10.1016/0040-4039(96)00456-X
6. Broddefalk, J.; Bergquist, K.-E.; Kihlberg, J. *Tetrahedron* **1998,** *54,* 12047–12070. doi:10.1016/S0040-4020(98)83057-3
7. Myers, A. G.; Liang, J.; Hammond, M. *Tetrahedron Lett.* **1999,** *40,* 5129–5133. doi:10.1016/S0040-4039(99)00880-1
8. Myers, A. G.; Glatthar, R.; Hammond, M.; Harrington, P. M.; Kuo, E. Y.; Liang, J.; Schaus, S. E.; Wu, Y.; Xiang, J.-N. *J. Am. Chem. Soc.* **2002,** *124,* 5380–5401. doi:10.1021/ja012487x
9. Schindler, C. S.; Bertschi, L.; Carreira, E. M. *Angew. Chem., Int. Ed.* **2010,** *49,* 9229–9232. doi:10.1002/anie.201004047
10. Tschamber, T.; Adam, S.; Matsuya, Y.; Masuda, S.; Ohsawa, N.; Maruyama, S.; Kamoshita, K.; Nemoto, H.; Eustache, J. *Bioorg. Med. Chem. Lett.* **2007,** *17,* 5101–5106. doi:10.1016/j.bmcl.2007.07.017
11. Valiullina, Z. R.; Khasanova, L. S.; Gimalova, F. A.; Selezneva, N. K.; Spirikhin, L. V.; Miftakhov, M. S. *Russ. J. Org. Chem.* **2014,** *50,* 1527–1533. doi:10.1134/S1070428014100194
12. Kopitzki, S.; Dilmaghani, K. A.; Thiem, J. *Tetrahedron* **2013,** *69,* 10621–10636. doi:10.1016/j.tet.2013.10.027
13. Gervay-Hague, J. *Acc. Chem. Res.* **2016,** *49,* 35–47. doi:10.1021/acs.accounts.5b00357
14. Schombs, M.; Park, F. E.; Du, W.; Kulkarni, S. S.; Gervay-Hague, J. *J. Org. Chem.* **2010,** *75,* 4891–4898. doi:10.1021/jo100366v
15. Kulkarni, S. S.; Gervay-Hague, J. *Org. Lett.* **2008,** *10,* 4739–4742. doi:10.1021/ol801780c
16. Hsieh, H.-W.; Schombs, M. W.; Witschi, M. A.; Gervay-Hague, J. *J. Org. Chem.* **2013,** *78,* 9677–9688. doi:10.1021/jo4013805
17. Hsieh, H.-W.; Schombs, M. W.; Gervay-Hague, J. *J. Org. Chem.* **2014,** *79,* 1736–1748. doi:10.1021/jo402736g
18. Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraser-Reid, B. *J. Am. Chem. Soc.* **1988,** *110,* 5583–5584. doi:10.1021/ja00224a060
19. Zhang, Z. Y.; Ollmann, I. R.; Ye, X.-S.; Wischnat, R.; Baasov, T.; Wong, C.-H. *J. Am. Chem. Soc.* **1999,** *121,* 734–753. doi:10.1021/ja982232s
20. Pedersen, C. M.; Marinescu, L. G.; Bols, M. *Chem. Commun.* **2008,** 2465–2467. doi:10.1039/b801305e
21. Jensen, H. H.; Pedersen, C. M.; Bols, M. *Chem. – Eur. J.* **2007,** *13,* 7576–7582. doi:10.1002/chem.200700947
22. Heuckendorff, M.; Pedersen, C. M.; Bols, M. *Chem. – Eur. J.* **2010,** *16,* 13982–13994. doi:10.1002/chem.201002313
23. Jensen, H. H.; Bols, M. *Acc. Chem. Res.* **2006,** *39,* 259–265. doi:10.1021/ar050189p
24. Pedersen, C. M.; Nordstrøm, L. U.; Bols, M. *J. Am. Chem. Soc.* **2007,** *129,* 9222–9235. doi:10.1021/ja071955l
25. Heuckendorff, M.; Pedersen, C. M.; Bols, M. *J. Org. Chem.* **2012,** *77,* 5559–5568. doi:10.1021/jo300591k
26. Hsu, Y.; Lu, X.-A.; Zulueta, M. M. L.; Tsai, C.-M.; Lin, K.-I.; Hung, S.-C.; Wong, C.-H. *J. Am. Chem. Soc.* **2012,** *134,* 4549–4552. doi:10.1021/ja300284x
27. Bhat, A. S.; Gervay-Hague, J. *Org. Lett.* **2001,** *3,* 2081–2084. doi:10.1021/ol0160405
28. Daly, R.; McCabe, T.; Scanlan, E. M. *J. Org. Chem.* **2013,** *78,* 1080–1090. doi:10.1021/jo302487c
29. Liang, X.-Y.; Bin, H.-C.; Yang, J.-S. *Org. Lett.* **2013,** *15,* 2834–2837. doi:10.1021/ol401166x
30. Mydock, L. K.; Demchenko, A. V. *Org. Biomol. Chem.* **2010,** *8,* 497–510. doi:10.1039/B916088D
31. Mydock, L. K.; Kamat, M. N.; Demchenko, A. V. *Org. Lett.* **2011,** *13,* 2928–2931. doi:10.1021/ol2009818
32. Kamkhachorn, T.; Parameswar, A. R.; Demchenko, A. V. *Org. Lett.* **2010,** *12,* 3078–3081. doi:10.1021/ol101089u
33. Mydock, L. K.; Demchenko, A. V. *Org. Lett.* **2008,** *10,* 2107–2110. doi:10.1021/ol800648d
34. Mydock, L. K.; Demchenko, A. V. *Org. Lett.* **2008,** *10,* 2103–2106. doi:10.1021/ol800345j
35. Premathilake, H. D.; Mydock, L. K.; Demchenko, A. V. *J. Org. Chem.* **2010,** *75,* 1095–1100. doi:10.1021/jo9021474
36. Crich, D.; Li, M. *Org. Lett.* **2007,** *9,* 4115–4118. doi:10.1021/ol701466u
37. Heuckendorff, M.; Premathilake, H. D.; Pornsuriyasak, P.; Madsen, A. Ø.; Pedersen, C. M.; Bols, M.; Demchenko, A. V. *Org. Lett.* **2013,** *15,* 4904–4907. doi:10.1021/ol402371b
38. Wang, S.; Meng, X.; Huang, W.; Yang, J.-S. *J. Org. Chem.* **2014,** *79,* 10203–10217. doi:10.1021/jo5018684
39. Jensen, H. H.; Nordstrøm, L. U.; Bols, M. *J. Am. Chem. Soc.* **2004,** *126,* 9205–9213. doi:10.1021/ja047578j
40. Heuckendorff, M.; Pedersen, C. M.; Bols, M. *J. Org. Chem.* **2013,** *78,* 7234–7248. doi:10.1021/jo4012464
41. Pedersen, C. M. Super Armed Glycosyl Donors – The best aid for a better reaction rate & Radical Azidonation – Novel safe methodologies. Ph.D. Thesis, University of Aarhus, Denmark, 2007.
42. Okada, Y.; Mukae, T.; Okajima, K.; Taira, M.; Fujita, M.; Yamada, H. *Org. Lett.* **2007,** *9,* 1573–1576. doi:10.1021/ol070427b
43. Felice, F. D.; Rúveda, E. A.; Stortz, C. A.; Colombo, M. I. *Carbohydr. Res.* **2013,** *380,* 167–173. doi:10.1016/j.carres.2013.08.002
44. Hosoya, T.; Ohashi, Y.; Matsumoto, T.; Suzuki, K. *Tetrahedron Lett.* **1996,** *37,* 663–666. doi:10.1016/0040-4039(95)02227-9
45. Abe, H.; Terauchi, M.; Matsuda, A.; Shuto, S. *J. Org. Chem.* **2003,** *68,* 7439–7447. doi:10.1021/jo030128+
46. Abe, H.; Shuto, S.; Matsuda, A. *J. Am. Chem. Soc.* **2001,** *123,* 11870–11882. doi:10.1021/ja011321t
47. Ichikawa, S.; Shuto, S.; Matsuda, A. *J. Am. Chem. Soc.* **1999,** *121,* 10270–10280. doi:10.1021/ja992608h
48. Tamura, S.; Abe, H.; Matsuda, A.; Shuto, S. *Angew. Chem., Int. Ed.* **2003,** *42,* 1021–1023. doi:10.1002/anie.200390261
49. Romero, J. A. C.; Tabacco, S. A.; Woerpel, K. A. *J. Am. Chem. Soc.* **2000,** *122,* 168–169. doi:10.1021/ja993366o
50. Kasai, Y.; Michihata, N.; Nishimura, H.; Hirokane, T.; Yamada, H. *Angew. Chem., Int. Ed.* **2012,** *124,* 8150–8153. doi:10.1002/ange.201203305
51. Yamada, H. *Trends Glycosci. Glycotechnol.* **2011,** *23,* 122–133. doi:10.4052/tigg.23.122
52. Heuckendorff, M.; Pedersen, C. M.; Bols, M. *J. Org. Chem.* **2012,** *77,* 5559–5568. doi:10.1021/jo300591k
53. Farias, M. A. M.; Kincaid, V. A.; Annamalai, V. R.; Kiessling, L. L. *J. Am. Chem. Soc.* **2014,** *136,* 8492–8495. doi:10.1021/ja500622v

54. Totani, K.; Shinoda, Y.; Shiba, M.; Iwamoto, S.; Koizumi, A.; Matsuzaki, Y.; Hirano, M. *RSC Adv.* **2015,** *5,* 75918–75922. doi:10.1039/C5RA16659D
55. Imamura, A.; Ando, H.; Korogi, S.; Tanabe, G.; Muraoka, O.; Ishida, H.; Kiso, M. *Tetrahedron Lett.* **2003,** *44,* 6725–6728. doi:10.1016/S0040-4039(03)01647-2
56. Hada, N.; Oka, J.; Nishiyama, A.; Takeda, T. *Tetrahedron Lett.* **2006,** *47,* 6647–6650. doi:10.1016/j.tetlet.2006.06.181
57. Gold, H.; Boot, R. G.; Aerts, J. M. F. G.; Overkleeft, H. S.; Codée, J. D. C.; van der Marel, G. A. *Eur. J. Org. Chem.* **2011,** 1652–1663. doi:10.1002/ejoc.201001690
58. Laroussarie, A.; Barycza, B.; Andriamboavonjy, H.; Kenfack, M. T.; Blériot, Y.; Gauthier, C. *J. Org. Chem.* **2015,** *80,* 10386–10396. doi:10.1021/acs.joc.5b01823
59. Chen, L.; Kong, F. *Tetrahedron Lett.* **2003,** *44,* 3691–3695. doi:10.1016/S0040-4039(03)00673-7
60. Zhu, X.; Kawatkar, S.; Rao, Y.; Boons, G.-J. *J. Am. Chem. Soc.* **2006,** *128,* 11948–11957. doi:10.1021/ja0629817
61. Fedina, K. G.; Abronina, P. I.; Podvalnyy, N. M.; Kondakov, N. N.; Chizhov, A. O.; Torgov, V. I.; Kononov, L. O. *Carbohydr. Res.* **2012,** *357,* 62–67. doi:10.1016/j.carres.2012.05.021
62. Crich, D.; Pedersen, C. M.; Bowers, A. A.; Wink, D. J. *J. Org. Chem.* **2007,** *72,* 1553–1565. doi:10.1021/jo061440x
63. Ishiwata, A.; Akao, H.; Ito, Y. *Org. Lett.* **2006,** *8,* 5526–5528. doi:10.1021/ol062198j
64. Abronina, P. I.; Fedina, K. G.; Podvalnyy, N. M.; Zinin, A. I.; Chizhov, A. O.; Kondakov, N. N.; Torgov, V. I.; Kononov, L. O. *Carbohydr. Res.* **2014,** *396,* 25–36. doi:10.1016/j.carres.2014.05.017
65. Balmond, E. I.; Benito-Alifonso, D.; Coe, D. M.; Alder, R. W.; McGarrigle, E. M.; Galan, M. C. *Angew. Chem.* **2014,** *126,* 8329–8333. doi:10.1002/ange.201403543
66. Dharuman, S.; Crich, D. *Chem. – Eur. J.* **2016,** *22,* 4535–4542. doi:10.1002/chem.201505019
67. Furukawa, T.; Hinou, H.; Nishimura, S.-I. *Org. Lett.* **2012,** *14,* 2102–2105. doi:10.1021/ol300634x

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Fluorescent carbon dots from mono- and polysaccharides: synthesis, properties and applications

Stephen Hill and M. Carmen Galan*

Review

Open Access

Address:
School of Chemistry, University of Bristol, Cantock's Close, Bristol BS8 1TS, UK

Email:
M. Carmen Galan* - m.c.galan@bristol.ac.uk

* Corresponding author

Keywords:
fluorescent carbon dots; monosaccharides; nanomaterials; nanotechnology applications; polysaccharides

Beilstein J. Org. Chem. **2017,** *13,* 675–693.
doi:10.3762/bjoc.13.67

Received: 23 January 2017
Accepted: 30 March 2017
Published: 10 April 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Fluorescent carbon dots (FCDs) are an emerging class of nanomaterials made from carbon sources that have been hailed as potential non-toxic replacements to traditional semiconductor quantum dots (QDs). Particularly in the areas of live imaging and drug delivery, due to their water solubility, low toxicity and photo- and chemical stability. Carbohydrates are readily available chiral biomolecules in nature which offer an attractive and cheap starting material from which to synthesise FCDs with distinct features and interesting applications. This mini-review article will cover the progress in the development of FCDs prepared from carbohydrate sources with an emphasis on their synthesis, functionalization and technical applications, including discussions on current challenges.

Introduction

Nanotechnology applied to biological and biomedical problems has seen an explosion of research in recent years [1]. Functional nanomaterials that can carry biologically relevant molecules have become very useful for drug delivery, sensing and catalysis to name just a few applications. As a result, nanomaterials exhibiting novel electronic and optical properties, having controlled size, geometry, surface distribution and functionality have been developed as materials for probing biological interactions and in biomedical applications [2-6]. Among these novel type of probes, luminescent semiconductors, quantum dots (QDs), which possess a narrow emission spectra and common excitation, superior photostability and electron density when compared to organic fluorophores, in addition to bright visible emission, have become particularly popular for their versatility as non-isotopic detection labels which are amenable to live cell imaging and immunoassay applications [7]. In particular, cadmium-based QDs (e.g., CdS, CdSe, CdSe/ZnS) are commonly used for in vitro biological studies due to their well-established synthesis and functionalisation strategies, tuneable emission profiles and high quantum yields of fluorescence

(QYs) [8-11]. However, the presence of heavy metals like Cd^{2+}, and the associated concerns surrounding heavy metal toxicity has meant that their in vivo applications are restricted [12]. Therefore, the development of fluorescent nanoparticles that are able to replicate QD fluorescence properties without exhibiting long term toxicity profiles, has become very relevant.

The term carbon dots (CDs) has been coined to describe a new class of carbon-based nanomaterials which are typically discrete, quasi-spherical nanoparticles, with sizes usually less than 10 nm in diameter (although bigger sizes have recently been reported). These relatively new nanomaterials have found many applications in the fields of photo- and electrocatalysis, chemical sensing, biosensing, bioimaging and nanomedicine, due to their unique tuneable photoluminescence (PL) properties, chemical inertness, high water solubility, ease and low cost fabrication and more importantly, low toxicity profiles. The latter makes these fluorescent nanomaterials attractive for a wide range of in vivo applications, which has been the topic of several recent reviews [13-15]. Following the serendipitous discovery by Xu et al. during the separation and purification of single-walled carbon nanotubes (SWCNTs) [16], the development of synthetic methodologies to access these fluorescent nanomaterials combined with their myriad of applications, has led to CDs being hailed as the potential non-toxic successors to traditional semiconductor QDs, particularly in the areas of live imaging and drug delivery.

Synthetic approaches to access CDs can be classified into two broad categories: top-down or bottom-up syntheses. Top-down methods are characterised by using a bulk carbon substrate as the starting material; using conditions that remove nanoparticles from the bulk substrate such as electrochemistry, chemical oxidation, arc discharge or laser ablation, carbon-based nanoparticles can be obtained. Typical substrates used are single/multi-walled carbon nanotubes, graphite, graphene or candle soot, amongst many others [15,17]. The crystalline make-up of top-down derived CDs is usually highly sp^2 in character, which is transferred from the sp^2-enriched starting materials, e.g., graphite or graphene. Conversely bottom-up methodologies rely on the use of a molecular precursor which can be treated in such a way as to seed the formation of a CD. Typical starting materials include amino acids, citric acid, biomass and carbohydrates to name but a few, which can be reacted using thermal decomposition, chemical or hydrothermal oxidation, microwave, acid-mediated reflux, ultrasonic irradiation or silica nanoparticle-templated synthesis [18-23]. Unlike their top-down equivalents, the CDs derived from these methods are usually less sp^2 crystalline and tend to have more amorphous morphologies. It should be stated that no two CD preparations lead to the same type of nanoparticle, as any changes to the ratio and composition of starting materials, additives, solvent, temperature, type of vessel, etc., does have an effect on the final molecular composition and architecture of the CD. Resultantly, differential properties are easily acquired through minor manipulations of the CD synthesis. To date, the de novo rational design of bottom-up syntheses of CDs for advanced application is limited in the literature.

Carbohydrates are one of the most diverse and important class of biomolecules in nature and offer well-defined chiral scaffolds primed for modification at the anomeric position and alcohol functionalities. Therefore, the use of carbohydrates as a starting material for synthesizing CDs is extremely attractive not only due to their abundance, availability and heterogeneity, but also due to their high water solubility, low-carbonisation temperatures, low cost and typically inherently lack toxicity. With all these options available to tune the synthesis of CDs, it is no surprise that researchers have already began to see the benefits of carbohydrates when considering the synthesis of novel FCDs with improved properties. For example, simple monosaccharides such as glucose, glucosamine, mannose, fructose and their derivatives and common disaccharides, e.g., sucrose, lactose, and maltose have been employed to prepare fluorescent carbon dots (FCDs) using different methodologies [13,24]. Similarly, important carbohydrate-based biopolymers such as cellulose, chitin, chitosan, dextran, cyclodextrin, and hyaluronic acid, which differ not only in elemental composition, but also in chemo-physical properties, have also been successfully utilised in the preparation of CDs, where their differences allow tailoring of the CD structure and properties [25].

In this review, we focus on the most recent approaches developed to prepare fluorescent CDs using mono-, oligo- and polysaccharides as the main carbon source.

Review

Fluorescent carbon dots synthesised from monosaccharides

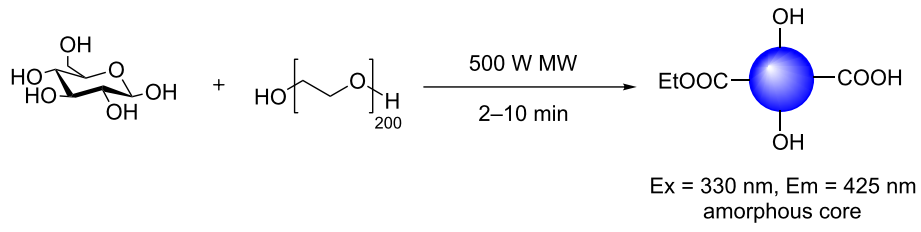
Glucose-based fluorescent carbon dots

Sustainable syntheses of CDs have driven researchers to find readily available, cheap and renewable carbon sources of which the monosaccharide glucose is an ideal candidate. Not only is glucose cheap and commercially available, but also has a low carbonisation temperature, ring-opens readily to afford a reactive aldehyde moiety which can be further exploited for conjugations, polymerisations and (hetero)aromatic formation, which are all ideal for generating CDs [26,27]. For these reasons, in addition to the inherent low toxicity and high water solubility of glucose, this particular monosaccharide has been extensively used as an ideal carbon source for CD formation, under a range of experimental conditions.

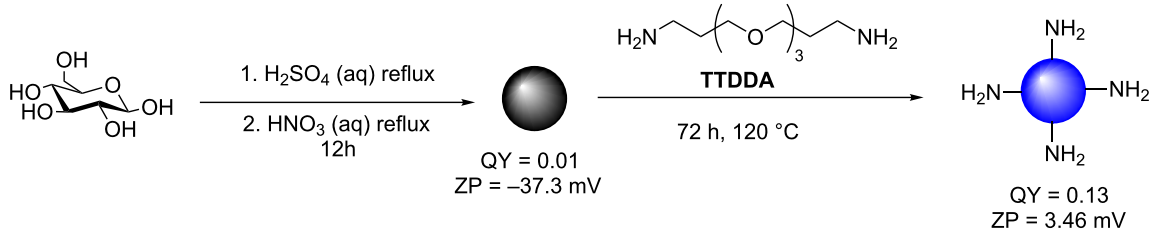
The microwave-assisted synthesis of FCDs from a glucose solution in the presence of poly(ethylene glycol)-200 (PEG-200) by Yang et al. is, to the best of our knowledge, the first reported example involving a carbohydrate moiety (Scheme 1) [18]. The water-soluble nanoparticles exhibited an amorphous core, as deduced by X-ray diffraction (XRD), while Fourier-transformed infrared (FTIR) spectroscopy analysis indicated the presence of a range of oxygen-containing functionalities, e.g., alcohols, ethers and carboxylic acids on the CD surface, which are likely the reason for the high water solubility exhibited by the nanomaterial. This type of chemical profile is typical of standard bottom-up synthesised CDs [15,28]. Interestingly, the team was also able to show that the use of PEG-200, as a surface passivation agent (SPA), was crucial for favourable photoluminescence (PL) properties and QYs of up to 6.3% were achieved. The use of SPAs is among one of two main techniques that are widely employed to improve the PL properties of FCDs. SPAs are argued to provide uniform PL trapping sites on the CD surface, alongside promoting new functionality that can work, in tandem with the core, to turn-on fluorescence.

Another example, which highlights the importance of surface passivation and how SPAs can be used to modify and tune CD chemical and physical properties, was reported by Travas-Sejdic et al. [23]. They also employed glucose as the carbon precursor which, after refluxing in aqueous H_2SO_4, yielded carbonaceous nanoparticles with observable PL (Scheme 2). Further treatment with aqueous HNO_3 under reflux, yielded nanoparticles of weak PL (QY = 1%). The PL properties could be improved upon introducing surface passivation, which was achieved by heating the weakly fluorescent CDs in a solution of 4,7,10-trioxa-1,13-tridecanediamine (TTDDA) for 72 hours at 120 °C to give a nanomaterial with QY values of up to 13%. FTIR studies suggested that TTDDA incorporation onto CDs occurred via amide formation, from the reaction between surface carboxylic acids and the corresponding amine SPA. This was further supported by the change in zeta-potential (ZP) values which shifted from −37.3 mV (non-passivated CD) to 3.46 mV (TTDDA passivated CD). Similar carbonaceous materials were obtained when the team used sucrose or starch as starting carbohydrate materials.

In addition to microwave and acid reflux-mediated glucose dehydration reactions, the group of Wang developed an alternative protocol that combined glucose with monopotassium phosphate (KH_2PO_4) in a Teflon-lined autoclave chamber with heating to 200 °C for 12 h (Scheme 3) [29]. The fluorescence emission could be tuned by changing the ratio of sugar and KH_2PO_4. For instance a molar ratio of 1:26 (glucose/KH_2PO_4) afforded blue-fluorescent CDs (QY = 0.02), whereas a 1:36 ratio yielded green-fluorescent CDs (QY = 0.01). In the absence of KH_2PO_4, irregular black carbon aggregates were obtained. Raman and TEM analysis showed both types of FCDs had graphitic crystallinity. This example highlights that an inorganic-based dehydrating agent could be used instead of a traditional diamine SPA to induce CD dehydration and affect their PL properties. Most carbohydrate-derived CDs emit in the blue area of the visible section of the electromagnetic spectrum under UV/high energy blue excitation. However, most mammalian cells are also autofluorescent in this particular

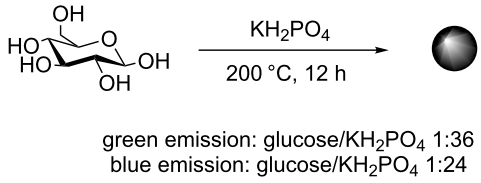


Scheme 1: Microwave-driven reaction of glucose in the presence of PEG-200 to afford blue-emissive CDs.



Scheme 2: Two-step synthesis of TTDDA-coated CDs generated from acid-refluxed glucose.

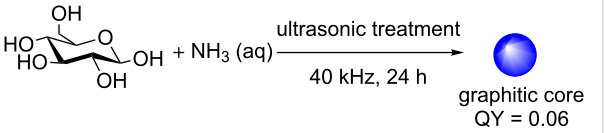
region [30]. As a result, the majority of CDs produced with blue emission have QYs that are not suitable for bioimaging applications. CDs with multicolour/excitation-dependent emission that can be red shifted and avoid the cellular autofluorescence window, are a good alternative. Unfortunately, the CD fluorescence tends to lose intensity upon red-shifting the excitation. An ideal CD probe for bioimaging applications will have either a high QY in the blue, or adequate green to red emission. Thus, the green-emissive glucose-based CDs produced by Wang et al. are ideal for this type of application and the team showed their applicability in cell internalisation studies with HepG2 cells [29]. The green CDs were non-toxic to cells at concentrations of up to 625 μg/mL and exposures of 72 h. Laser scanning confocal microscopy (LSCM) demonstrated cell internalization, making these materials a good candidate as a bioimaging agent.



Scheme 3: Glucose-derived CDs using KH_2PO_4 as a dehydrating agent to both form and tune CD's properties.

In 2011 Qu et al. developed a tuneable synthesis of FCDs by selecting a different inorganic ion and carbohydrate combinations using microwave irradiation as the heating source, demonstrating that both the starting material and dehydrating agent of choice can allow tuning/manipulation of the fluorescence properties of the system [31]. It was found that irradiation times of 14 min could be employed to afford CDs from glycerol, glycol, glucose or sucrose. The source of the inorganic ion was important too, as increasing the valency of either the anion or cation would lead to a greater ability to dehydrate the carbon precursor. An ideal balance of cation and anion valence was found when using $CuSO_4$ which afforded CDs with QY of up to 9.5%, in-line with the state-of-the-art at the time.

As an alternative, Kang et al. showed the following year that nitrogen-doped water-soluble fluorescent CDs could be afforded in a one-step ultrasonic reaction of glucose and aqueous ammonia [NH_3 (aq)] in solution (Scheme 4) [32]. The CDs are generated by the formation of small vacuum bubbles in solution by alternating high and low pressure waves. This process leads to temperature increases, hydrodynamic shear forces and the formation of high-speed liquid jets in solution. All of these effects facilitate the degradation of glucose and the incorporation of ammonia into the CD structure. The introduction of N-doping allows the injection of electrons into the CD structure, which allows for new PL and fluorescence properties to be established. This is a widely employed strategy for improving the QY of CDs. The team was able to show that the presence of the dopant yielded N-doped CDs with a QY of 6%, which was superior to the N-free CDs. The resultant CDs afforded from the ultrasonic treatment were well dispersed, with TEM indicating graphite crystallinity with blue-green emission.



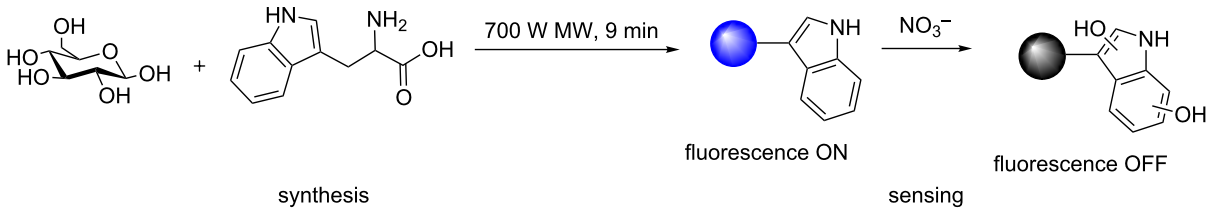
Scheme 4: Ultrasonic-mediated synthesis of glucose-derived CDs in the presence of ammonia.

Most reported CD syntheses, regardless of the type of starting material or synthetic method, tend to produce CDs with blue-green fluorescence emission. Jana et al. reported a carbohydrate-based preparation to access yellow and red emissive CDs, demonstrating that fine-tuning the reaction conditions, combined with the use of additives, can lead to modifications in the emissive properties of the nanoparticles [20]. The team was able to show that the use of sulfuric acid with a carbohydrate, although the exact carbohydrate used is not disclosed in the article, generally gave blue-green emission, whereas phosphoric acid-mediated CD formation gave particles with red-shifted emission profiles (Red-CD). Interestingly, Red-CDs only had stable red emission in strong acidic conditions; changing to green emission in neutral or basic pH. The origin of the tuneable fluorescence was attributed, by the authors, to a number of contributing factors including the different sizes of the nanoparticles as key. As the CD size increases, the size of sp^2 domains can also increase (controlled by the dehydrating agent). Additionally, the introduction of certain heteroatoms can allow for red-shifted emission with increasing size. Moreover, the presence of defect sites, which are associated to the PL properties, is confirmed using room temperature electron paramagnetic resonance (EPR) spectroscopy, by the presence of free electrons in the spectra. Further functionalization of the hydrophobic red fluorescent CDs via surface passivation and polymer coating, in which hydrophilic anhydride groups of the polymer can react with PEG-diamine, via a ring-opening to afford a free acid and an amide-linked SPA, lead to water-soluble CDs ready for biolabelling applications. The CDs were then labelled with either TAT (a cell-penetrating peptide), or folate and then incubated with HeLa cells. Fluorescence microscopy images confirmed that incubation times of 3–6 hours were adequate to allow for CD labelling of the cells. Further toxicity assays indicated that concentrations of up to 200 μg/mL were tolerated, as determined by cell viability studies (MTT assay). More recently, the

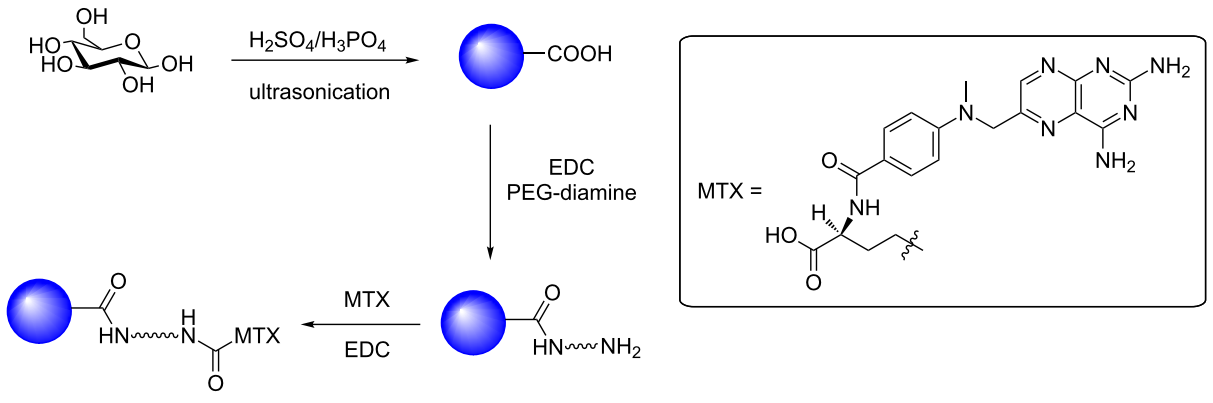
same team developed hydrophobic yellow and red emissive CDs via the degradation of ascorbic acid in the presence of oleophilic oleylamine [13]. The CDs were similarly polymer coated with poly(maleic anhydride-*alt*-1-octadecene) that could subsequently be functionalised with hydrophilic PEG-diamine, providing an amine functionality for conjugation with glucosamine, histidine, arginine and folate. The yellow/red emissive loaded-CDs were shown to be viable bioimaging probes in live cells, since their emission did not overlap with the cell autofluorescence.

Recent years have seen an increase in synthetic reports of large scale N-doped CDs with good QYs from carbohydrate starting materials. For example in 2014, Leitão et al. described the microwave synthesis of CDs using 2.5 g of glucose and 0.3 g of tryptophan as the N-dopant/surface passivation agent (Scheme 5) [33]. The resultant CDs had a QY of up to 12% (34-times higher than that of the undoped CDs). Interestingly, the N-doped CDs in this report had a 20 nm diameter, as determined by TEM, which is contrary to the generally held belief that CDs' particular properties are only observed below a diameter of 10 nm, which is not the case here and has since been observed in one other carbohydrate-derived CD synthesis [34]. The team demonstrated the utility of the glucose/tryptophan-derived CDs as a sensor of peroxynitrite anions (NO_3^-) in solution. The peroxynitrite anion is one of the key reactive species which is implicated in various metabolic and physiological processes [35]. Thus, it is important to provide analytical methods to detect and quantify its presence, however, due to its high reactivity, low concentration levels and quick diffusion, it has been traditionally difficult to detect. The team was able to show significant quenching of the CDs via tryptophan oxidation of the exposed residues on the surface of CDs (Scheme 5). Post-oxidation fluorescence is compromised and therefore can be used as a signal for selectively sensing peroxynitrite up to concentrations of 1.5 μM (with a linear regression between 2.5–50 μM). The sensing ability of the nanoparticles was exhibited in serum-fortified samples, which can be regarded as a biomimetic for complex biological media.

A number of glucose-based CDs has been reported in recent years as drug-delivery vehicles. In 2015, Yunus et al. synthesized CDs by the ultrasonication of glucose or sucrose in the presence of oxidising conditions afforded by H_3PO_4/H_2SO_4 (Scheme 6) [36]. The resultant CDs were blue emissive and the use of strong oxidising conditions during their synthesis afforded CDs with surface carboxylic acids that could be functionalised. Surface conjugation with PEG-diamine afforded a steric blocking, enhanced permeation and retention (EPR) shell, whilst providing an amine functionality for further surface conjugation. The anticancer drug methotrexate (MTX), which is a well-studied drug used to treat various types of cancer includ-



Scheme 5: Tryptophan-derived CDs used for the sensing of peroxynitrite in serum-fortified cell media.



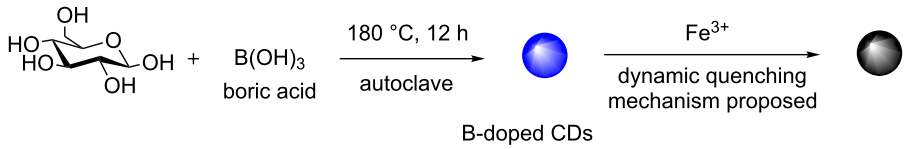
Scheme 6: Glucose-derived CDs conjugated with methotrexate for the treatment of H157 lung cancer cells.

ing lung cancer, was then conjugated via EDC-mediated amide coupling chemistry (Scheme 6) [37]. The MTX-CDs were internalised into H157 lung cancer cells and compared with cells exposed to unfunctionalised amine-bearing CDs. While the amine-CDs showed no cellular toxicity, MTX-CDs were highly toxic to H157 cell cultures, highlighting the potential applicability of carbohydrate-derived CDs as vehicles for the delivery of conventional cancer therapeutics.

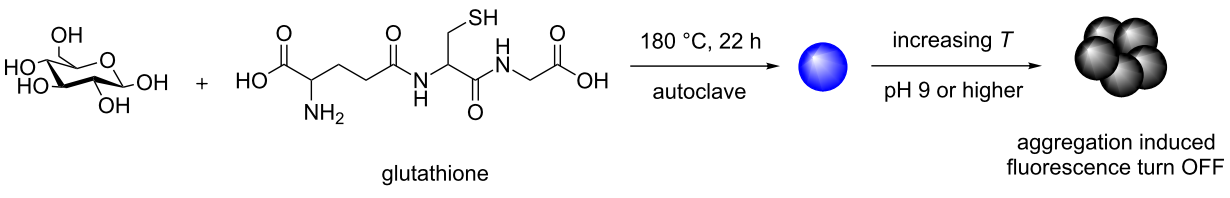
More recently, in addition to the introduction of electron-donating heteroatoms such as N or S as dopant agents to improve the PL properties of CDs, the use of boron as an additive, which is an electron-accepting element, has also been explored by Hao et al. [38]. The CDs were produced by the addition of boric acid ($B(OH)_3$) into the hydrothermal carbonisation of glucose, using a Teflon autoclave at 180 °C for 12 h (Scheme 7). The resultant fluorescent nanoparticles had an average diameter of 4 nm and were negatively charged with ZP values of −40.7 mV. XPS and FTIR analysis confirmed the presence of B in the CD structure. Although, the addition of boron did not change the typical blue fluorescence profile significantly, when compared to other reported heteroatom-doping syntheses, the fluorescence of the B-CDs was dynamically quenched by Fe^{3+} ions. Mechanistic studies suggested that a dynamic quenching model was prevalent at low concentrations due to interactions between Fe^{3+} and the CD surface, possibly indicating the interception of an excited CD state by the Fe^{3+} ion that leads to fluorescence quenching. The group exemplified the applicability of the material by demonstrating the ability of the B-CD to sense Fe^{3+} in tap water samples with a limit of detection of 242 nM, which complies with U.S. Environmental Protection Agency standards.

Having shown that chemical doping with heteroatoms within the CD synthesis can lead to materials with improved PL and physicochemical properties, this has encouraged research groups to focus their efforts to study the effect of using different heteroatoms simultaneously. For example, in 2015 Zhang et al. demonstrated that the energy intensive, hydrothermal treatment of glucose in the presence of glutathione (acting as both SPA and N/S heteroatom dopant) conducted at 180 °C for 22 h, resulted in CDs with QYs of up to 7%. The obtained CDs had blue-emissive fluorescence under UV excitation, a standard feature of bottom-up produced carbohydrate-derived CDs (Scheme 8) [39]. Whereas most synthetic CDs from glucose sources generally show a fluorescence-decay response to one or several transition metals. Surprisingly, the CDs produced in the presence of glutathione had a very stable fluorescence output, which was unaffected by a wide-range of transition metal cations. The new CD's fluorescence intensity was, however, sensitive to changes in both pH and temperature. The CDs were shown to aggregate and change emission from pH 3 to 9, which the authors attribute to the ionisation of the surface functionality (Scheme 8). The feature is reversible as demonstrated by monitoring the PL intensity at a given excitation at different pHs and during several iterations. Similarly the CDs were shown to have an emission-intensity dependence on the temperature. Upon increasing the temperature from 15 to 90 °C, 52% of the fluorescence was lost, without any red or blue shift in the emission maximum. The mechanism for this change was also attributed to nanoparticle aggregation, with CD agglomeration occurring at higher temperatures.

Two different groups reported, nearly concurrently their results, in the use of both N and P as dopants in their CD syntheses. For



Scheme 7: Boron-doped blue-emissive CDs used for sensing of Fe^{3+} ion in solution.



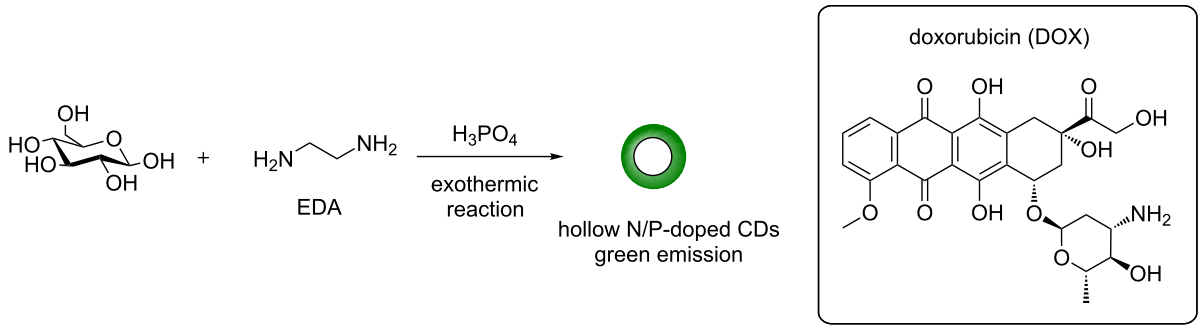
Scheme 8: N/S-doped CDs with aggregation-induced fluorescence turn-off to temperature and pH stimuli.

instance, Dong et al. described the development of dual-doped hollow CDs using a wet-chemical method from glucose in the presence of 1,2-ethylenediamine (EDA) and conc. H_3PO_4 (Scheme 9) [40]. During the exothermic reaction, a foam was produced which resulted in green-emissive N,P-doped CDs that were hollow and had a diameter of 10 nm as determined by HRTEM and AFM. Despite the hollowness of the CDs, the nanoparticles exhibit an excitation-dependent emission profile akin to other glucose-derived, bottom-up synthesised CDs. The material's unique properties deemed them ideal candidates for drug-delivery purposes. The team chose doxorubicin (DOX) as the model drug for this purpose and loading of DOX onto the CD was demonstrated via a change in the ZP from −9.3 mV to −0.13 mV, suggesting that an electrostatic interaction between the positively charged amino group in DOX and the negatively charged groups on the CD surface can take place. Also van der Waals and π–π stacking interactions were attributed as contributing factors to the DOX loading and DOX incorporation into the hollow CD cavity. Drug release at acidic pHs, further supported the proposed electrostatic interactions between DOX and the CDs. Further studies showed initial efficacy of the DOX–CD adduct as a beneficial drug-delivery system, even in animal models.

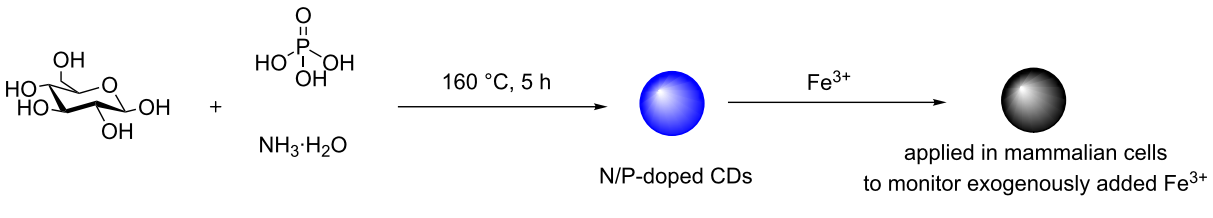
On the other hand, Zhao et al. described an alternative synthesis for N/P-doped CDs. Hydrothermal oxidation of glucose, phosphoric acid and aqueous ammonia, as the nitrogen source, in a Teflon-lined autoclave followed by heating at 160 °C for 5 h afforded blue-emissive CDs under 365 nm excitation (Scheme 10) [41]. A high QY of 30% was obtained, which is one of the highest reported for a carbohydrate-derived CD to date. Interestingly, it was observed that the fluorescence of these N/P-CDs was strongly dependent on the local concentration of Fe^{3+}. With increasing concentrations of the metal leading to fluorescence decay of the CDs, which was attributed to the interception of an excited state on the CD by the Fe^{3+} ion. The selectivity towards Fe^{3+} was demonstrated against a panel of other transition and alkali metals and a detection limit for Fe^{3+} of 1.8 nM was established. The glucose-derived blue-emissive CD could be readily internalised into T24 cells, without significant cell death, and used to detect the presence of exogenously added Fe^{3+} (Scheme 10).

It is important to highlight that small changes in the nitrogen source (EDA vs ammonia), ratio of reagents and reaction conditions can lead to marked differences in fluorescence, physical and chemical properties of the nanomaterials, as demonstrated with these two parallel reports for hollow-green and solid blue-emitting CDs.

A systematic study has been performed by the Travas-Sejdic group on the synthesis of CDs from either citric acid or glucose starting materials in the presence of either TTDDA or dopamine, in order to evaluate how the choice of carbon and nitrogen sources plays a key role in the final properties of these nanomaterials [42]. The authors found that the average size of



Scheme 9: N/P-doped hollow CDs for efficient drug delivery of doxorubicin.



Scheme 10: N/P-doped CDs applied to the sensing of Fe^{3+} ions in mammalian T24 cells.

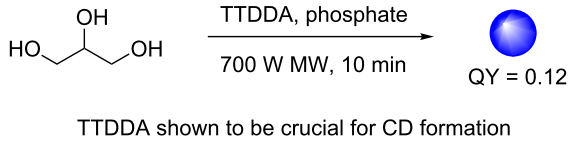
CDs prepared was dependent on both, the carbon source, e.g., CDs from citric acid were larger than the ones derived from glucose, and the nitrogen source, e.g., CDs derived from dopamine were larger than those using TTDDA. The authors attribute this observation to the fact that citric acid possesses readily available carbonyl groups (as opposed to the masked aldehyde in the carbohydrate) that can readily react with basic TTDDA or dopamine to form stable intermediates; while glucose mostly interacts with the amine dopants through intramolecular forces such as van der Waals' forces and hydrogen bonds. The latter weaker interactions cause the intermediates to break down into small fragments during the heating process resulting in smaller CDs. In the case of N-dopant agents, the presence of a bulky phenyl ring in dopamine was reasoned to be the possible cause for the somewhat larger sizes observed. In addition, it was found that PL properties were mostly dependent on the N-source, with optimum QY of up to 29.5% (for glucose) or 33.9% (for citric acid) when using TTDDA, as opposed to dopamine (Scheme 11). These results show that by experimentally probing the reaction conditions and fully characterising the obtained materials, a better understanding of the underpinning mechanisms of CD formation and PL mechanisms will be gained, which in turn will lead to improved materials with high QYs.

Non-glucose monosaccharide-based fluorescent carbon dots

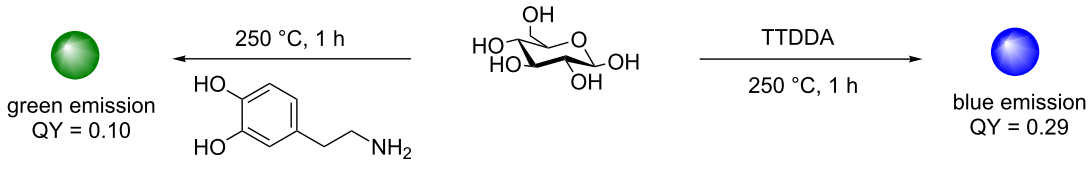
In addition to glucose, different monosaccharides and polyols have also been utilised as carbon sources for the synthesis of FCD, although this approach is less common.

The ability of glycerol to undergo dehydration and polymerisation in the presence of amino groups makes it a cheap and suitable candidate as a molecular precursor for CD synthesis. To that end, Liu et al. demonstrated in 2011 that the microwave-assisted pyrolysis of glycerol in the presence of TTDDA afforded blue-emissive CDs with a QY of 12% (Scheme 12) [43]. The particles had preeminent multicolour emission, which was excitation-dependent. Importantly, the team demonstrated that TTDDA was crucial as a passivating agent for optimal levels of fluorescence. The method was also applicable to other carbon sources such as glucose, sucrose, glucan and starch. The novel nanomaterials were found to be useful in live cell bioimaging applications. The team carried out cell viability studies (MTT assay) and after treatment of HepG2 cells with these multicolour emissive CDs, 100% cell viability was recorded with concentrations of up to 240 μg/mL of the CDs, while significant toxicity was seen at concentrations at and above 400 μg/mL. CDs (100 μg/mL) were also incubated with HepG2 for 24 h and laser scanner confocal microscopy (LCSM) was used to image the internalization of the CDs within the cells using the green, yellow and red channels, demonstrating their utility.

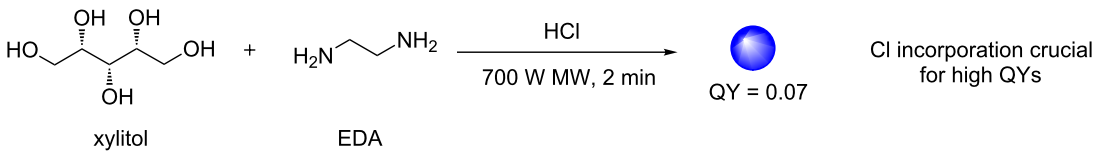
In a similar fashion, xylitol was used as a CD molecular precursor, in the presence of HCl and ethylene diamine (EDA), in a 2 min microwave-mediated synthesis of CDs developed by Kim et al. [44]. The team successfully demonstrated that to improve the blue emission of the nanoparticles, HCl was crucial. In the absence of HCl as an additive, the QY was only 0.38%, whereas in the presence of HCl, a significant increase of the QY to 7% was observed (Scheme 13). Interestingly, Cl atoms are incorporated as part of the CD structure in as much as 9.14%, based on the elemental analysis, demonstrating that in addition to N-dopant agents, Cl sources such as HCl, are key as SPAs that can improve the PL properties of FCDs. The Cl/N-doped CDs were incubated with WI38 and HeLa cell lines and the cell viability was studied by an MTT assay. It was found that no cytotoxicity was observed up to CD concentrations of 100 μg/mL, while slight levels of toxicity were detected at concentrations of up to 1000 μg/mL. Fluorescence microscopy analysis of HeLa cells treated with CDs at 100 μg/mL for 24 h showed cell internalisation as monitored by their multicolour emissive properties, in addition LCSM confirmed their remarkable photostability too, as long exposure times lead to no obvious photobleaching.



Scheme 12: Formation of blue-emissive CDs from the microwave irradiation of glycerol, TTDDA and phosphate.



Scheme 11: Comparative study of CDs formed from glucose and N-doped with TTDDA and dopamine.

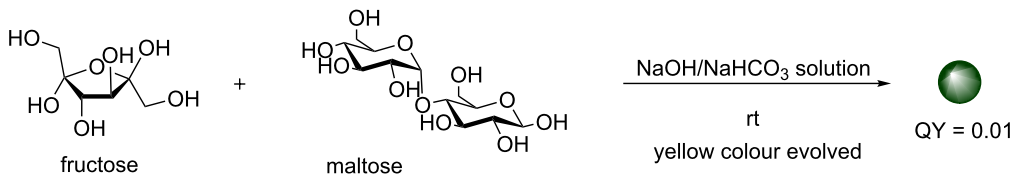


Scheme 13: Xylitol-derived N-doped CDs with excellent photostability demonstrating the importance of Cl incorporation to the fluorescence properties.

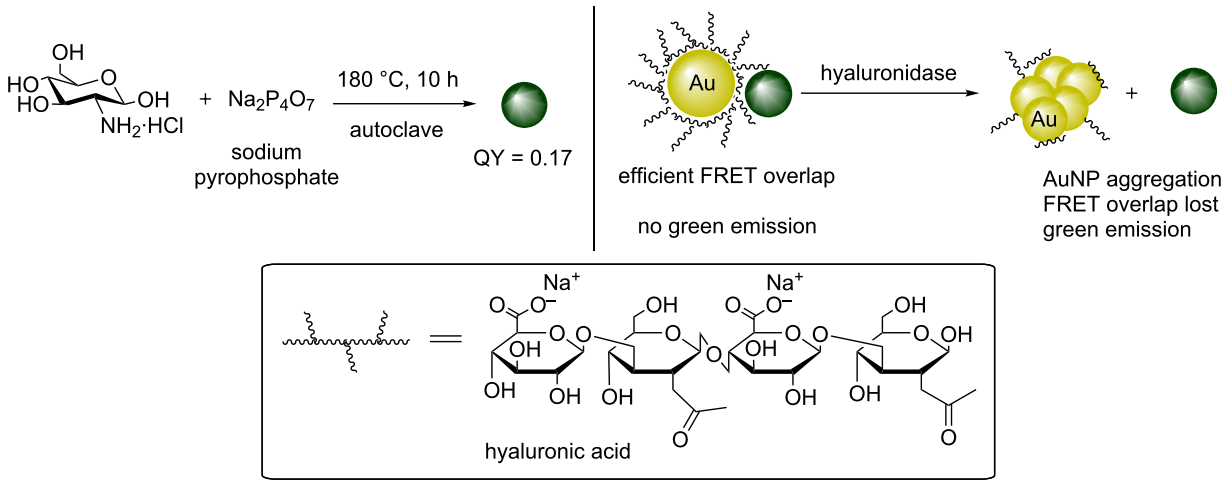
Fructose and maltose combinations have also been used as an alternative to glucose as the carbon source. The Ostrikov team developed a room temperature preparation of weakly emissive CDs (QY 2%) by mixing a 500 mM aqueous solution of fructose and maltose (a glucose 1,4-linked disaccharide) with a 500 mM solution of NaOH and $NaHCO_3$ also dissolved in water (Scheme 14) [45]. The resulting clear mixture was monitored until a colour change towards a yellow colouration was observed after approximately 60 minutes of mixing. Upon excitation by 405 nm lasers, a green fluorescence was recorded. It was found that the concentration of the solution was essential for the formation of CDs. Although this method does not produce highly fluorescent CDs, the example shows that green-emitting CDs can be made in a reaction without either strong heating, N-doping or surface passivation occurring. Also remarkably, the CDs produced by this method were found by HRTEM to have graphite crystallinity. This feature is interesting as, until now, it was thought that this type of crystallinity in a bottom-up constructed nanomaterial was only possible under energy intensive/forcing conditions.

We have already established that an effective method for modulating the properties of CDs is to introduce heteroatoms, with the use of N-dopant agents being the most common. The majority of methods discussed thus far, utilise cheap, readily available neutral carbohydrate such as glucose as the carbon source in combination with a nitrogen-containing molecule. Glucosamine hydrochloride, which is a byproduct from the hydrolysis of chitosan and chitin polysaccharides found on crustacean shells, bears an amine functionality at C-2 and offers all the advantages of glucose, while already containing an N atom. A few examples in the literature have already utilised this sugar as the starting material in the synthesis of CDs with interesting results. One of the earliest examples of the hydrothermal preparation of FCDs using glucosamine hydrochloride was shown by Wang et al. [34]. A one-step process whereby an aqueous (deionised) solution of the amine-containing glucoside was heated in an autoclave to 140 °C for 12 h, which after several days of dialysis, led to strongly green-emitting CDs with a 35 nm average diameter. Interestingly, the authors observed that under the same reaction conditions, glucose did not generate CDs. The authors proposed that polymerisation of glucosamine molecules followed by aromatisation via intramolecular dehydration, leads to a burst of nucleation when the aromatic cluster supersaturation is reached. This burst of nucleation takes place and the carbon nuclei grow to partially nanocrystalline CDs with certain hydrophilic functional groups in the surface. Raman, FTIR and XPS data confirmed the presence of aromatic amines, hydroxy and carboxy groups on the CD surface.

Subsequently, Liu et al. reported the hydrothermal synthesis of amino-functionalised green fluorescent CDs using glucosamine hydrochloride in the presence of excess sodium pyrophosphate ($Na_4P_2O_7$) (Scheme 15) [19]. The team showed that heating an aqueous mixture of the sugar and $Na_4P_2O_7$ for 10 h to 180 °C in a Teflon-lined autoclave resulted in green fluorescent N/P-doped CDs with QYs of up to 17% with an excitation independent emission which could be modulated by varying the concentration of $Na_4P_2O_7$ in the starting mixture. Also, the higher the concentration of pyrophosphate, the less aggregation product was observed. The resultant CDs were then effectively coupled to hyaluronate (a long-chain polymer containing repeating disaccharide units of glucuronate-β1->3-*N*-acetylglucosamine) stabilised gold nanoparticles (AuNPs) and used as a sensitive and selective probe to monitor hyaluronidase enzymatic activity, which is an enzyme that breaks down



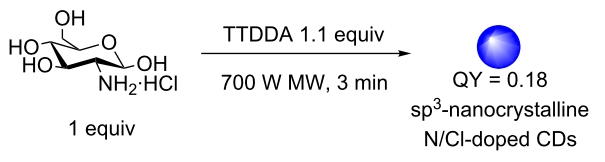
Scheme 14: Base-mediated synthesis of CDs with nanocrystalline cores, from fructose and maltose, without forcing reaction conditions.



Scheme 15: N/P-doped green-emissive CDs working in tandem with hyaluronic acid-coated AuNPs to monitor hyaluronidase activity.

hyaluronate (Scheme 15). As hyaluronate is used to stabilise the AuNPs, any enzymatic activity that degrades the polymer would result in AuNP aggregation, which in turn modulates the absorption properties of the AuNPs. The latter has a favourable overlap with the emission spectra of CDs, when stabilised. Hence a turn-on of the CD fluorescence is indicative of enzyme activity.

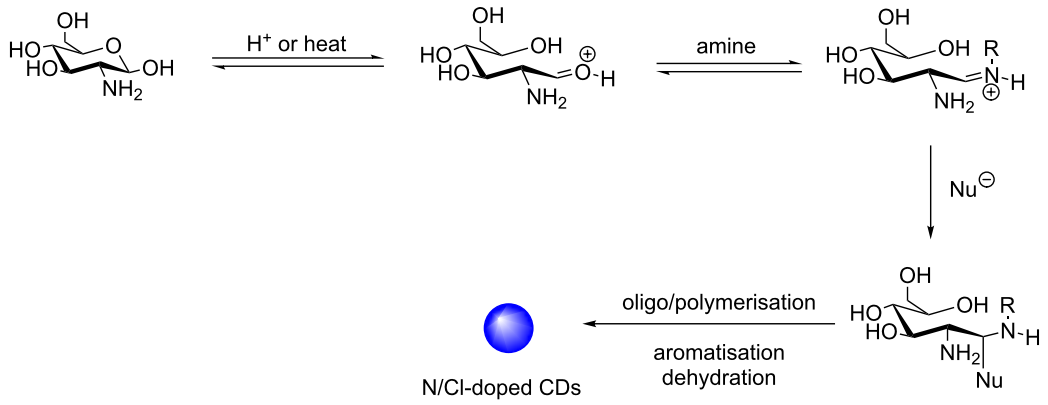
More recently, Galan et al. reported the 3 min one-step synthesis of blue-emitting CDs from glucosamine hydrochloride in the presence of TTDDA using microwave irradiation with QYs of up to 17% (Scheme 16) [46]. While most reported syntheses afford CDs with sp^2 crystalline or amorphous cores, the team showed that the resultant nanoparticles had an sp^3 nanocrystalline core, as determined by HRTEM and Raman spectroscopy. The authors attributed this observation to the relatively mild conditions used. They also showed that the presence of HCl was critical for the PL properties of the CD and that the formation of C–Cl bonds, as determined by Raman and FTIR spectroscopy, yielded the chlorine as a crucial auxochrome, which is in agreement to results previously reported by Kim et al. [44].



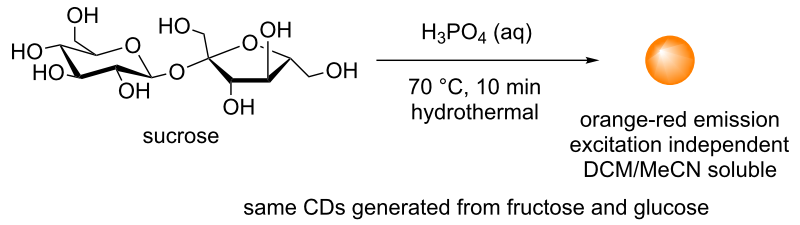
Scheme 16: Three-minute microwave synthesis of Cl/N-doped CDs from glucosamine hydrochloride and TTDDA to afford bottom-up synthesised CDs with an sp^3 nanocrystalline core.

Mechanism studies of the reaction by ^1H, ^{13}C, FTIR and React-IR helped to identify the key reaction intermediates (Scheme 17). The loss of the anomeric proton/carbon with formation of an aldehyde was observed within the first 90 seconds of the reaction, after which time, amide formation and sp^2-centre formation/aromatisation were also observed. React-IR studies under hydrothermal conditions, but at a lower temperature of 70 °C, helped the team to identify a reactive iminium species, which is formed from the reaction between the sugar aldehyde and an amine present in the reaction mixture, and is a key intermediate in the initial stages of nanoparticle formation. Trapping of the iminium electrophile could allow oligomer formation and dehydration, leading to the formation of the sp^3-enriched nanocrystalline core. In the second phase of the reaction, following the loss of bulk water, further carbonisation occurs and aromaticity is then generated on the outer layers of the core. Surface passivation by TTDDA can now take place via either incorporation of TTDDA into the surface heteroaromatics or amide bond formation. Amide formation can occur either through surface-bound carboxylic acids reacting directly with an amine (e.g., TTDDA or sugar-derived amine) or through the nucleophilic attack of an alcohol to the iminium electrophile, followed by rearrangement of the resulting imidate.

The work by Mandal et al. has also recently sought to provide some insights into nanoparticle formation and PL mechanism for sugar-derived CDs [47]. The team studied the reaction between sucrose and H_3PO_4 to afford excitation-independent orange-red emissive CDs (Ex = 365 nm), which were readily soluble in organic solvents such as DCM and MeCN (Scheme 18). Mechanistic investigations showed evidence of



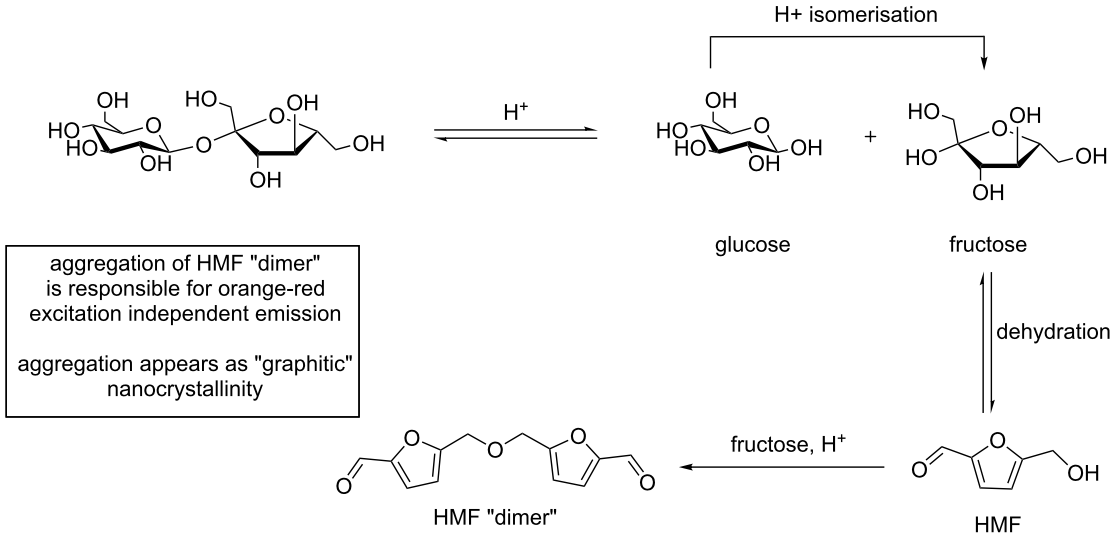
Scheme 17: Mechanism for the formation of N/Cl-doped CDs via key aldehyde and iminium intermediates, monitored by ^{1}H and ^{13}C NMR, FTIR and React-IR studies.



Scheme 18: Phosphoric acid-mediated synthesis of orange-red emissive CDs from sucrose.

hydroxymethylfurfural (HMF) derivatives as the major component in the preparation of this type of CD, as evidenced by ^{1}H, ^{13}C NMR, FTIR and MALDI–MS (Scheme 19).

The authors proposed that initial acid catalysed degradation of the sucrose disaccharide to its monosaccharide constituents fructose and glucose, followed by glucose isomerisation to fructose, leads to HMF formation following three dehydration steps. Indeed, HMF formation has been identified as a dehydration product in reactions with glucose, fructose and sucrose under acidic conditions [48]. Furthermore, the team was able to show that instead of polymeric furfural structures, HMF dimers are



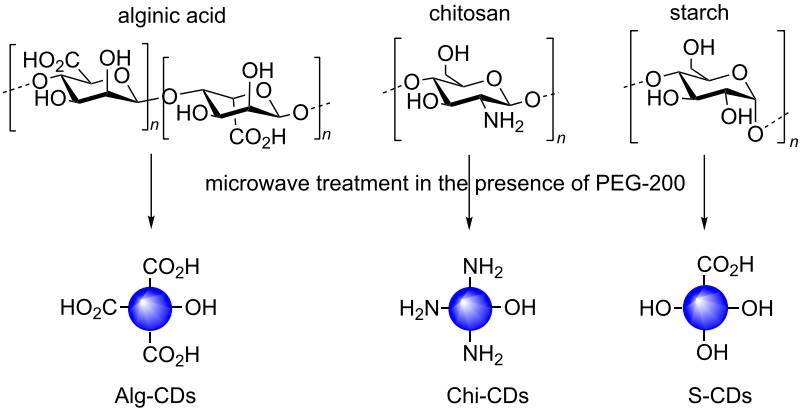
Scheme 19: Proposed HMF dimer, and its formation mechanism, that upon aggregations bestows orange-red emissive on sucrose-derived CDs.

produced via the acid-catalysed ether formation between HMF and fructose followed by subsequent dehydration, and undergo aggregation to form fluorescent CDs. Although, it is clear that small changes in the reaction conditions and reagents do have a significant effect in the final nanoparticle properties. These results provide evidence that aggregation of furfural intermediates or other heteroaromatic species could be responsible for the PL and physicochemical properties observed.

Fluorescent carbon dots synthesised from polysaccharides

Polysaccharides are essentially polymeric sugar molecules composed of monosaccharide units coupled together via glycosidic linkages to form long linear or branched chains. Some of the most common polysaccharides found in nature include cellulose, starch, glycogen or chitin [49]. Upon hydrolysis, these structures break down into smaller fragments such as oligosaccharides or monosaccharide units. Thus, it is unsurprising that these naturally occurring materials have also been used as CD precursors. Many CD syntheses report the use of biomass, particularly sourced from plant matter, which is essentially a huge source of naturally derived polysaccharides combined with smaller amounts of other organic molecules, e.g., amino acids, which can act as dopant agents. Some examples include the use of garlic [50], orange juice [51], onion waste [52] and general kitchen waste [53]. For the purpose of this review, we will concentrate on describing examples where defined and commercially available polysaccharides are used for the synthesis of FCDs and how these materials compare to CDs made using their monomeric counterparts.

Many different polysaccharides with different elemental composition and structural morphologies are available and as seen for monosaccharide-derived CDs, the different features and functional groups present in those distinct carbohydrate chains, will have an effect in the final properties of the CDs synthesised from them. Pramanik et al. exploited this hypothesis in the synthesis of CDs from three different polysaccharides: chitosan (Chi-CDs), alginic acid (Alg-CDs) and starch (S-CDs) in the presence of PEG-200 under identical microwave conditions (Scheme 20) [54]. TEM analysis of the samples highlighted that a range of morphologies and sizes were obtained depending on the polymer used. For example, S-CDs afforded the smallest particle size distribution (1–2 nm) but little morphological uniformity. On the other hand, Chi-CDs appear to have a distinctly spherical morphology with a size range of 2–10 nm and Alg-CDs also exhibit a distinct spherical morphology with a size range of 2–4 nm. Interestingly, an inverse correlation between the size of the CD and the fluorescence output was established, the smallest S-CDs gave the best fluorescence intensity of the three samples, while the largest Chi-CDs had the lowest. FTIR analysis provided evidence that the starting polysaccharide functional group composition is conferred onto the CDs. For instance, alginic acid has one carboxylic acid group per monomer unit, whereas chitosan is an amine-containing polysaccharide; analysis of the different CDs showed higher intensities for peaks attributed to carboxylic acid C=O bonds in both Alg-CDs and Chi-CDs. Similarly Chi-CDs showed an abundance of amine functionality, while S-CDs spectra had many more signals that could be assigned to alcohol groups and some carboxylic acid functionality, with the latter probably generated during the reaction. The authors further demonstrated the applicability of the different materials in heavy metal sensing. To that end, each CD sample was exposed to the same concentration (0.001 M) of divalent metal cations Cu^{2+}, Cd^{2+}, Sn^{2+} or Zn^{2+} in solution and the fluorescence response monitored. The starch-based S-CDs showed an interesting PL response, whereas the fluorescence output increased when in solution with all

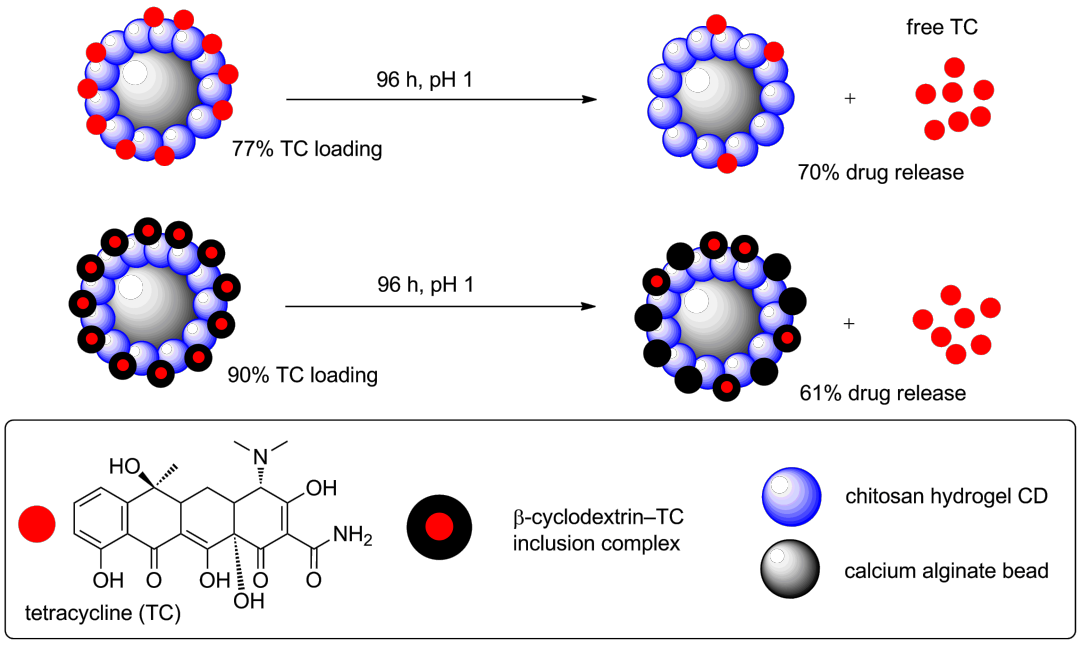


Scheme 20: Different polysaccharide-derived CDs in the presence of PEG-200 and how the starting material composition is conferred to the CD products.

metal ions tested, in the case of Cu^{2+}, a significant reduction in fluorescence was recorded. The authors proposed that Cu^{2+}, due to its paramagnetic nature, could quench the S-CD fluorescence via a photoinduced electron transfer mechanism, in which Cu^{2+} is reduced to Cu^{+}. The presence of Cu^{+} was confirmed by selected area electron diffraction (SAED) in the TEM analysis of the CD surface.

CDs from chitosan hydrogels have also been reported by Chowdhury et al. [55]. The hydrogels were synthesised from a mixture of acetic acid, glycerol and chitosan, as a more stable starting material for CDs. Microwave irradiation of the hydrogel yielded UV-blue emissive CDs with a range of sizes from 0.6–8.7 nm (as determined by DLS). Zeta-potential analysis yielded a value of +27 mV, indicative of an abundance of amino groups and as expected from an amino group containing chitosan starting material. In addition, the group also investigated CDs prepared from chitosan/Ag and chitosan/Au nanocomposites, which were incorporated while preparing the chitosan hydrogels. It was observed that although the emission of the new CDs was broad and less well defined, there was an enhancement in the PL emission for the Ag or Au-doped CDs. Subsequently in 2014, the same group was able to show that coating of calcium alginate (CA) beads with chitosan hydrogel-based CDs yielded a new nanomaterial that could be employed as a pH-responsive drug-delivery vehicle (Scheme 21) [56]. The CDs were used as a protective layer onto the CA beads and tetracycline (TC) was loaded onto the CD–CA beads. It was shown that a two-fold increase on drug loading was seen when compared to uncoated beads. Subsequently, TC release at a range of pH values was studied over a 96 h period and it was found that 70% of TC release takes place at low pH (pH 1) when compared to 36% release at pH 7 and 27% at pH 12. In order to improve the drug delivery profile of the complex, the authors developed a β-cyclodextrin/tetracycline (β-TC) host–guest inclusion complex, which allows a second "barrier" of release. The pre-formulated β-TC complex was loaded onto the CD–CA beads and not only were higher loading levels measured (90%), but also a slower rate of TC release at each pH value was recorded, as expected from a more stable drug/nanocomplex. The results reported here are good examples of the potential applications of amine-coated CDs as important components in drug-delivery applications.

Chitin, which is a cheap and readily available linear polysaccharide comprised of β-1,4-linked *N*-acetyl-D-glucosamine units, is the second most abundant biopolymer in nature and forms the backbone of crustaceans and insects exoskeleton and is also found in the cell wall of yeast and fungi [57]. Chitin is also the precursor of chitosan, which is formed by N-deacetylation to partially free amino groups, and is notoriously insoluble in water. Despite this fact, Shchipunov et al. demonstrated in 2015, the first hydrothermal synthesis of CDs derived from chitin in a Teflon-lined autoclave at 180 °C for 3 h in the presence of HNO_3, all in deionised water [58]. The CDs produced in this manner were purified from unreacted/unsolubilised



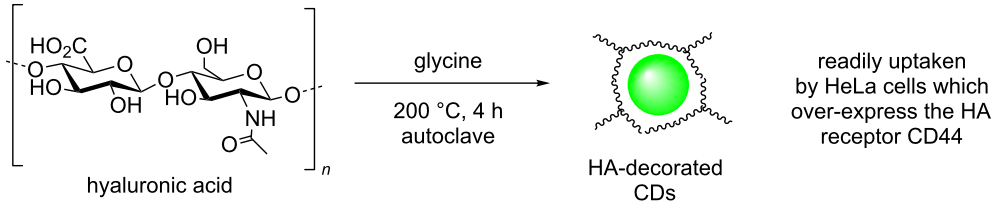
Scheme 21: Tetracycline release profiles for differentially-decorated CDs.

chitin via several filtration, centrifugation and dialysis steps. The N-doped CDs were blue-emitting under UV excitation with apparent long-term, bench stable fluorescence. These results might suggest further opportunities in the field for these type of less water soluble N-containing polysaccharides.

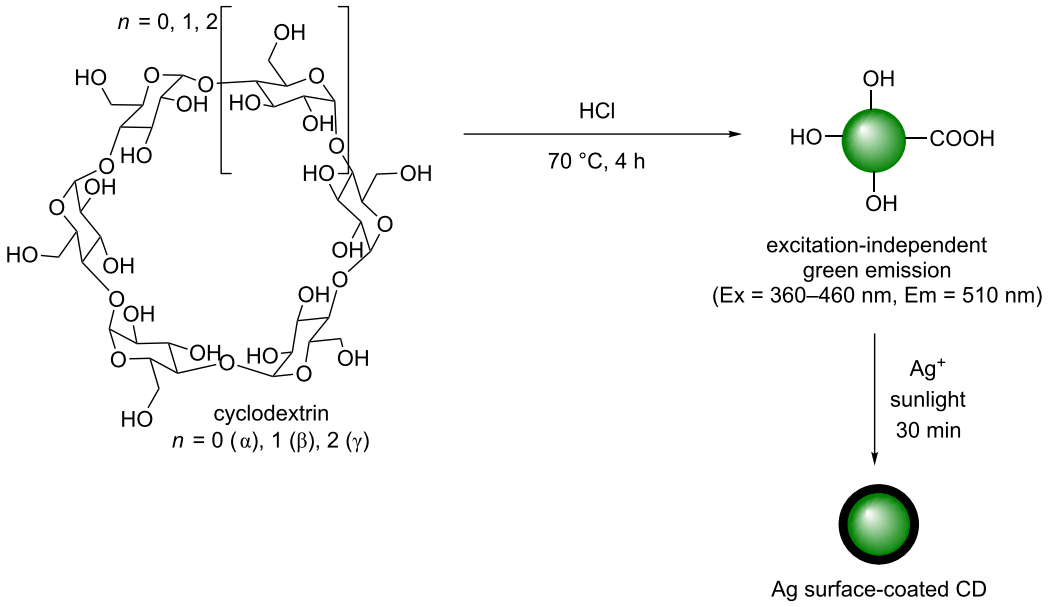
Hyaluronic acid is another N-containing polysaccharide composed of repeating dimeric units of glucoronic acid and *N*-acetyl-D-glucosamine units and which forms the core of complex proteoglycan aggregates found in the extracellular matrix [57]. The team of Du and Shao et al. reported the synthesis of N-doped hyaluronic acid-derived CDs and their application as drug delivery vectors [59]. Following standard hydrothermal synthetic procedures as previously described for other CD preparations, hyaluronate was heated in a Teflon-lined autoclave in the presence of glycine, which was found to be key, to 200 °C for 4 h to yield CDs of under 10 nm in size (Scheme 22). Structural analysis of the resultant CDs indicated the presence of carbonyl-containing functional groups such as carboxylic acids and amides, which coated a graphitic-type core. The nanoparticles exhibit excitation-dependent emission and were blue-emissive under UV excitation, but green when excited at 496 nm. Although no NMR characterisation was carried out on the samples, the authors proposed that due to the polymeric nature of the starting material, the resulting N-doped CD cores might be decorated by unreacted/fragmented hyaluronic acid (HA–CD). Subsequent cell feeding experiments with HA–CD with HeLa and U251 cells, revealed that upon internalisation the CDs where found to localise in the cytoplasm and particularly around the nucleus. Due to the large amounts of internalisation a receptor-mediated endocytosis was proposed. The particles were used as fluorescent probes to target CD44 high expression in tumour cells, opening the door for these types of polysaccharide-based nanomaterials in other targeted live cell labelling, imaging and drug-delivery applications.

In addition to amine containing polysaccharide, other neutral carbohydrate-based polymers have also been reported in the synthesis of CDs. Cyclodextrin is a cyclic glucose polymer that is commonly available in its α, β and γ forms, each corresponding to the number of glucose units (6, 7 and 8, respectively). In 2014, Wang et al. reported the synthesis of CDs from each of the different cyclodextrins via an acidic, hydrothermal treatment at 70 °C for 4 h (Scheme 23) [60]. Reaction of each type of cyclodextrin afforded quasi-spherical nanoparticles with a size range of 2.5 ± 0.8 nm and with an amorphous carbon core. The materials obtained had a range of alcohol and carbonyl-containing functionalities present on their surface. QYs were measured to range from 9% to 13%, which were dependant on the type of cyclodextrin utilised, with each CD showing a green emission under UV irradiation and excitation-independent emission from 360 to 460 nm excitation, which is typical of a uniform morphology. The authors proposed that the uniform emission could be attributed to either the uniform size distribution or the uniform surface state giving a single quantum dot-like emission profile. This interesting report highlights the fact that a reducing sugar is not essential to produce CDs and that under acidic and forcing conditions this type of starting materials can still undergo acetal hydrolysis, dehydration, aromatisation and carbonisation to yield CDs. The resulting cyclodextrin-derived CDs were then used for the detection of Ag^+ ions in solution. It was found initially that mixing $AgNO_3$ in an aqueous solution of CDs in sunlight resulted in the formal reduction of Ag^+ to elemental Ag^0, which was thought to proceed via the adhesion of Ag^+ to the CD surface, followed by reduction in the presence of sunlight, which promotes the excitation of the reducing electron to a higher energetic state (Scheme 23). Through UV–vis absorbance and TEM measurements it was evident that a surface layer of plasmonic Ag existed on their surface. The PL intensity of the CDs was modified in a linear manner with Ag^+ concentrations, and as such the nanoparticles could be utilised as a fluorescence probe to detect Ag^+ in solution up to concentrations of 0–25 μM.

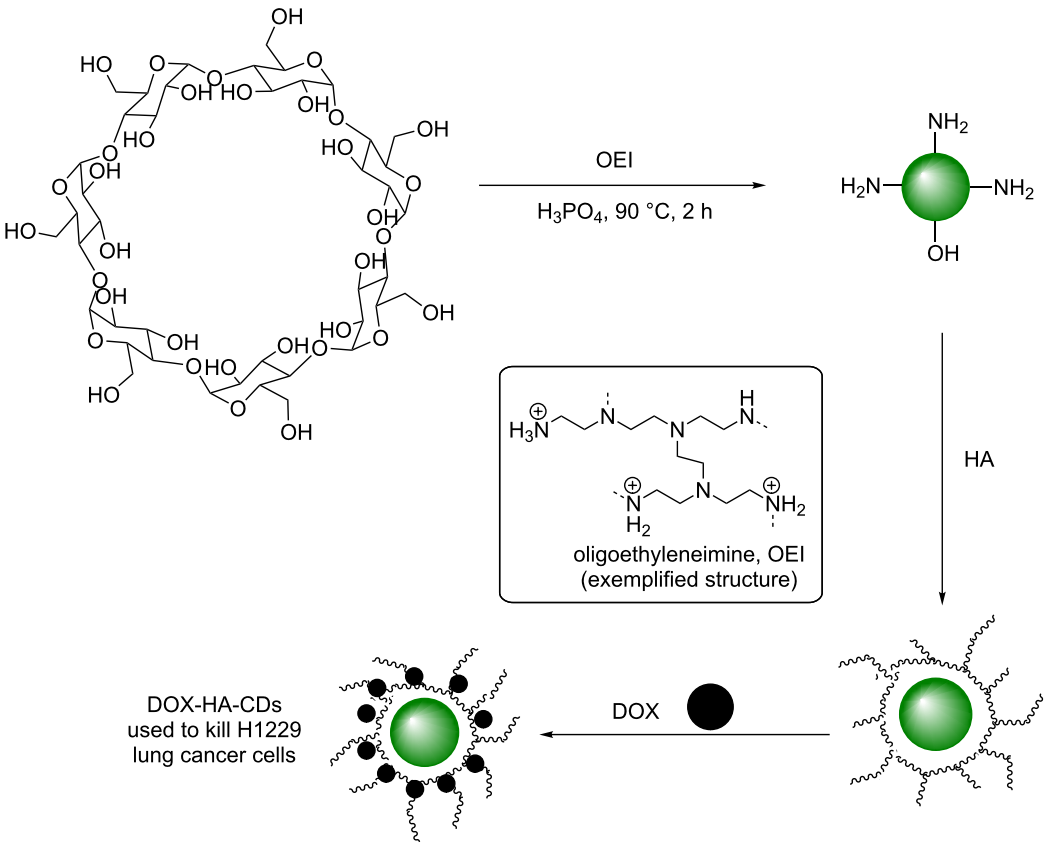
β-Cyclodextrin has also been utilised in the synthesis of CDs using a surface passivation and inorganic dehydration method. The groups of Yang and Teo et al. demonstrated that the synthesis of excitation-independent green emissive CDs could be achieved through the reaction of β-cyclodextrin in the presence of oligoethyleneimine (OEI) and phosphoric acid under thermolysis conditions (90 °C) for 2 h (Scheme 24) [61]. It was



Scheme 22: Hyaluronic acid (HA) and glycine-derived CDs, suspected to be decorated in unreacted HA, allowing receptor-mediated cell uptake.



Scheme 23: Cyclodextrin-derived CDs used for detection of Ag^+ ions in solution, based on the formal reduction of Ag^+ to afford plasmonic Ag.



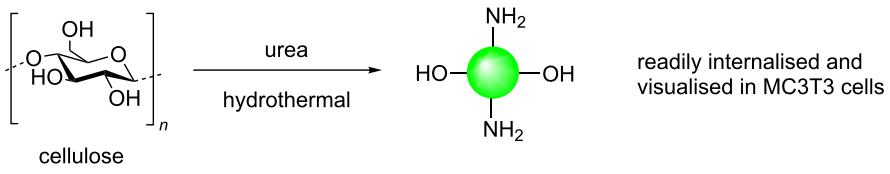
Scheme 24: Cyclodextrin and OEI-derived CDs, coated with hyaluronic acid and DOX, to produce an effective lung cancer cell drug-delivery vehicle.

demonstrated that the presence of phosphoric acid was crucial for the formation of fluorescent CDs, as the control reaction, in the absence of acid, did not produce emissive CDs. AFM and TEM indicated quasi-spherical CDs of 2–4 nm, while FTIR and XPS indicated that nitro groups were present within the CD structure. The green-emissive CDs were photostable at a wide range of pH values (1–13) and over long-exposure to excitation sources. These results indicated the advantages of inorganic-ion mediated dehydration and the use of N-doping via surface passivation to achieve QYs up to 30%. Due to the use of OEI, the CDs were positively charged as measured by ZP and as a result the novel nanomaterial could form nanocomplexes with negatively-charged polymers. To demonstrate their applicability, CDs were successfully decorated with hyaluronic acid (HA), a negatively-charged polysaccharide, and shown by AFM, DLS and TEM, to have formed nano-aggregates of up to 250 nm. Interestingly, the emissive properties of the CDs were unchanged upon complexation to HA. The resultant nano-aggregates were then loaded with doxorubicin (DOX) and a strong correlation between dose and cell death was demonstrated in lung cancer H1299 cells (Scheme 24).

Cellulose is the most abundant organic molecule on Earth and is a linear polysaccharide comprised of repeating β-1,4-linked glucose units. Similarly to cyclodextrin, cellulose does not contain N-functionalities, which has been shown to be crucial for superior PL properties in CDs. In order to exploit the abundance, renewable and cheap advantages offered by using cellulose as the starting material, an N-doping strategy is needed. The group of Yao described recently the formation of CDs from cellulose (the specific type is not defined in the original report) via hydrothermal treatment in the presence of urea (Scheme 25) [62]. The resultant CDs were blue-green emissive with excitation-dependent emission and a QY of up to 21%. The high QY and favourable fluorescence properties were, in-part, attributed to the presence of auxochromic N within the architecture of the CDs. Subsequent CD internalisation experiments in MC3T3 osteoblast cells indicated that after exposure times of up to 24 h, the cell viability was unchanged when using concentrations of CDs of up to 250 μg/mL. LCSM experiments showed that the CDs were readily internalised into the cells and could potentially find uses in drug-delivery applications.

Conclusion

FCDs have only been around for a little over a decade and yet, it has become clear that these novel fluorescent nanomaterials have tremendous potential in many applications such as metal sensing, photocatalysis and as probes for bioimaging and biomedical applications and they offer a cheaper and non-toxic alternative to other metal-based fluorescent nanomaterials, e.g., semiconductor QDs. In this review, we have described a number of synthetic approaches to access FCDs using mono-, oligo- and polysaccharides, as cheap and readily available starting materials and the data has been collated in Table 1. Methods described include thermal decomposition, chemical or hydrothermal oxidation under autoclave, ultrasonic or microwave-assisted conditions. The presence of defects in the CD structure has been proposed to be important with regards to their PL properties. Additionally, the use of surface passivating agents to provide uniform PL trapping sites on the CD surface and the introduction of electron-donating heteroatoms as dopant agents, have been shown to improve and help tune the PL properties of these interesting nanomaterials. Not one synthesis is the same, it has been made evident that small changes in the synthetic scheme employed to access CDs, have an impact on the final chemical and physical properties of the nanoparticles obtained (See Table 1). Thus, careful consideration needs to be given to the type of carbon source used (carbohydrates being inherently heterogeneous provide an abundant and cheap source to be explored among other materials), reagent ratios/concentration, presence or absence of dopant agent/s (N, P, S or B) and their sources, and type of chemical process employed. Although a full mechanism of CD formation has not been elucidated to date, initial mechanistic studies on the formation of CDs from carbohydrates, have suggested that carbohydrate ring-opening to the aldehyde, which can then react with available nucleophiles in the reaction mixture is key. Subsequent dehydration/aromatization events can take place, which lead to the production of N-heteroaromatic structures on the CD surface. The ability to tune the CD synthesis to produce different nanodots, offers unique opportunities and renders these materials amenable to a wide range of applications, as we have briefly described in this review. On the other hand, CD quantum yields are still lower in comparison to their direct competitors (semiconductor QDs) and efforts are currently being devoted to



Scheme 25: Cellulose and urea-derived N-doped CDs with green-emissive fluorescence.

Table 1: Summary of carbohydrate-derived CDs synthetic protocols and properties.

Carbohydrate	Heteroatom dopant/SPA[a]	Synthetic conditions	Fluorescence profile[b]	Size [nm]	Crystallinity	Principle functionality	Ref.
glucose	PEG-200	microwave	blue to green (Dep)	2–4	amorphous	C=C, C=O (acid), OH, C-O	[18]
glucosamine HCl	$Na_4P_2O_7$	Teflon-autoclave, reflux	green (Ind)	4	–	C=C, C=O, C-O, C-N, -OH, $-NH_2$	[19]
not specified	various	hydrothermal	blue to Red	various	various	various	[20]
glucose	TTDDA	H_2SO_4, HNO_3 reflux	blue (Dep)	5	graphitic	C=C, C=O (acid/amide), OH, C-O	[23]
glucose	KH_2PO_4	Teflon-autoclave, reflux	blue or green (Dep)	2–5	graphitic	C=C, C=O (acid), OH, C-O	[29]
glucose	phosphate	microwave	blue to green (Dep)	2	–	C=C, C=O (acid), OH, C-O	[31]
glucose	NH_3	ultrasonic	blue (Dep)	10	graphitic	C=C, C=O (acid), OH, N-aromatics, C-O	[32]
glucose	tryptophan	microwave	blue (Ind)	20	–	C=C, C=O (acid), OH, N-aromatics, C-O	[33]
glucosamine HCl	–	Teflon-autoclave, reflux	green (Ind)	30	crystalline	C=C, C=O, C-N, C-O, O-H	[34]
glucose	PEG-diamine	H^+/ultrasonic	blue (-)	10	–	C=C, C=O (acid/amide), OH, C-O, N-H	[36]
glucose	boric acid	Teflon-autoclave, reflux	blue (Dep)	3–5	–	C=C, C=O (acid), -OH, B-OH	[38]
glucose	glutathione	hydrothermal	blue to green (Dep)	2.5	amorphous	C=C, C=O, C-O, N-H, Oxidised S	[39]
glucose	EDA, conc. H_3PO_4	hydrothermal	blue to green (Dep)	10	hollow	C=C, C=O, C=N, -OH, P=O, P-C	[40]
glucose	NH_3, H_3PO_4	Teflon-autoclave, reflux	blue (Dep)	3	graphitic	C=C, C=O, P-C, P-N, P-O	[41]
glucose	TTDDA or dopamine	hydrothermal	blue or green (Dep)	2–7	crystalline	C=C, C=O, -OH, $-NH_2$	[42]
glycerol	TTDDA	microwave	blue to green (Dep)	3.5	amorphous	C=C, C=O (amide), -OH, $-NH_2$	[43]
xylitol	EDA, HCl	microwave	blue (Dep)	4–5	graphitic	C=C, C=O (amide), -OH, C-N, -Cl	[44]
fructose/maltose	–	NaOH/$NaHCO_3$, rt	green (Dep)	3–5	graphitic	C=C, C=O, C-O, -OH	[45]
glucosamine HCl	TTDDA	microwave	blue (Dep)	2–5	sp^3 crystalline	C=C, C=O (amide), C-O, C-N, C-Cl	[46]
sucrose	–	H_3PO_4, hydrothermal	orange-red (Ind) associated to HMF dimer aggregation	4	graphitic (molecular crystallinity)	C=C, C=O (acid), C-O, -OH	[47]
several polysaccharides	PEG-200	microwave	blue (Dep)	1–10 (substrate dependent)	–	substrate dependent	[54]
chitosan	glycerol, AcOH hydrogel	microwave	UV to blue (Dep)	1–8	–	C=C, $-NH_2$, C-O, -OH	[55,56]
chitin	–	HNO_3, Teflon-autoclave, reflux	blue (Dep)	4-8	graphitic	C=C, C=O (amide), $-NH_2$, -OH	[58]

Table 1: Summary of carbohydrate-derived CDs synthetic protocols and properties. (continued)

hyaluronic acid	glycine	Teflon-autoclave, reflux	blue (Dep)	2–4	graphitic	C=C, C=O (amide), -NH_2, C-O	[59]
cyclodextrin	–	HCl, hydrothermal	green (Ind)	2.5	amorphous	C=C, C=O (acid), C-O, -OH	[60]
cyclodextrin	OEI	hydrothermal	green (Ind)	2–4	–	C=C, C=O (anhydride, amide), C-O, -OH, -NH_2	[61]
cellulose	urea	hydrothermal	blue (Dep)	4	graphitic	C=C, C=O (amide), C-N, C-O, -NH_2, -OH	[62]

[a]Surface passivating agent, [b]major fluorescence emission range highlighted; Ind = excitation-independent emission, Dep = excitation-dependent emission.

improve their PL properties. As we gain a better understanding at the molecular level of the mechanism of photoluminescence and chemical formation of these exciting nanomaterials, we will be able to devise procedures to access designer materials for specific applications. It is clear that the future of this field is "CD" bright.

References

1. El-Boubbou, K.; Huang, X. *Curr. Med. Chem.* **2011,** *18,* 2060. doi:10.2174/092986711795656144
2. Katz, E.; Willner, I. *Angew. Chem., Int. Ed.* **2004,** *43,* 6042. doi:10.1002/anie.200400651
3. Marradi, M.; Martín-Lomas, M.; Penadés, S. *Adv. Carbohydr. Chem. Biochem.* **2010,** *64,* 211. doi:10.1016/S0065-2318(10)64005-X
4. Wu, C.; Chiu, D. T. *Angew. Chem., Int. Ed.* **2013,** *52,* 3086. doi:10.1002/anie.201205133
5. Vácha, R.; Martinez-Veracoechea, F. J.; Frenkel, D. *ACS Nano* **2012,** *6,* 10598. doi:10.1021/nn303508c
6. Canton, I.; Battaglia, G. *Chem. Soc. Rev.* **2012,** *41,* 2718. doi:10.1039/c2cs15309b
7. Reichardt, N. C.; Martin-Lomas, M.; Penadés, S. *Chem. Soc. Rev.* **2013,** *42,* 4358. doi:10.1039/c2cs35427f
8. Gao, X.; Cui, Y.; Levenson, R. M.; Chung, L. W. K.; Nie, S. *Nat. Biotechnol.* **2004,** *22,* 969. doi:10.1038/nbt994
9. Hou, B.; Benito-Alifonso, D.; Kattan, N.; Cherns, D.; Galan, M. C.; Fermin, D. J. *Chem. – Eur. J.* **2013,** *19,* 15847. doi:10.1002/chem.201302722
10. Benito-Alifonso, D.; Tremel, S.; Hou, B.; Lockyear, H.; Mantell, J.; Fermin, D. J.; Verkade, P.; Berry, M.; Galan, M. C. *Angew. Chem., Int. Ed.* **2014,** *53,* 810. doi:10.1002/anie.201307232
11. Hou, B.; Benito-Alifonso, D.; Webster, R.; Cherns, D.; Galan, M. C.; Fermin, D. J. *J. Mater. Chem. A* **2014,** *2,* 6879. doi:10.1039/C4TA00285G
12. Derfus, A. M.; Chan, W. C. W.; Bhatia, S. N. *Nano Lett.* **2004,** *4,* 11. doi:10.1021/nl0347334
13. Lim, S. Y.; Shen, W.; Gao, Z. *Chem. Soc. Rev.* **2015,** *44,* 362. doi:10.1039/C4CS00269E
14. Miao, P.; Han, K.; Tang, Y.; Wang, B.; Lin, T.; Cheng, W. *Nanoscale* **2015,** *7,* 1586. doi:10.1039/C4NR05712K
15. Baker, S. N.; Baker, G. A. *Angew. Chem., Int. Ed.* **2010,** *49,* 6726. doi:10.1002/anie.200906623
16. Xu, X.; Ray, R.; Gu, Y.; Ploehn, H. J.; Gearheart, L.; Raker, K.; Scrivens, W. A. *J. Am. Chem. Soc.* **2004,** *126,* 12736. doi:10.1021/ja040082h
17. Dong, Y.; Pang, H.; Yang, H. B.; Guo, C.; Shao, J.; Chi, Y.; Li, C. M.; Yu, T. *Angew. Chem., Int. Ed.* **2013,** *52,* 7800. doi:10.1002/anie.201301114
18. Zhu, H.; Wang, X.; Li, Y.; Wang, Z.; Yang, F.; Yang, X. *Chem. Commun.* **2009,** 5118. doi:10.1039/b907612c
19. Liu, S.; Zhao, N.; Cheng, Z.; Liu, H. *Nanoscale* **2015,** *7,* 6836. doi:10.1039/C5NR00070J
20. Bhunia, S. K.; Saha, A.; Maity, A. R.; Ray, S. C.; Jana, N. R. *Sci. Rep.* **2013,** *3,* 1473. doi:10.1038/srep01473
21. Hou, J.; Yan, J.; Zhao, Q.; Li, Y.; Ding, H.; Ding, L. *Nanoscale* **2013,** *5,* 9558. doi:10.1039/c3nr03444e
22. Pan, D.; Zhang, J.; Li, Z.; Zhang, Z.; Guo, L.; Wu, M. *J. Mater. Chem.* **2011,** *21,* 3565. doi:10.1039/c0jm03763j
23. Peng, H.; Travas-Sejdic, J. *Chem. Mater.* **2009,** *21,* 5563. doi:10.1021/cm901593y
24. Li, H.; Kang, Z.; Liu, Y.; Lee, S.-T. *J. Mater. Chem.* **2012,** *22,* 24230. doi:10.1039/c2jm34690g
25. Yang, Y.; Cui, J.; Zheng, M.; Hu, C.; Tan, S.; Xiao, Y.; Yang, Q.; Liu, Y. *Chem. Commun.* **2012,** *48,* 380. doi:10.1039/C1CC15678K
26. Klinger, K. M.; Liebner, F.; Fritz, I.; Potthast, A.; Rosenau, T. *J. Agric. Food Chem.* **2013,** *61,* 9004. doi:10.1021/jf4019596
27. Hwang, H.-I.; Hartman, T. G.; Rosen, R. T.; Lech, J.; Ho, C.-T. *J. Agric. Food Chem.* **1994,** *42,* 1000. doi:10.1021/jf00040a031
28. Cayuela, A.; Soriano, M. L.; Carrillo-Carrión, C.; Valcárcel, M. *Chem. Commun.* **2016,** *52,* 1311. doi:10.1039/C5CC07754K
29. Yang, Z.-C.; Wang, M.; Yong, A. M.; Wong, S. Y.; Zhang, X.-H.; Tan, H.; Chang, A. Y.; Li, X.; Wang, J. *Chem. Commun.* **2011,** *47,* 11615. doi:10.1039/c1cc14860e
30. Berezin, M. Y.; Achilefu, S. *Chem. Rev.* **2010,** *110,* 2641. doi:10.1021/cr900343z
31. Wang, X.; Qu, K.; Xu, B.; Ren, J.; Qu, X. *J. Mater. Chem.* **2011,** *21,* 2445. doi:10.1039/c0jm02963g
32. Ma, Z.; Ming, H.; Huang, H.; Liu, Y.; Kang, Z. *New J. Chem.* **2012,** *36,* 861. doi:10.1039/c2nj20942j

33. Simões, E. F. C.; da Silva, J. C. G. E.; Leitão, J. M. M. *Anal. Chim. Acta* **2014,** *852,* 174. doi:10.1016/j.aca.2014.08.050
34. Yang, Z.-C.; Li, X.; Wang, J. *Carbon* **2011,** *49,* 5207. doi:10.1016/j.carbon.2011.07.038
35. Pacher, P.; Beckman, J. S.; Liaudet, L. *Physiol. Rev.* **2007,** *87,* 315. doi:10.1152/physrev.00029.2006
36. Ajmal, M.; Yunus, U.; Matin, A.; Ul Haq, N. *J. Photochem. Photobiol., B* **2015,** *153,* 111. doi:10.1016/j.jphotobiol.2015.09.006
37. Gottesman, M. M.; Fojo, T.; Bates, S. E. *Nat. Rev. Cancer* **2002,** *2,* 48. doi:10.1038/nrc706
38. Wang, F.; Hao, Q.; Zhang, Y.; Xu, Y.; Lei, W. *Microchim. Acta* **2016,** *183,* 273. doi:10.1007/s00604-015-1650-1
39. Wang, C.; Xu, Z.; Cheng, H.; Lin, H.; Humphrey, M. G.; Zhang, C. *Carbon* **2015,** *82,* 87. doi:10.1016/j.carbon.2014.10.035
40. Gong, X.; Zhang, Q.; Gao, Y.; Shuang, S.; Choi, M. M. F.; Dong, C. *ACS Appl. Mater. Interfaces* **2016,** *8,* 11288. doi:10.1021/acsami.6b01577
41. Shi, B.; Su, Y.; Zhang, L.; Huang, M.; Liu, R.; Zhao, S. *ACS Appl. Mater. Interfaces* **2016,** *8,* 10717. doi:10.1021/acsami.6b01325
42. Peng, H.; Li, Y.; Jiang, C. L.; Luo, C. H.; Qi, R. J.; Huang, R.; Duan, C. G.; Travas-Sejdic, J. *Carbon* **2016,** *100,* 386. doi:10.1016/j.carbon.2016.01.029
43. Liu, C.; Zhang, P.; Tian, F.; Li, W.; Li, F.; Liu, W. *J. Mater. Chem.* **2011,** *21,* 13163. doi:10.1039/c1jm12744f
44. Kim, D.; Choi, Y.; Shin, E.; Jung, Y. K.; Kim, B.-S. *RSC Adv.* **2014,** *4,* 23210. doi:10.1039/c4ra01723d
45. Li, Y.; Zhong, X.; Rider, A. E.; Furman, S. A.; Ostrikov, K. *Green Chem.* **2014,** *16,* 2566. doi:10.1039/c3gc42562b
46. Hill, S. A.; Benito-Alifonso, D.; Morgan, D. J.; Davis, S. A.; Berry, M.; Galan, M. C. *Nanoscale* **2016,** *8,* 18630. doi:10.1039/C6NR07336K
47. Gude, V.; Das, A.; Chatterjee, T.; Mandal, P. K. *Phys. Chem. Chem. Phys.* **2016,** *18,* 28274. doi:10.1039/C6CP05321A
48. Vigier, K. D. O.; Benguerba, A.; Barrault, J.; Jérôme, F. *Green Chem.* **2012,** *14,* 285. doi:10.1039/C1GC16236E
49. Shi, L. *Int. J. Biol. Macromol.* **2016,** *92,* 37. doi:10.1016/j.ijbiomac.2016.06.100
50. Zhao, S.; Lan, M.; Zhu, X.; Xue, H.; Ng, T.-W.; Meng, X.; Lee, C.-S.; Wang, P.; Zhang, W. *ACS Appl. Mater. Interfaces* **2015,** *7,* 17054. doi:10.1021/acsami.5b03228
51. Sahu, S.; Behera, B.; Maiti, T. K.; Mohapatra, S. *Chem. Commun.* **2012,** *48,* 8835. doi:10.1039/c2cc33796g
52. Bandi, R.; Gangapuram, B. R.; Dadigala, R.; Eslavath, R.; Singh, S. S.; Guttena, V. *RSC Adv.* **2016,** *6,* 28633. doi:10.1039/C6RA01669C
53. Himaja, A. L.; Karthik, P. S.; Sreedhar, B.; Singh, S. P. *J. Fluoresc.* **2014,** *24,* 1767. doi:10.1007/s10895-014-1465-1
54. Chandra, S.; Pathan, S. H.; Mitra, S.; Modha, B. H.; Goswami, A.; Pramanik, P. *RSC Adv.* **2012,** *2,* 3602. doi:10.1039/c2ra00030j
55. Chowdhury, D.; Gogoi, N.; Majumdar, G. *RSC Adv.* **2012,** *2,* 12156. doi:10.1039/c2ra21705h
56. Gogoi, N.; Chowdhury, D. *J. Mater. Chem. B* **2014,** *2,* 4089. doi:10.1039/c3tb21835j
57. Rinaudo, M. *Prog. Polym. Sci.* **2006,** *31,* 603. doi:10.1016/j.progpolymsci.2006.06.001
58. Shchipunov, Y. A.; Khlebnikov, O. N.; Silant'ev, V. E. *Polym. Sci., Ser. B* **2015,** *57,* 16. doi:10.1134/S1560090415010121
59. Zhang, M.; Fang, Z.; Zhao, X.; Niu, Y.; Lou, J.; Zhao, L.; Wu, Y.; Zou, S.; Du, F.; Shao, Q. *RSC Adv.* **2016,** *6,* 104979. doi:10.1039/C6RA22210B
60. Hu, M.; Yang, Y.; Gu, X.; Hu, Y.; Huang, J.; Wang, C. *RSC Adv.* **2014,** *4,* 62446. doi:10.1039/C4RA11491D
61. Yang, C.; Thomsen, R. P.; Ogaki, R.; Kjems, J.; Teo, B. M. *J. Mater. Chem. B* **2015,** *3,* 4577. doi:10.1039/C5TB00467E
62. Shen, P.; Gao, J.; Cong, J.; Liu, Z.; Li, C.; Yao, J. *ChemistrySelect* **2016,** *1,* 1314. doi:10.1002/slct.201600216

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Glycoscience@Synchrotron: Synchrotron radiation applied to structural glycoscience

Serge Pérez*[1] and Daniele de Sanctis[2]

Review

Open Access

Address:
[1]Department of Molecular Pharmacochemistry, CNRS-University Grenoble Alpes, France and [2]ESRF – The European Synchrotron, Grenoble, France

Email:
Serge Pérez* - spsergeperez@gmail.com

* Corresponding author

Keywords:
antibodies; carbohydrate binding domains; cellulose; glycosaminoglycans; glycolipids; glycosyl hydrolases; glycosyl transferases; kinetic crystallography; lectins; polysaccharides; powder diffraction; small-angle X-ray scattering; starch; synchrotron radiation; transporters; X-ray crystallography

Beilstein J. Org. Chem. **2017,** *13,* 1145–1167.
doi:10.3762/bjoc.13.114

Received: 30 January 2017
Accepted: 17 May 2017
Published: 14 June 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Synchrotron radiation is the most versatile way to explore biological materials in different states: monocrystalline, polycrystalline, solution, colloids and multiscale architectures. Steady improvements in instrumentation have made synchrotrons the most flexible intense X-ray source. The wide range of applications of synchrotron radiation is commensurate with the structural diversity and complexity of the molecules and macromolecules that form the collection of substrates investigated by glycoscience. The present review illustrates how synchrotron-based experiments have contributed to our understanding in the field of structural glycobiology. Structural characterization of protein–carbohydrate interactions of the families of most glycan-interacting proteins (including glycosyl transferases and hydrolases, lectins, antibodies and GAG-binding proteins) are presented. Examples concerned with glycolipids and colloids are also covered as well as some dealing with the structures and multiscale architectures of polysaccharides. Insights into the kinetics of catalytic events observed in the crystalline state are also presented as well as some aspects of structure determination of protein in solution.

Introduction

Over the last decade, glycoscience has greatly benefited from the development of structural biology and the investigation of macromolecular structure and function relationships. Major contributions also came from considerable advances in high resolution NMR spectrometry and electron microscopy along with the continuous evolution of synchrotron radiation and free electron laser light sources. Since its discovery, X-ray radiation has been an invaluable tool to investigate the structure of matter. The range of wavelengths, in the region of an angstrom, and energies, extending over electronic shell levels, make them

the perfect probe to study material at the atomic scale. Nevertheless, the low availability and versatility of sources had for a long time represented a limitation on the use of X-rays for scientific applications. A major breakthrough came from the advent of synchrotron science. Over the years, they became an indispensable resource in the exploration of matter, thanks to the continuous spectrum of emitted radiation, the extremely high flux and brightness. Those features allow a wide range of experiments, spanning virtually all branches of sciences and technological applications, particularly those akin to nanoscience. Developments in neutron sources have paralleled those of synchrotron sources. Figure 1 summarizes the differences and the complementarity of the information that can be gathered from analyses performed with the respective sources. The synergistic use of both sources becomes particularly relevant when accurate hydrogen details are necessary.

Structural glycobiology gained recognition with the elucidation of glycosyl hydrolases mechanism by X-ray crystallography, but the scope of applications in glycobiology is much broader: it encompasses the range of glycan containing (macro)-molecules and their conjugates. The present article reviews the application of synchrotron radiation to some key areas of glycoscience potentially of interest to the growing number of non-specialist users.

Structural characterization of protein–carbohydrate interactions are covered as well as some involving glycolipids and colloids and the structure and architecture of polysaccharides. Insights into the kinetics of catalytic events occurring in the crystalline state are also described as well as some aspects of the determination of structure of proteins in solution.

Review

Synchrotron radiation

Synchrotrons are particle accelerators in which charged particles circulate along a closed path. Storage rings are a particular kind of synchrotron in which the charged particles, usually electrons, are accelerated to speeds close to c, the speed of light, and kept orbiting at constant energy (Figure 2). In practice, the terms synchrotron and storage ring are often used interchangeably. The application of magnetic fields induces curvature in the trajectories of the particles, which lose energy by emitting electromagnetic radiation, known as synchrotron light. The electrons are forced to deviate from a straight trajectory either by bending magnets that present a constant dipolar magnetic field and ensure the closing of the orbit, or by insertion devices, such as undulators. Undulators are a much more efficient way to produce X-ray beams and force electrons along an oscillating path in the horizontal plane (Figure 3). In this manner, the X-ray emitted at one oscillation is in phase with the

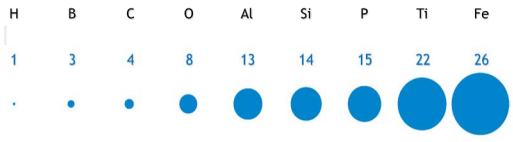
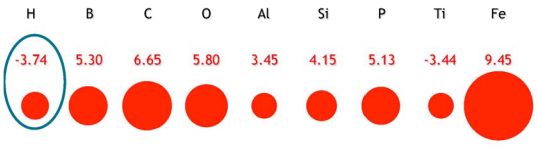
Synchrotron Radiation	Neutrons
Light beam (electromagnetic wave)	Particle beam (neutral subatomic particle)
Interactions with the electrons surrounding the nuclei (in the sample)	Interactions with the nuclei and the magnetic moment of unpaired electrons (in the sample)
Mainly scattered by heavy elements (Scattering proportional to Z)	Scattered by all elements, also the light ones like the hydrogen isotopes (Scattering not proportional to Z)
	
Small penetration depth (surface studies)	Deep penetration depth (bulk studies)
Very intense beam measuring small or ultra-dilute samples	Less intense beam measuring larger samples
Applications • Protein-crystal structures • Fast chemical reactions • Surface studies (defects, corrosion) • High-energy spectroscopy e.g. measurements of electron energy-levels	**Applications** • Magnetic structures & excitations • Organic structures using the H-D isotope effect • Bulk studies (strains, excitations) • Low-energy spectroscopy e.g. molecular vibrations

Figure 1: Complementarity of synchrotron radiation and neutron sources to investigate the structure of matter.

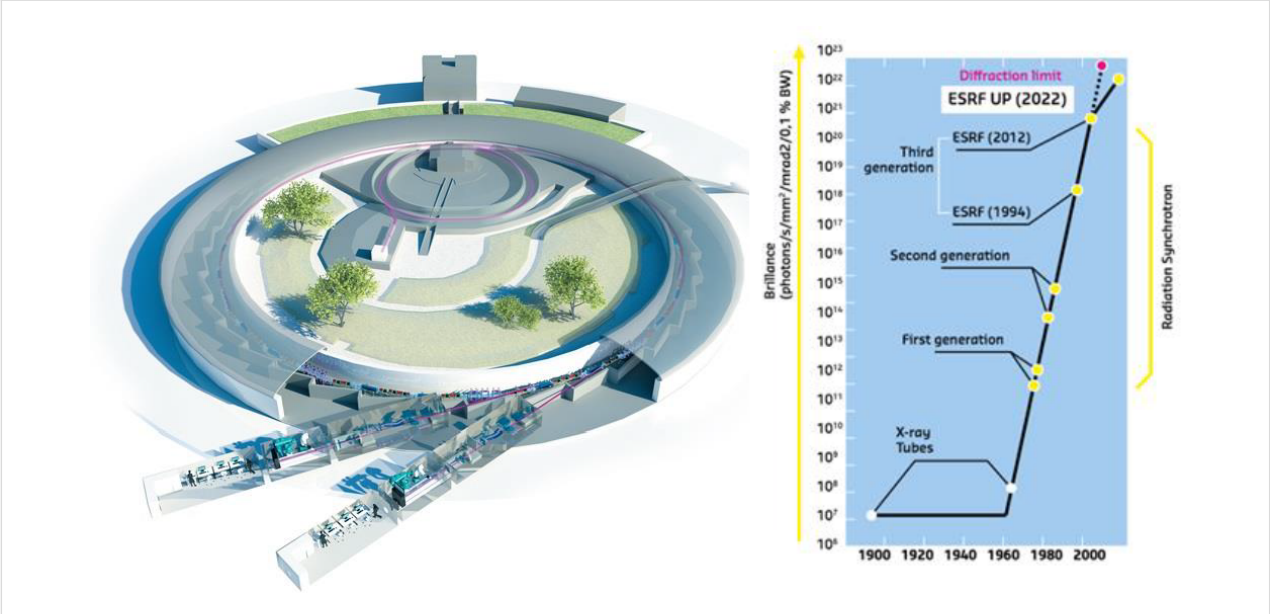


Figure 2: A representation of a synchrotron storage ring, including linear accelerator, booster and two beamlines (left) and the increase in X-ray brilliance since the first X-ray tubes to the current ESRF configuration and the predicted next generation after the machine upgrade planned in 2020. Credit: S. Gerlier/ESRF with permission.

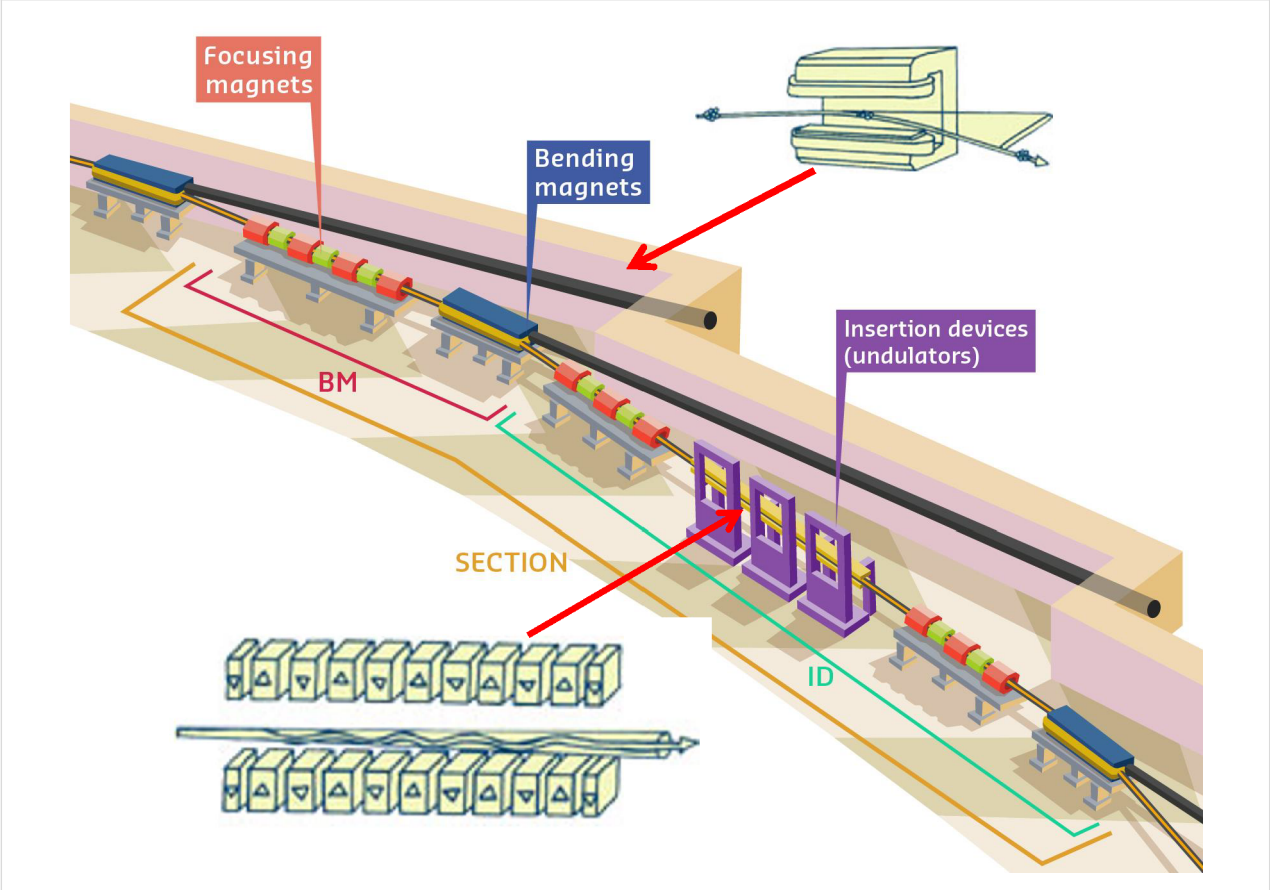


Figure 3: Schematic representation of a sector of a storage ring. Bending magnets and insertion devices are alternated. Bending magnets emit X-rays over a large angular range (top right) and are responsible for maintaining the closed trajectory in the storage ring. Insertion devices such as undulators (bottom left) produce X-rays with higher brilliance, which propagate along the electron beam. Credit: S. Gerlier/ESRF with permission.

radiation from the following oscillations, resulting in an intrinsic higher brilliance. The improvements in insertion devices have made storage rings the most versatile intense X-ray sources, and many storage rings have been constructed around the world and planning for the construction and commissioning of a new generation of storage rings is under way [1].

In the quantum mechanics wave-particle duality, X-rays produced by a synchrotron can be regarded as a linearly polarized electromagnetic plane wave or as photons with energy given by Planck's law. An X-ray photon that interacts with an atom can either be scattered or absorbed. Scattering that occurs with the same momentum (where there is no change in wavelength between scattered and incident waves) is called elastic or Thomson scattering. This is not generally the case, as an incident photon can transfer part of its energy to the electron and is scattered at a lower frequency by a phenomenon known as Compton scattering.

Photoelectric absorption occurs instead when an atom absorbs an X-ray photon. The excess energy is transferred to an electron, which is expelled and the atom is ionized. When the incident photon has an energy above the atomic K shell (so called K-edge energy), it expels an electron from the inner shell and creates a hole, which is eventually filled by an electron decaying from an outer shell. The difference in energy is emitted as a photon of energy equal to the difference of the two atomic shells. This effect is known as X-ray fluorescence and the photon energy provides a unique fingerprint of that atom. Moreover, modulation in absorption around the edge reflects the local structure of the material [2]. Photoelectric absorption, besides depending on the energy, varies with the Z atomic number (approximately proportional to Z^4). This phenomenon produces the contrast that is used in X-ray imaging techniques.

The synchrotron light spectrum is polychromatic, with a spectral bandwidth that depends on the type and configuration of the sources. Many experiments require a narrower bandwidth, or a monochromatic beam, and this is produced using a perfect crystal, a monochromator, and the desired wavelength is selected by changing the angle of the incident beam, in accordance with Bragg's law [3]. The monochromatic beam is then focused by using, for example, a system of X-ray mirrors. Consequently, an X-ray beam of the desired size and shape is delivered to the sample position. Continuous development of focusing systems has led to the use of beam sizes as small as a few nanometers. This has allowed the study of smaller samples with an enhanced signal to noise and higher spatial resolution [4-7].

The tunability of the wavelength to reach the values that are optimum for a given experiment provides the most powerful way to determine the three-dimensional features of macromolecular structures. Diffraction performed at an energy close to a heavy element absorption edge produces a resonant effect for which scattered waves are reemitted with a phase delay, inducing small variations in the diffraction intensities. The differences in the intensities can be used to determine the position of the heavier atoms and ultimately the electron density map of the macromolecule's structure. This effect is known as multiwavelength anomalous diffraction (MAD) and today, with its single-wavelength variant (SAD), it is the most successful and widely used techniques to determine the 3D structure of complex systems such as biological molecules, which can be composed of thousands of atoms [8,9].

Molecular structures

As early as 1930, the first crystal structures of organic compounds to be investigated were carbohydrates of low molecular weight. Over the following years, only eight additional crystal structures were reported. The determination of the three dimensional structure of the dehydrated form of sucrose, in 1947, was considered a significant contribution to the field. A major breakthrough occurred in 1951, when Bijvoet confirmed, without ambiguity, the D-configuration of glucose, which had been assigned from indirect reasoning by Emil Fischer in 1891 [10]. At the present time, the Cambridge Structural Database contains a few thousand entries for carbohydrate crystal structures, among which a limited number of molecules are relevant to glycobiology.

With the exception of sucrose and cyclic compounds, such as cyclodextrins or cyclo-amyloses, carbohydrates are reluctant to crystallize in form and size suitable for X-ray crystallographic analysis. This is even more pronounced for compounds having molecular weights ranging from 1000 to 5000 Da. Among the reasons, there is the difficulty to produce sufficient amount of material or the intrinsic occurrence of molecular disorder in solution, where several forms coexists (linear, five and six membered rings, anomeric mixture, etc.). It is also true that much less effort has been devoted to the production of organic crystals of medium sized biomolecules compared to biological macromolecules. Nevertheless, in many instances, ordered samples may be obtained, either in the form of molecular crystals of micrometric dimensions or in the form of polycrystalline materials.

Small molecule crystals

In the quest to solve the crystal structures of cello-oligosaccharides, as model compounds of cellulose, several attempts to grow crystals of β-D-cellotetraose of a size suitable for X-ray

diffraction had failed. Despite many attempts, the best crystals ever obtained for cellotriose and cellotetraose were very thin laths having dimensions of only about 10 µm in thickness. In the case of cellotetraose, single crystals as small as 0.40 × 0.15 × 0.015 mm could be processed by an X-ray synchrotron beam and 3800 independent reflections were collected. The molecular and crystal structure was solved using molecular replacement methods, and refined to an R factor of 0.048 [11]. Those results were useful in the elucidation of the crystalline structure of cellulose.

The family of resin glycosides offers another example of difficulty in terms of single crystal growth. Glycolipids (or lipo-oligosaccharides) comprise a carbohydrate moiety covalently linked to a lipid that confers on them an amphiphilic character, which makes them reluctant to crystallize. One member of the family is tricolorin A (L-rhamnopyranosyl (1->3) α-L-rhamnopyranosyl (1->2) β-D-glucopyransyl (1->2) β-D-fucopyranoid linked to japinoli acid forming a 19-membered ring macrocyclic ester, extracted from *Convolvulaceous* species which have been used in traditional medicine throughout the world since ancient times. Small crystals, with dimensions of 0.5 × 0.01 × 0.01 mm, could be grown using protein crystallization methods. Data were collected using synchrotron radiation, and the structure was solved using direct methods. Four independent molecules were found in each asymmetric unit (which contains 284 non-hydrogen atoms) in a highly hydrated unit cell (Figure 4) [12].

Polycrystalline material

Powder diffraction is a standard technique in material science that is used to investigate polycrystalline materials as many micrometer-sized crystals instead of a large single crystal. A powder diffraction pattern captures all possible crystal orientations simultaneously. The development of synchrotron radiation instrumentation dedicated to powder diffraction [13] allows to perform the experiments that were considered to be impractical before.

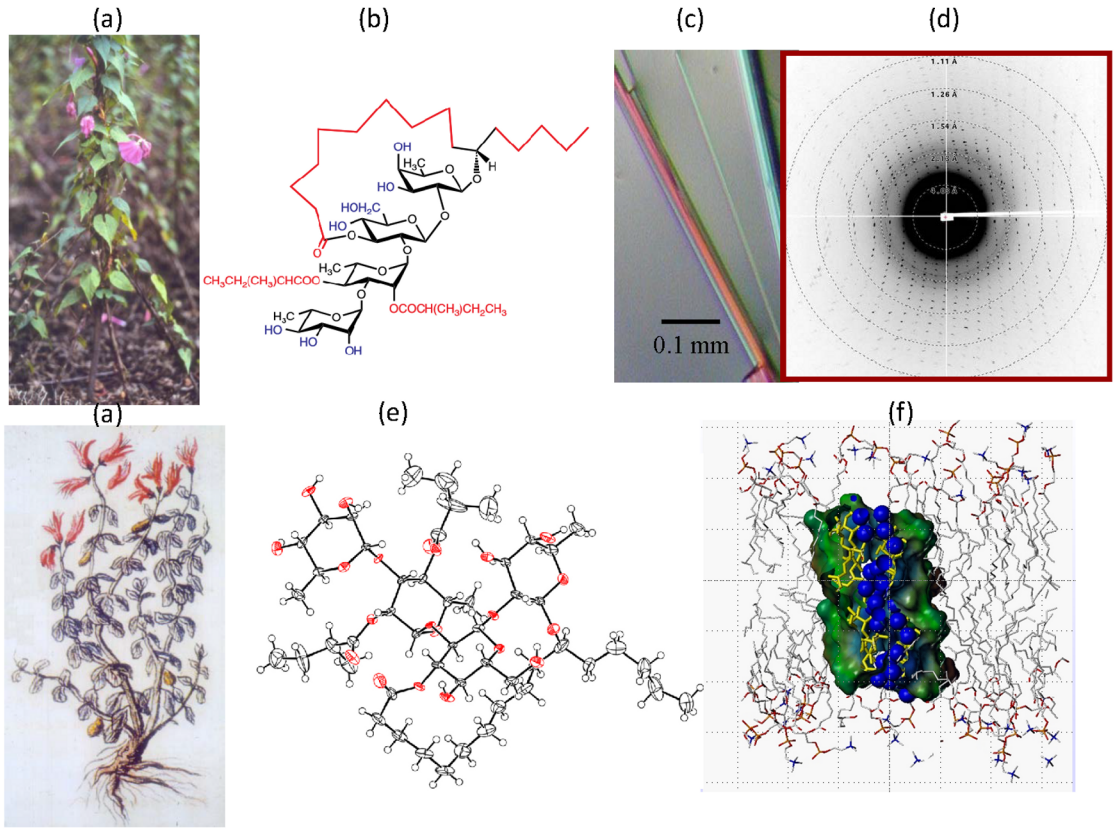


Figure 4: Structural features of the resin glycoside tricolorin A. (a) Extracted from the Mexican variety of the morning glory plant *Impomea tricolor Cav*. (b) Chemical structure of tricolorin A (L-rhamnopyranosyl (1->3)-α-L-rhamnopyranosyl-(1->2)-β-D-glucopyranosyl-(1->2)-β-D-fucopyranoside linked to japinoli acid. (c) Single crystals of tricolorin A. (d) Diffractogram from X-ray synchrotron. (e) Molecular structure of one, out of the four crystallographically independent molecules in the unit cell. (f) Molecular modelling of the insertion of tricolorin A within a fluid phospholipid bilayer [12].

When using high quality data in conjunction with advanced computational methods, it is possible to solve and refine crystals structures of small organic molecules with limited torsional freedom. This approach is less powerful than single crystal diffraction because of a loss of information by reducing the 3D space on a 1D spectra. Nevertheless, the resolution of the crystalline structure of a synthetic pentasaccharide from heparin, illustrates the potential of this technique. From the experimentally recorded X-ray powder diffractogram (Figure 5), the unit cell dimensions and the space group were determined. The process was continued with a computational building of the pentasaccharide and a simulated annealing procedure in direct space to locate the molecule in the unit cell. Once the carbohydrate backbone was positioned, the refinement continued by an adjustment of the rotations of the glycosidic linkages and side chains. The final construction and model completion provided the crystal and molecular structure with a high confidence factor [14].

Recently remarkable examples of protein structures determined by using this technique were also reported [15-17]. The clear advantage over single crystal diffraction is the easier preparation of the crystalline samples. As a result of improvement in the technique's high resolution, new possibilities exist, such as the investigations of the occurrence of phase transitions in large macromolecules as a function of temperature.

Macromolecular structures

X-ray diffraction with synchrotron radiation is the most powerful method for revealing the three-dimensional structure of biological macromolecules. Among the 128,000 structures deposited in the Protein Data Bank (January 2017) more than 80% have been measured and solved at synchrotron radiation facilities [18].

Macromolecular crystallography beamlines underwent a constant evolution over last decade that had a dramatic impact on the throughput and on the complexity of the structures determined. However, despite the development in nano-volume liquid handling for high-throughput screens, the crystallization of biological macromolecules still represents an important bottleneck in structure determination. Nanoliter handling devices allow the screening of hundreds of crystallization conditions even with a limited amount of sample of a few tens of microliters [19]. Furthermore, a successful example of automation in crystal harvesting were recently reported [20],

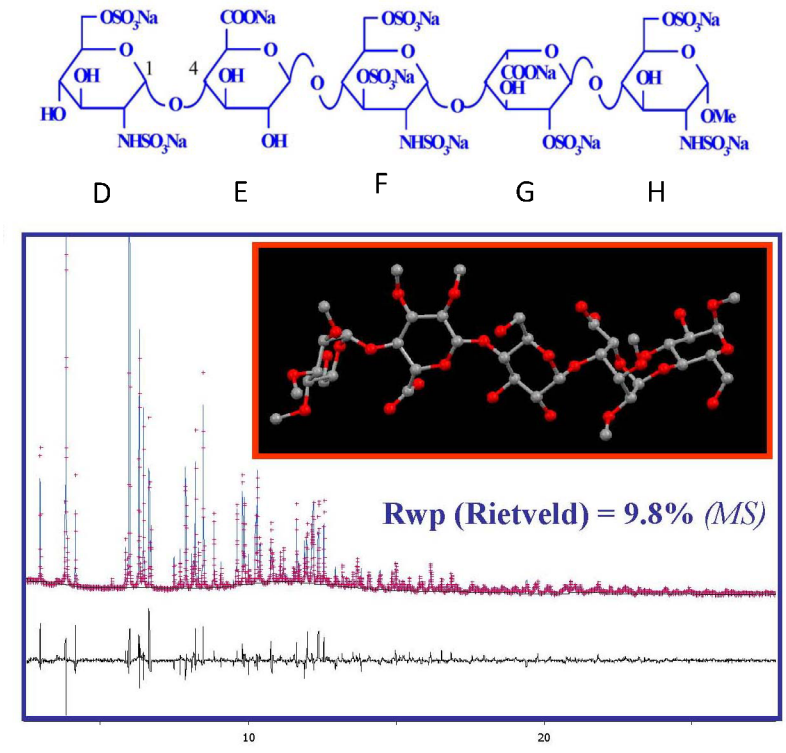


Figure 5: Powder diffractogram measured on a synthetic pentasaccharide from heparin, at ESRF beamline ID31, λ: 0.8 Å). The unit cell constant and the space group symmetry were assigned to: a = 15.54; b = 8.83, c = 17.67, β = 94.6; Monoclinic, $P2_1$. A 3-dimensional model of the structure in the unit cell was obtained using a molecular model where the sulfated idose residue was kept in $^1C_4$ conformation [14]. (Courtesy Drs J. Kieffer and Philippe Ochsenbein with permission).

while robots are now used to handle cryo-cooled samples at most synchrotron sources. Automation allows for reliable sample exchange and the evaluation of hundreds of samples per day. The development in pixel array detector technology has reduced the time for data collection to minutes or less and significantly improved the quality of the data thanks to no read-out noise and point spread across the pixels. Furthermore the advent of microfocus [21] and microbeam beamlines [22,23] completely dedicated to macromolecular crystallography permits diffraction data collection from smaller samples, of the order of a few micrometers. By matching the X-ray beam to the crystal size, it maximizes the diffraction signal-to-noise and reduces background scattering from crystal holder and mother liquor. New beamline graphical control software [24-26] facilitates beamline operation without exposing the complexity of the hardware. This allows the implementation of elaborate experiments even to users that are less familiar with computational tools. Beamline control software is interfaced with a laboratory information management system (LIMS), a metadata management system [27]. It is used to track samples, record experiment details and report experimental protocols and results from automatic post-experiment data processing protocols [28]. The synergy among these components has recently given rise to completely automated data collection experiments [29].

Glycoproteins

In recent years, the expression and production of recombinant proteins was of great benefit to the whole structural biology community, with more than 85% of the protein structures deposited in the Protein Data Bank being expressed in *Escherichia coli*. However, many proteins require post-translational modifications for correct biological activity and it is estimated that more than 50% of all human proteins are glycosylated, whereas proteins expressed in *E. coli* do not contain any glycan chains. For proteins that require post-translational modification, eukaryotic expression systems are usually preferred [30].

The crystallization of glycoproteins faces several obstacles, including the micro-heterogenity of glycans at the surface of the protein. For a given glycoprotein, there may exist considerable variations of N-linked glycan chains from protein to protein. Such a heterogeneous macromolecular mix is not suitable for crystal formation. Large post-translational modifications also have the effect of increasing surface entropy and hinder crystal packing. For this reason, it is sometime necessary to manipulate the glycoform to facilitate the crystallization. In the case of the human IgE-FcεRIα [31], Man5-GlcNAc-GlcNAc-Asp-linked glycoforms produced better crystals than in the case where only the Man-GlcNAc-GlcNAc-Asp form was present. There are other cases where crystallization may be facilitated by the presence of glycans that form stabilizing intermolecular contacts within the crystal. Platforms for the expression and crystallization of glycoproteins are available and can typically be successful in a few weeks [32].

Nevertheless, in the large majority of glycosylated structures, only the electron density map of the initial N-linked GlcNAc is present and can be modelled. In most of the cases, the glycan chains are exposed to the solvent and highly flexible. In such instances, the glycan can be modelled only up to the last visible residue (Figure 6).

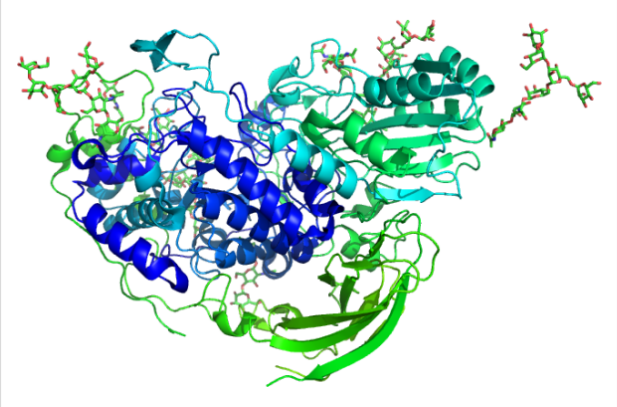


Figure 6: Three dimensional ribbon representation of a heavily N-glycosylated *Aspergilllus* sp. Family GH3 β-D-glucosidase protein (PDB 5FJI) [33].

The rise in the deposition of glycosylated protein structures reinforced the need for appropriate model restraints for model building and refinement crystallographic software. Model refinement without correct restraints will nearly always result in distortion and particular caution should be given to crystallographic reports where there is a wrong linkage distance specification or a mistaken anomer and handedness. Automated detection, building and validation of sugar models starting from X-ray diffraction data are being implemented [34].

Carbohydrate interacting proteins

The carbohydrate-mediated recognition events that have a high biological relevance give a pivotal role to the study of protein–carbohydrate interactions. Those interactions drive several distinct biological events, going from the enzymes involved in the biosynthesis, to the hydrolysis and modifications. Transporters and proteins purely involved in recognition (lectin, antibodies, carbohydrate binding modules, glycosaminoglycan binding proteins) are the other important classes of carbohydrate-binding proteins. Figure 7 shows the evolution of the number of carbohydrate interacting proteins that have been solved over the last 25 years, with a particular emphasis on the number of structures determined at high resolution.

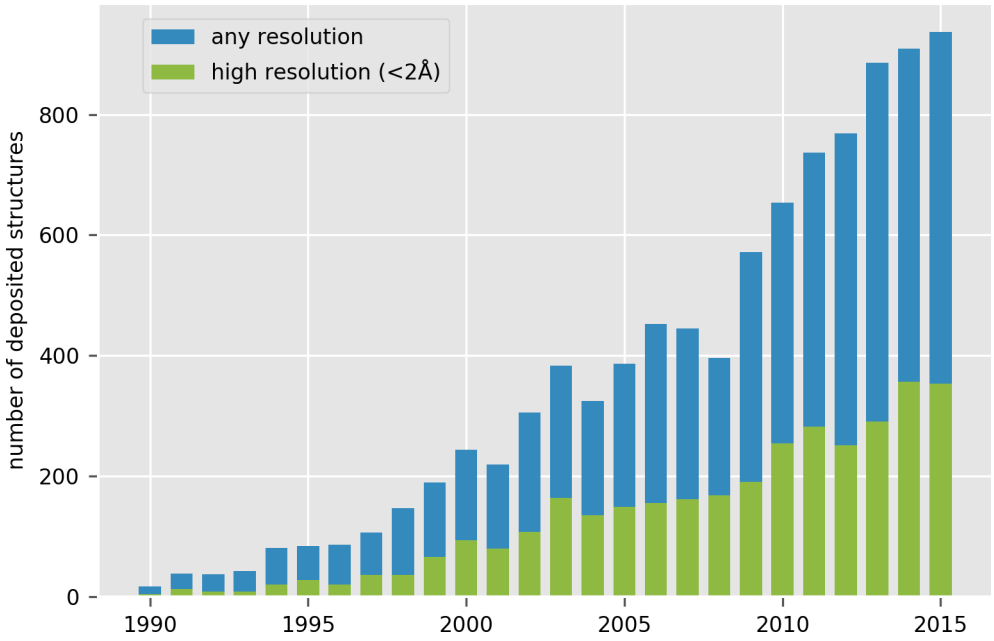


Figure 7: Histogram of the number of deposited crystal structures of glycan-binding proteins deposited over the years. Structures being resolved at high resolution (<2 Å) are displayed in green. (Courtesy Dr. J. Hendrickx with permission).

Glycosyl transferases: The biosynthesis of oligosaccharides is performed by a ubiquitous class of enzymes: the glycosyl transferases (GTs). The catalytic mechanism underlying the biosynthesis of glycosidic linkage requires the transfer of a sugar residue from a donor to an acceptor [35]. Acceptor substrates are carbohydrates, proteins, lipids, DNA, flavonol, antibiotics and steroids. In contrast, glycosyl donor substrates are mostly sugar nucleotides, such as UDP-GlcNAc, UDP-Gal, GDP-Man, and the GTs that process them are often referred to as Leloir enzymes. In certain cases, lipid-linked sugars, e.g., dolichol phosphate saccharides and unsubstituted phosphates are also utilized. The transfer of saccharides by GTs is regio-specific and stereo-specific: depending on the anomeric configuration of the transferred saccharide, two possible stereo-chemical outcomes occur, either inversion or retention. Based on the CAZy classification, the number of GT families amounts to 90, in a context where sequence homology is low http://www.cazy.org) [36]. The increased number of sequenced genomes is paralleled by an increasing number of accession entries for the GTs crystal structures in the PDB, which amounts to 900. Unlike glycoside hydrolases which display a large variety of different folds, the structures of GTs solved today can be clustered in two types of folds (and variants of these folds), namely GT-A and GT-B (Figure 8). Different folds are nevertheless observed for GTs that use lipid phosphate donor substrates. The achievement of the enzyme-transition

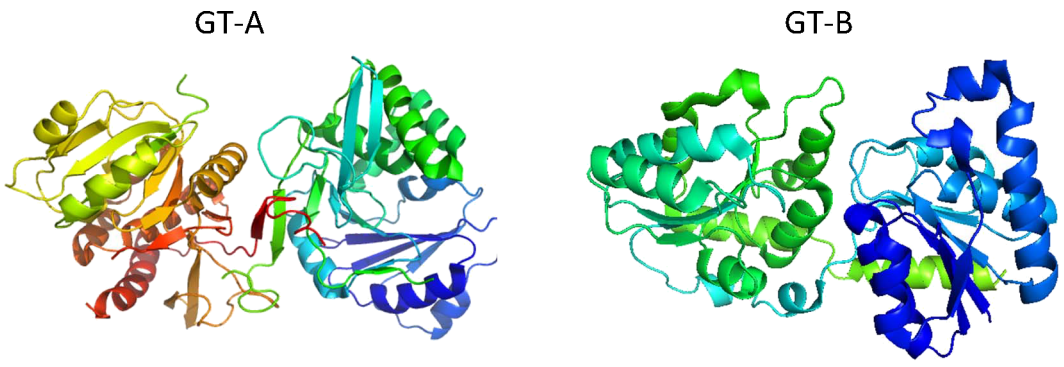


Figure 8: Ribbon diagram representations of prototypical members of the GT-A and GT-B super-family fold, respectively. PDB 1OMZ [38] and PDB 1NLM [39].

state complex requires a particular arrangement of the active site that is the result of concomitant protein dynamics, plasticity of GTs and conformation changes that allow for substrate recognition and catalysis [37].

In plants, GTs are also involved in the biosynthesis of hemicellulose. Xyloglucan is one of the main hemicellulose components in the cell walls of dicots. Its biosynthesis involves different GTs, including a fucosyltransferase, FUT1 that belongs to the glycosyltransferase family 37. The determination of the crystal structure revealed yet another variant of a GT-B fold and could explain FUT1 substrate specificity (Figure 9). Furthermore, the determination in complex with a minimal xyloglucan oligosaccharide acceptor and GDP lead to the understanding of the FUT1 mechanism [40].

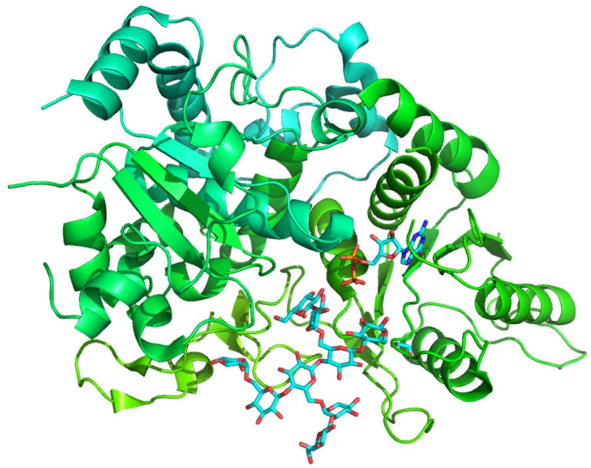


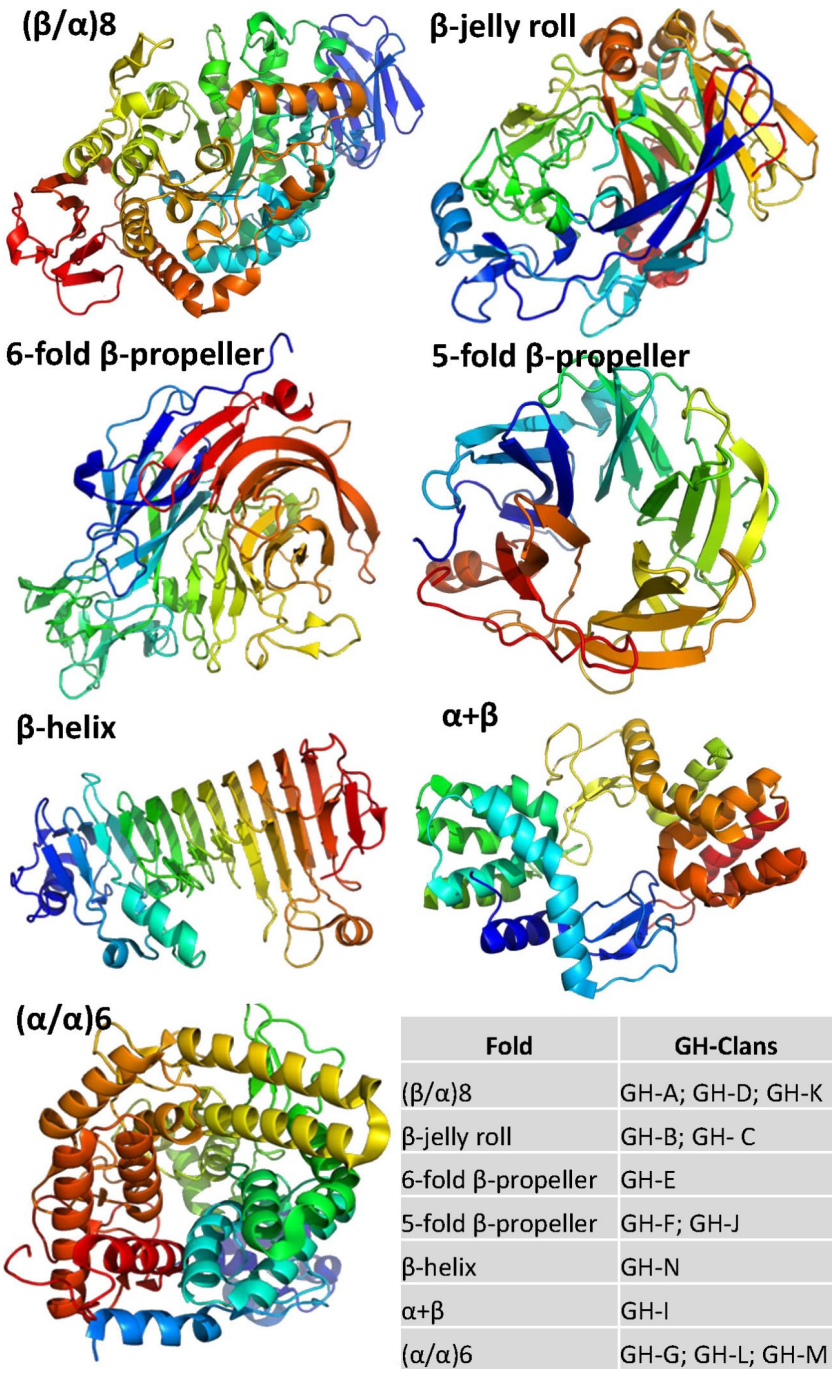
Figure 9: Representation of the FUT1 structure determined in complex with the acceptor (carbon atoms in green) and with end product GDP (carbon atoms in yellow) (PDB 5KOR) [40].

Carbohydrate esterases: Carbohydrate esterases perform the de-O or de-N-acylation of carbohydrates. From a mechanistic point of view, this family of enzymes is divided into two classes, according to the dual role played by the carbohydrate. One class is exemplified by the pectin methyl esterase in which case the carbohydrate plays the role of the "acid". In another class, the carbohydrate acts as an alcohol, as in acetylated xylan. A classification based on amino acid sequence similarities has been proposed yielding 16 families [41]. Among the 100 crystal structures which have been solved, 30 were obtained in complex with carbohydrates, mainly pectic oligosaccharides.

Polysaccharide lyases: Polysaccharide lyases (PLs) constitute a family of enzymes that cleave uronic acid-containing polysaccharide chains. The underlying mechanism is a β-elimination mechanism which generates an unsaturated hexenuronic acid residue and a new reducing end of the polysaccharide. At the present time, the reported number of crystal structures amounts to 190, among which 64 are complexed with carbohydrate ligands. These enzymes show a large variety of folds. Based on amino acid sequence similarities, polysaccharide lyases have been classified in 24 families [41].

Glycoside hydrolases: The hydrolysis of carbohydrates is the result of the action of a wide spread group of enzymes: the glycosyl hydrolases (GHs). GHs cleave the glycosidic linkage between two or more carbohydrates or between a carbohydrate and a non-carbohydrate moiety. They can catalyse the hydrolysis of O-, N-, S-linked glycosides, as well. The catalytic event can occur either in the middle (-endo) or at the end (-exo) of the substrate. The hydrolysis of the glycosidic bond implies a general acid (proton donor) and nucleophile/base, and involves two amino acid residues of the enzyme. Depending upon the position of these catalytic residues with respect to the substrate cleavable bond, the outcome of the reaction is either an inversion (inverting mechanism) or a retention (retaining mechanism) of the anomeric configuration. At the present time, about 4500 crystal structures of GHs have been deposited in the PDB. Approximately 30% of them are complexed with carbohydrate ligands. A classification of GH (more than 100 families) has been established, first based on amino acid sequences similarities and further consolidated by the availability of 3D dimensional structures [42]. The analysis of the GH structures present in the CAZy database helped not only to decipher the hydrolytic mechanism, but also reveal the evolutionary relationships between these enzymes. An extended classification based on the fold of the proteins, allowed the identification of 14 main clans (Figure 10).

Carbohydrate binding modules: Carbohydrate binding modules (CBM) are defined as a sequence of contiguous amino acids within a carbohydrate-active enzyme with a discrete fold having carbohydrate-binding activity. Initially CBMs were classified as a cellulose-binding domain, but their occurrence in other carbohydrate active enzymes required a dedicated classification, separate from other non-catalytic proteins, and similiar to lectins, antibodies and sugar-transport molecules. Depositions in the PDB for CBMs amount to 900. In the CAZy database, CBMs are classified within 80 families based on amino acid sequence similarities, while a three-dimensional structural classification clusters CBM into seven fold families [43]. The most represented fold is the β-sandwich comprised of two β-sheets, each consisting of three to six anti-parallel β-strands. As a large proportion of crystal structures are complexed with carbohydrates (from monosaccharides to oligosaccharides),



Fold	GH-Clans
(β/α)8	GH-A; GH-D; GH-K
β-jelly roll	GH-B; GH- C
6-fold β-propeller	GH-E
5-fold β-propeller	GH-F; GH-J
β-helix	GH-N
α+β	GH-I
(α/α)6	GH-G; GH-L; GH-M

Figure 10: Representation of the seven folds most commonly found in glycoside hydrolases. From the classification of glycoside hydrolases into more than 100 families, a hierarchical clustering into 14 clans has been proposed based on similarities of folds [42]. Some folds are common to different clans.

three CBM types have been classified based on their sugar recognition modes: surface binders, ”endo-type” binders and “exo-type” binders [44].

Lectins: Lectins constitute a unique and diverse family of proteins that reversibly bind monosaccharides and oligosaccharides, with utmost specificity, without displaying any catalytic or immunological activity. At the present time, the number of crystal structures of lectins deposited in the PDB amounts to about 1,500. Interestingly, about 60% of them were obtained ligated to carbohydrates, which range from monosaccharides to 10 residue-long oligosaccharides. Lectins occur in plants, animals, algae, bacteria, fungi and yeasts, and viruses. Their involvement in key biologically-important recognition pro-

cesses is well documented, as in the case of embryogenesis, fertilization, inflammation and metastasis. Lectins play a key role in parasite-symbiotic recognition in microbes and invertebrates of plants and vertebrates. The present role assigned to lectin lies in their ability to decipher sugar-encoded information, i.e., they are a molecular reader of the glyco-code.

The plethora of three-dimensional structures of lectins, both in unbound form or complexed with oligosaccharides, lead to their organization in a dedicated database, available at http://glyco3d.cermav.cnrs.fr [45]. Information contained in the database provided description of the main features of this important class of proteins. Lectins exhibit a variable oligomeric assembly that ranges from mono- to deca-valency (Figure 11).

Lectins present sugar-binding sites that are in most cases relatively shallow, and are located near the surface and therefore accessible to solvent. One or two calcium ions, identified in several lectin families of different origins, are involved in the carbohydrate binding by direct coordination to the sugar hydroxy groups. The comparison of detailed conformational features of oligosaccharides and their modes of interaction with the protein led to the development of different molecular modelling methods.

A somehow indirect application of the fine specificity of the binding of oligosaccharides to lectins has been elegantly developed to solve the phase problem in protein crystallography. Selenium-labelled carbohydrates can bind to the combining site of lectins at relatively low concentration, and provide sufficient

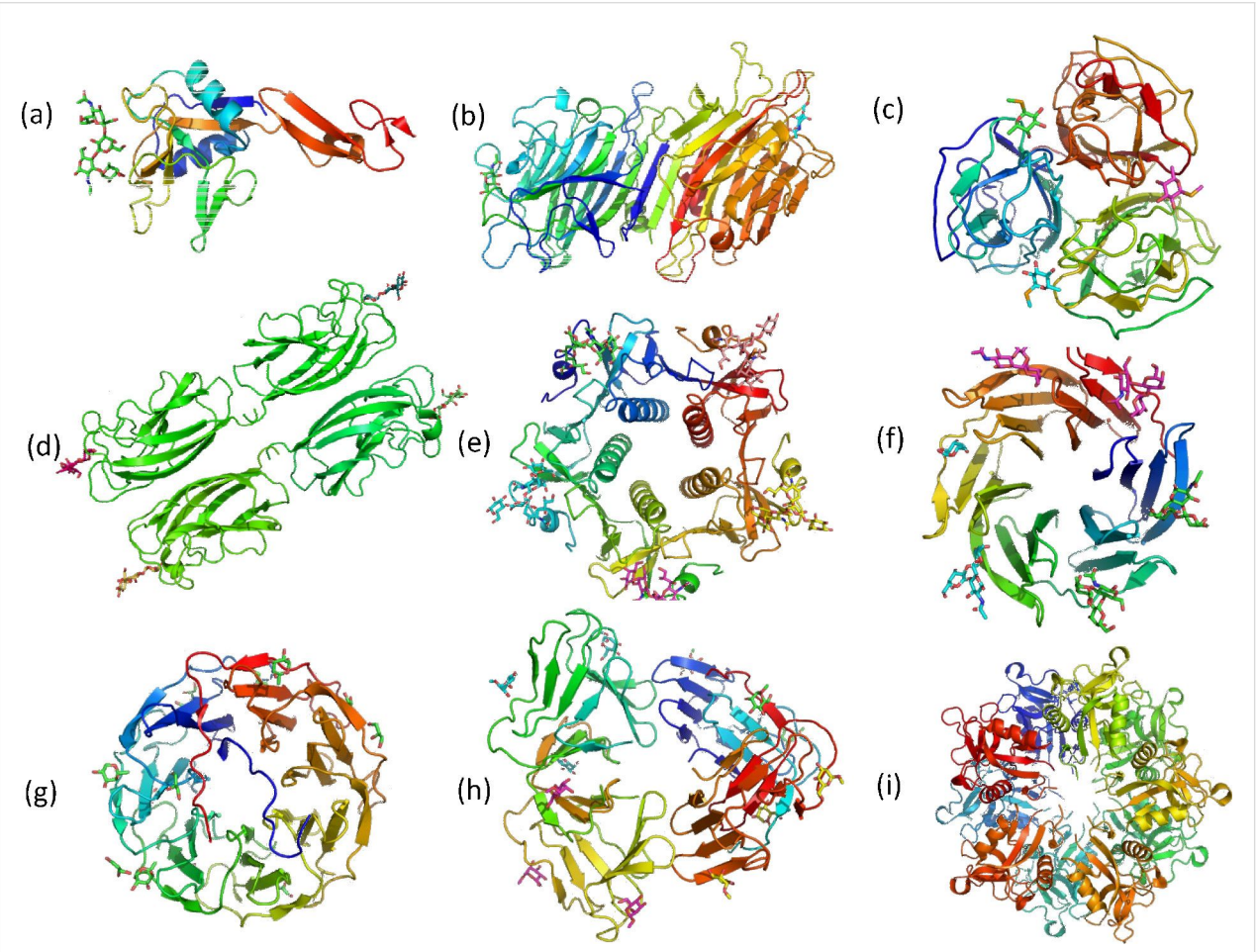


Figure 11: The multivalent carbohydrate binding features of lectins from X-ray structures. (a) Monovalent. E-selectin with bound sialyl LewisX: NeuAc α2->3 Gal β1->4 (Fuc α1->3) GlcNAc (PDB: 1G1T) [46]. (b) Divalent. *Dolichos bifluorus* seed lectin in complex with the blood group A trisaccharide (PDB: 1LU2) [47]. (c) Trivalent. N-terminal domain of BC2L-C lectin from *Burkholderia cenocepacia* with specificity for fucosylated human histo-blood group antigens (PDB: 2WQ4) [48]. (d) Tetravalent. *Pseudomonas aeruginosa* II lectin complexed to iso-globoside Gal α1->3 Gal β1->4 Glc (PDB: 2VXJ) [49]. (e) Pentavalent. Cholera toxin B subunit bound tGM1 pentasacharide: Gal β1->3 GalNAc β1->4 (Neu5Ac α2->3) Gal β1->4 Glc (PDB: 3CHB) [50]. (f) Hexavalent. *Burkholderia Ambifaria* lectin (BambL) complexed with H type2 trisaccharide, Fuc α1->2 Gal β1->4 GlcNAc (PDB: 3ZZV) [23]. (g) Heptavalent. Lectin from *Photorhabdus luminescens* complexed to L-fucose. PDB: 5C9P) [51]. (h) Octavalent. Lectin from *Galanthus nivalis* complexed with Me α-D-Man (PDB: 1MSA) [52]. (i) Decavalent. C-type lectin from *Bothrops jararacussu* (PDB: 5F2Q) [53].

anomalous signals for MAD or SAD methods of phasing to work, as was exemplified by the structure solution of the F17-G fibrial adhesion [54]. This elegant approach was used to elucidate the crystal structure of *Ralstonia solanacearum* lectin [55], *Parkia platycephala* lectin [56] and *Psathyrella velutina* lectin [57].

Anti-carbohydrate antibodies: Carbohydrate determinants are expressed on the cell surface through glycoproteins and glycolipids where they are exposed to a wide range of contexts, surroundings and surface densities. It is within such a landscape that antibodies recognize carbohydrate determinants. Data from many systems have shown that the minimum epitopes are often found at the extremity of the determinant. As a result, the presentation of the carbohydrate on the target cell may be such that antibodies with similar specificity exhibit different selective cell-profiling. Up to now, crystallographic studies of carbohydrate-antibodies mainly concentrated on systems where carbohydrates are complexed with antibody (Fab) or variable fragments (Fv). The organization of a small database of high-resolution three-dimensional structures of carbohydrate–antibody complexes [45] provides a way to classify the different types of bindings. Antibodies that recognize a terminal carbohydrate motif present a cavity-like binding feature, while a groove-like binding site is found in antibodies that bind to internal carbohydrate motifs. This is a mechanism typically found in bacterial polysaccharides. There is an occurrence of very large cavities which are open at both ends. The "side-on" entry of the antigen is often at the origin of the occurrence of conformational antigens.

Glycosaminoglycan–protein complexes: As members of the proteoglycan family, glycosaminoglycans (GAGs) are linear polysaccharides, constituted by 40 to 100 repeating disaccharide units, which are usually found to be linked to core proteins. These polysaccharides are components of the peri- and extracellular matrix and are present on surfaces or close to surfaces of animal cells. Based on their core repeat disaccharide units, glycosaminoglycans are classified in four groups: the heparin/heparin group; the chondroitin/dermatan sulfate group, the keratin sulfate group, and the hyaluronic group. With the exception of the later, many sources of structural micro-heterogeneities occur as the epimerization at the C-5 position of uronic acids, and N- and O-sulfation. Of paramount interest is the elucidation of the role of GAGs in their interactions with such important proteins as extracellular matrix proteins, chemokines, growth factors, complement proteins, enzymes, and viruses [58].

The three-dimensional structures of proteins co-crystallized in interaction with GAG fragments have been organized in a database (http://glyco3d.cermav.cnrs.fr/). Because of their relevance for pharmaceutical application, most of these fragments are heparin oligosaccharides. Crystallization of proteins in complex with GAGs is very difficult because of the high degree of heterogeneity and intrinsic flexibility of GAGs. The crystal structure of a fragment as long as a hexadecasaccharide could be co-crystallized as complexed with thrombin and antithrombin at 2.5 Å resolution. For the time being, this is one of the largest oligosaccharide structures ever established throughout macromolecular X-ray crystallography (Figure 12) [59].

Transporters: Soluble sugars serve many purposes in complex organisms. Their cellular exchange relies on transport proteins that are responsible for uptake or release. To date, three main families of eukaryotic transporters have been identified GLUTs, SGLTs, and SWEETs – the most recently discovered sugar transport family, which is responsible for cellular export. In mammals, 14 monosaccharide transport proteins GLUTs are responsible for the diffusion of glucose, galactose, fructose, urate, myoinositol, and dehydroascorbic acid. SGLTs are sodium-glucose symporters that couple the transport of glucose to sodium ions. SWEETs have been characterized the most recently. Major carbohydrate transporters mediate an active uptake and efflux of various mono- and disaccharides. The low affinity of these proteins for sugars seems to be a characteristic feature of transporters involved in high turnover rates, rather than a highly specific transport at low levels of substrates. The structure of the first transporter to be determined was the one of lactose permease LacY [60]. Later on, the structures of different bacterial homologues were also solved. It is only recently that the structure of human GLUT1 was reported [61]. Nevertheless, the joint difficulty to solubilize and crystallize membrane proteins, explains the paucity of crystal structures deposited in the database [62-72]. This is even worse for those proteins involved in the transport of sugars (Figure 13)

Kinetic crystallography

Since the biological activity of many proteins is preserved in the crystalline state, the possibility to investigate the dynamic process of their mechanisms is absolutely intriguing. In kinetic crystallography, a biological reaction is initiated in the crystal and the fates of the transient species formed are followed by determining the structural changes. Depending upon the time scale of the reaction and the scheme used for its initiation, time-resolved crystallography requires either the use of fast diffraction techniques such as Laue diffraction (polychromatic beam), or the capture of intermediates by trapping methods. These trapping strategies require the complementary use of UV–visible single-crystal spectroscopy. Providing extreme care is taken to avoid artefacts, these methods are in principle available to a

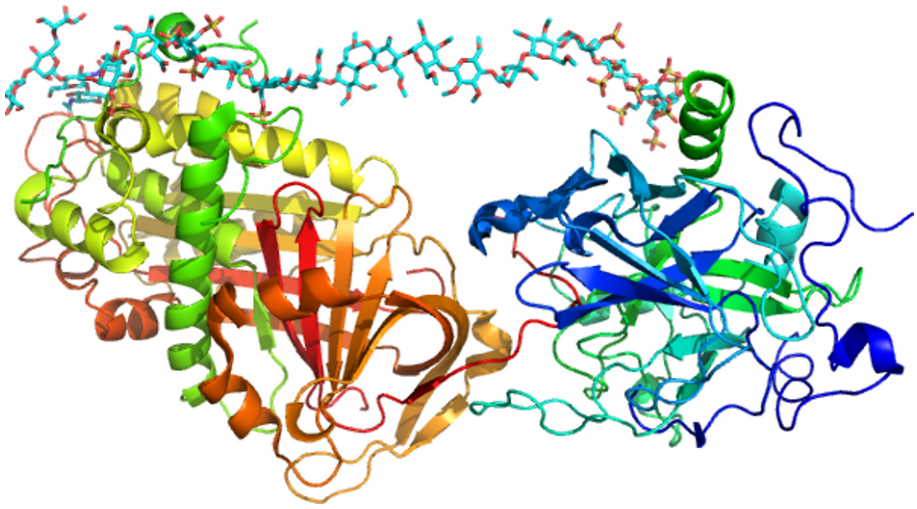


Figure 12: Three-dimensional depiction of the ternary complex formed by a heparin mimetic in interaction with antithrombin. The structure has been solved at 2.6 Å resolution (PDB 1TB6) [59]. The basis of the antithrombotic properties of therapeutic heparin could partly be deciphered by the availability of such a three-dimensional structure.

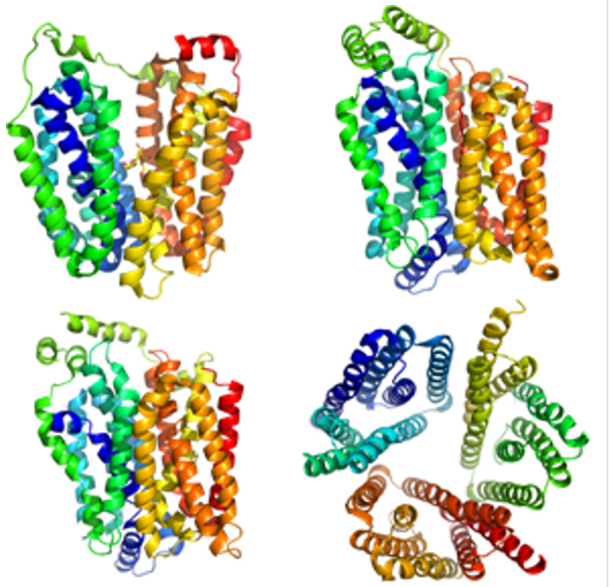


Figure 13: 3D representation of different sugar transporter structures: (left to right, top to down) lactose permease structure (PDB 1PV7, [60]), of the human glucose transporter GLUT1 (PDB 4PYP, [61]), of the bovine fructose transporter (GLUT 5) (PDB 4YB9, [73]) and of a SWEET transporter of *Orzyva sativa* (PDB 5CTG, [72]).

wide range of biological systems. Two types of intermediate trapping schemes are available.

In the so-called "trigger-freeze" experiment, a large fraction of molecules is brought into the intermediate state of interest at room temperature, which is trapped by flash-cooling. While in a "freeze-trigger" experiment, the sample is first flash-cooled, and then the reaction is triggered potentially after a transient and controlled temperature increase [74].

The 'trigger-freeze" approach consists in the use of various soaking times for the crystal with substrate sugars in presence of H_2O or glycerol ("trigger step") followed by the freezing step ranging from a few minutes to several hours. The use of a "freeze-trigger" approach solves the synchronization issue but introduces experimental complications as a photo-activable analogue called a caged-compound is required. For this, the substrates have to be modified by adding a photolabile group that prevents the reaction from occurring. The principles of the ideal cage-compound based kinetic crystallography experiment are presented in Figure 14.

The studies of the mechanism of blood group glycosyl transferase have been investigated by kinetic crystallography approaches with the aim of characterizing the double-displacement mechanism which involves the formation of a covalently bound glycosyl-enzyme intermediate, by trapping and solving the X-ray structure of this intermediate [75]. The A and B antigenic determinants are synthesised by the blood group A (GTA) and the blood group B (GTB) glycosyltransferases which transfer GalNAc from UDP-GalNAc for the A type and a Gal residue from UDP-Gal for the B-type. A mutant of the galactosyl transferase was created with the capacity to act as GTA and GTB [76,77]. The first attempt was conducted using the "trigger-freeze" method. The experiments were inconclusive presumably due to the lack of synchronisation of the reaction within the crystals and because the reaction time scale is shorter

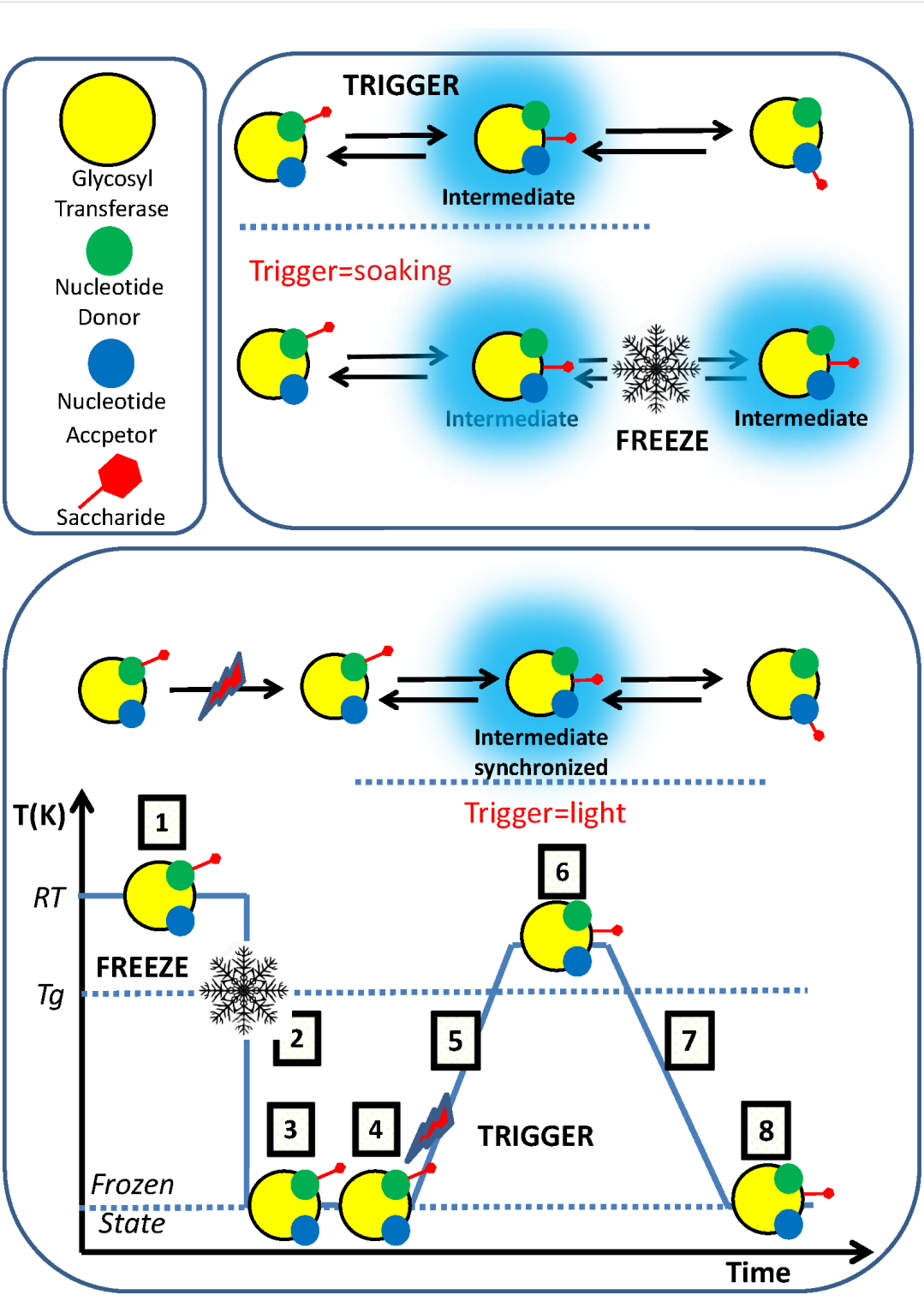


Figure 14: Kinetic crystallography. Protein crystals are soaked with the cage compound (Step 1) followed by flash-cooling (Step 2). The structure determination of the complex is solved (Step 3) to see if and how the cage-compound binds to the protein active site. Step 4 is the cleaving of the cage compound using a laser of appropriate wavelength. In this state, the substrate is available for hydrolysis but the frozen state prevents this from happening. Step 5 consists of a slow increase of the temperature to cross the glass transition and reach a temperature region where the protein has a greater degree of conformational flexibility but with a reaction rate slower than at room temperature. In Step 6, with the enzyme at room temperature, the reaction can proceed with a reorganization of the active site to transfer the sugar to a nucleophilic amino acid. An intermediate is formed and trapped during Step 7, which consists of decreasing the temperature to go back to the frozen state. Step 8 is the structural determination of the protein and elucidation of the intermediate. (Adapted from [75] with permission from Dr. G. Batot).

than the time scale of substrate diffusion in the crystals. For example, when UDP-GalNAc was soaked for 24 minutes, experiments resulted in structures with UDP-GalNAc in several conformations that are difficult to interpret. The "freeze-trigger" route was started using a series of cage compounds that had been synthesized. They all included a substrate donor, either UDP-Gal or UDP-GalNAc with an additional group on the sugar or on the uracil. Photolysis at 100 K was monitored by UV–vis absorption, both in solution and in crystals in order to assess the efficiency of the laser ablation in the crystalline glycosyl transferase. The four steps of the "freeze-trigger" process could be validated throughout by elucidation of the crystal structure of the glycosyl transferase, which has the active site occupied in a semi-closed conformation of the substrate with various levels of ordering of the internal flexible loop.

Small angle X-ray scattering

Small angle X-ray scattering is a universal technique whereby X-rays are recorded that have been elastically scattered at a low angle from samples in solution. Analysis of the scattered X-rays allows low-resolution structural information to be obtained, such as average particle size, distribution and shape. Different kinds of samples beside soluble proteins can be studied by this technique including nucleic acids, protein-based complexes, lipids, membrane proteins and surfactants, glycoproteins, virus, polymers and colloids [78,79].

Proteins: SAXS applied to biological materials (BioSAXS) is a complementary tool to protein crystallography and has become an invaluable resource for structural biologists [80]. Although at a much lower resolution than protein crystallography, BioSAXS permits the structural analysis of macromolecules at more physiological conditions, besides being suitable for the study of heterogenous systems that are unlikely to crystallise. Furthermore, the experiments in solution allow the effect of other factors, such as pH, ion concentration, or temperature, on the overall protein structure to be studied. Samples for structure analysis should be highly monodisperse. Besides sample quality control by using complementary analysis, such as dynamic light scattering, native gel, ultracentrifugation, many BioSAXS beamlines at synchrotrons are nowadays equipped with size exclusion chromatography devices directly connected with the sample cell and the data acquisition systems [81].

An illustration of how BioSAXS experiments can help to complete data obtained by protein crystallography is given by the characterization of the full structural assembly of the lectin of *Burkholderia cenocepacia,* an opportunistic bacterial pathogen. Throughout biochemical characterization, the lectin, BC2L-C was shown to be composed of two distinct domains, each displaying unique specificities and biological activities. The protein is a super lectin that binds independently to fucosylated human histo-blood group epitopes and to mannose/heptose glycoconjugates. The N-terminal domain is a fucose-binding lectin having similarity with tumour necrosis factor. The structure of the other domain (C-terminal part) which belongs to the superfamily of calcium-dependent lectins displays specificity for mannose and L-glycero-D-manno-heptose monosaccharides. The two domains are linked by a conformationally flexible sequence of 38 amino acids which was detrimental for crystallization. The respective crystal structures of the N- and C-domains could be solved separately and eventually used to establish the overall structure of the assembly by small-angle X-ray scattering (and further confirmed by electron microscopy). Figure 15 displays the reconstruction of the full macromolecular complex as a flexible arrangement of three mannose/heptose-specific dimers flanked by two fucose-specific TNF-α-like trimers [48]. This study (along with many other examples) highlights the potential of SAXS to decipher the global state of glycoproteins and carbohydrate binding proteins in solution so as to greatly amplify the high resolution 3D structural information derived from macromolecular crystallography of domains of small proteins.

Colloids: Scattering methods light, neutron and X-ray have long been the methods of choice to investigate the states of soft-condensed materials, which include solutions and gels. Differences in wavelengths and scatterers can be used for combined measurements yielding supplemental information. Following the instrumental developments of light sources, SAXS has become a common tool for the investigations of the state of materials in solution at the nanoscale. Many studies have been devoted to polysaccharides, for which structural change have been observed in real time. There exists extensive literature on this subject and the role played by polysaccharide association structures in food and in biomedical applications, as hydrogels, triggers the development of novel experiments and tools, such as optical tweezers, while making use of synchrotron radiation. Description of the details of molecular interactions occurring between complex materials such as polysaccharides and muccin is among the many new achievements that yields to the rational design of muco-adhesive polysaccharide-based nano-formulations [77].

The availability of new instrumentation that combines wide and small angle X-ray scattering and high resolution ultra-small angle X-ray scattering in a time-resolved manner is creating an opportunity to investigate the microstructure and non-equilibrium dynamics of soft matter on a length scale from a few angtroms to micrometers and on a timescale descending to the millisecond.

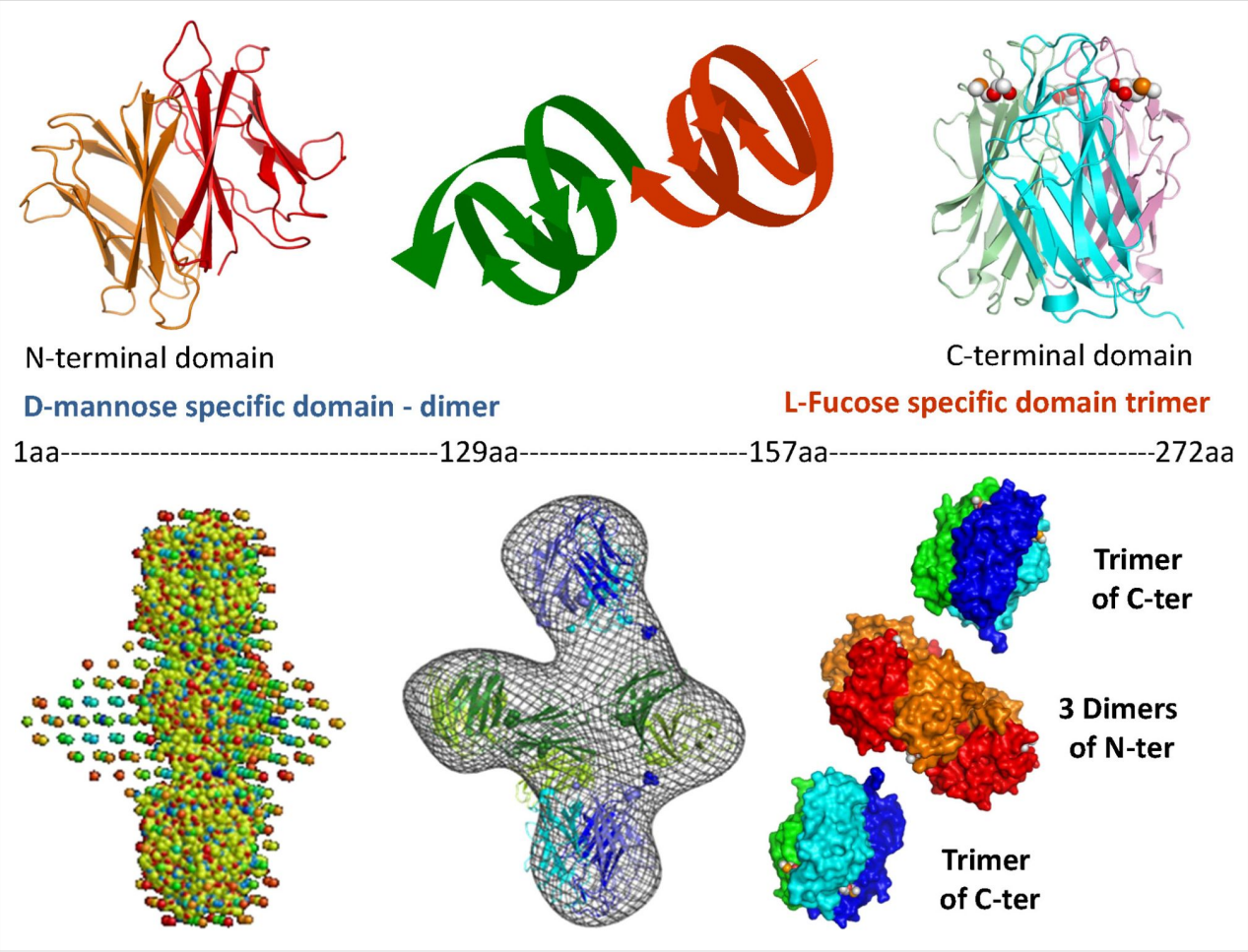


Figure 15: Reconstruction of the full three-dimensional structure of the soluble lectin (BC2L-C) from the opportunistic pathogen *Burkholderia cenocepacia* using Bio-SAXS experiments, from the knowledge of the respective crystalline structures of the N- and C-domains which had been solved separately because a sequence of 38 amino acids in the native protein was too flexible to allow crystal growth [48].

Grazing incidence X-ray reflectometry

Glycolipids: Despite their importance in the constitution and dynamics of plasma membranes, the structural and physicochemical features of gangliosides have been somehow neglected presumably because of the lack of appropriate experimental techniques. X-ray reflectometry is a surface-sensitive analytical technique based on the measure of the intensity of X-ray reflected by a flat surface. Any deviation from surface flatness will result in deviation of the reflected beam which can be analyzed to obtain the density profile of the interface normal to the surface [82].

Synchrotron X-ray reflectometry has been used to access the transverse structure of a biomimetic plasma membrane incorporating glycolipid rafts. The in situ chemical conversion of GD1a ganglioside into its metabolic product under the action of sialidase was investigated. The outcome of the sialidase action is not limited to the creation of GM1 and AsialoGM1 gangliosides as it is accompanied by a reshaping of the membrane which involves a rearrangement of the headgroups on the surface (Figure 16) [83].

Polysaccharide structures

In contrast to other macromolecules and because of the lack of regular crystalline order, X-ray diffraction of polysaccharides usually leads to an insufficient number of reflections to permit structural determination based on the data alone. Such a lack of experimental data must be complemented by modelling techniques. As such, the process of structural elucidation combines the calculation of diffraction intensities from various low energy models with those intensities collected on X-diffractograms. In this context, it is even most appropriate to use the term ‘model’ in place of ‘structure’. These experimental limitations explain why so few models of polysaccharides have been reported: there are just over 100, counting all polymorphs, variants, derivatives, and complexes. Because of their ubiquitous occurrence, polymorphs of celluloses and starches have attracted most of the attention.

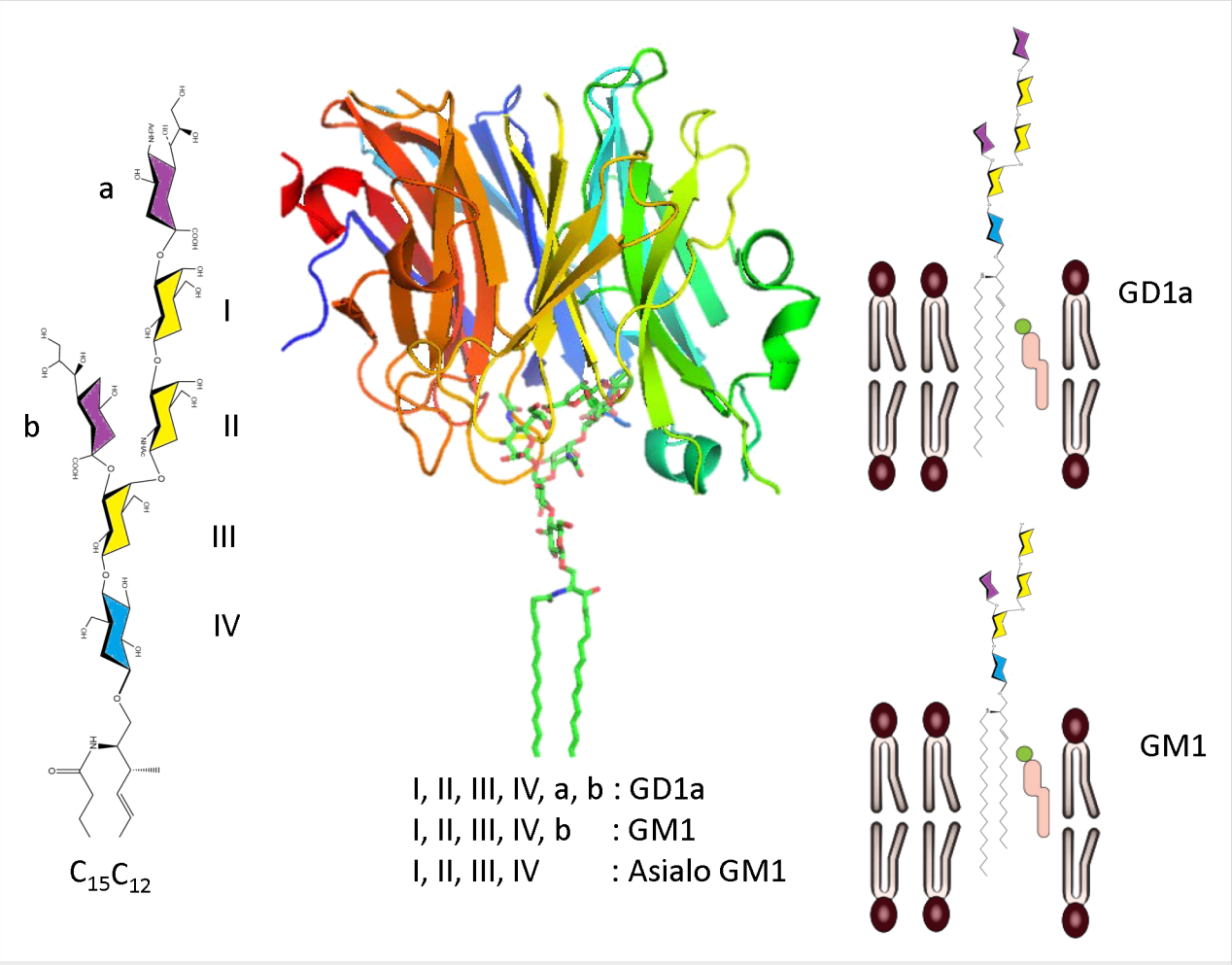


Figure 16: Characterization by synchrotron X-ray reflectometry of the transverse structures of a model membrane incorporating glycolipid rafts, under the action of a sialidase. (Adapted from [83] with permission from Dr. L. Cantù).

Fibrillar structures: cellulose and starch

Cellulose: The first X-ray fiber diffractograms of native cellulose were reported more than one century ago. The results of the investigations that have been undertaken left many of the structural details unclear as conflicting structural models were reported. One particular obstacle to be overcome in the study of cellulose microfibrils is the co-existence of a mixture of two crystal forms Iα [84] and Iβ [85]. In light of this allomorphism, the elucidation of the structure of cellulose I, awaited the mature developments of large scale facilities of synchroton and neutron sources, and the mastering of deuteration methods of the intra-crystalline regions of the native cellulose samples without altering the overall structural integrity. Through an ingenious combination of synchrotron and neutron fiber diffraction, a highly accurate structural model could be established. The samples diffracted to better than 1 Å resolution, and provided the determination of C- and O-atoms positions from the set of X-ray diffracted intensities (Figure 17). In addition, the position of hydrogen atoms were determined from Fourier-difference analysis from the set of neutron diffracted intensities collected from hydrogenated and deuterated samples [85].

This resulted in a description of the three-dimensional features of both allomorphs of native cellulose. Nevertheless, a detailed elucidation of the biosynthetic mechanism is still required to understand the occurrence of two different structural arrangements within the same microfibrils. Some unexpected features still need to be elucidated and this would require the use of complementary methods.

Starch: The complexity of starch in terms of the nature and size of its macromolecular components (amylopectin, amylose) has always been an obstacle to the elucidation of the structural components and their arrangements, which are at the origin of the birefringence of a starch granule. The structure of the crystalline domains of the two allomorphs of starch granules found in cereal and tubers had been established from a series of experimental observations (X-ray and electron crystallography) and

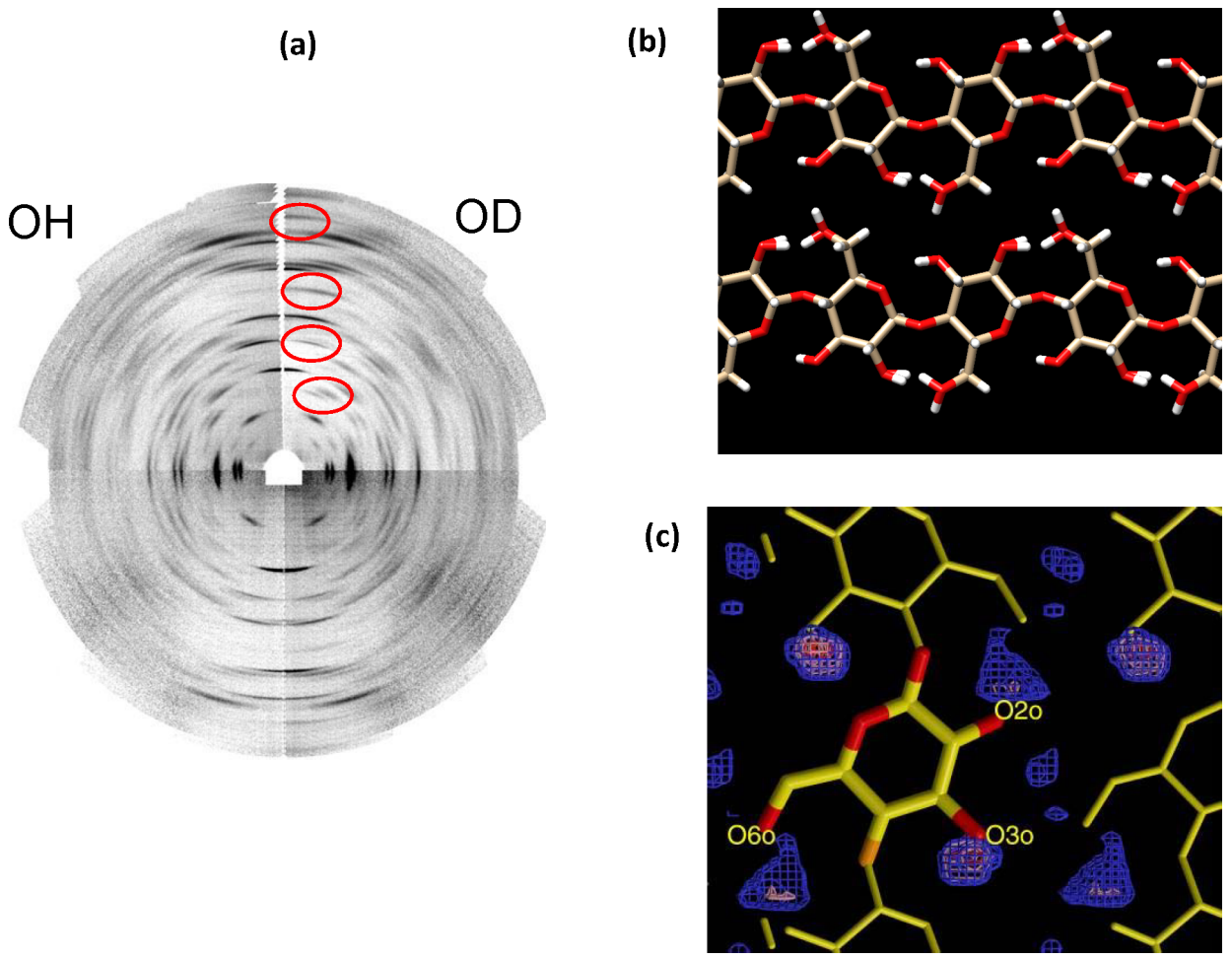


Figure 17: Complementary use of X-ray synchrotron and neutron fiber diffraction to unravel the three-dimensional structural organization of cellulose Iβ from *Halocynthia roretzi* (a). Composite fiber diffractogram of native sample (OH) and deuterated sample (OD). The differences in diffracted intensities are highlighted by the red contours (courtesy of Dr. Y. Nishiyama, with permission). (b) Depiction of conformation of the cellulose chains and their interactions in the unit cell, showing the disordered orientation of primary hydroxy groups. (c) Details of the difference of electron density highlighting the location of the deuterium atoms. Drawn from atomic coordinates taken from reference [85].

molecular modelling. While displaying differences in their mode of interactions, both allomorphs are characterized by a parallel arrangement of parallel-stranded left-handed double helices [86,87]. A second look at the crystal structure of the A-polymorph became possible when microcrystals were grown from short chains of synthetic starch and diffraction data collected using a micron-sized beam at a synchrotron source. While this new investigation corroborated the essential features of the original model, some additional fine details were revealed (Figure 18) [88].

Multiscale organization

Cellulose: Knowledge of the structure of a material is necessary to understand its properties. In the case of cellulose, it is also the key to ascertain the processes of biosynthesis. Cellulosic materials in nature often have many levels of structural complexity, whose organization depends on the source organism. In wood, a cohesive interlaced network of crystalline microfibrils of cellulose composes a first level of interacting components of the cell walls. The typical dimensions of the cellulosic fibers, which are composed of 30–40 cellulose chains, have lengths in the region of 1–2 nm and width of about 35 Å. The elucidation of the structural organization of these microfibrils came from the use of a micro-focused X-ray beam of 3 µm on a wood section of 10 µm thick and oriented perpendicular to the incident beam [89]. As depicted in Figure 19, a complete distribution map of the orientation of the axes of the cellulose microfibril, of a specimen of 52 × 42 µm, was established through a series of 546 diffraction patterns. These results can be translated into three-dimensions and establish the existence of an ultra-structural organisation in which the orientations of the cellulose fibrils follow a super-helicoidal fashion.

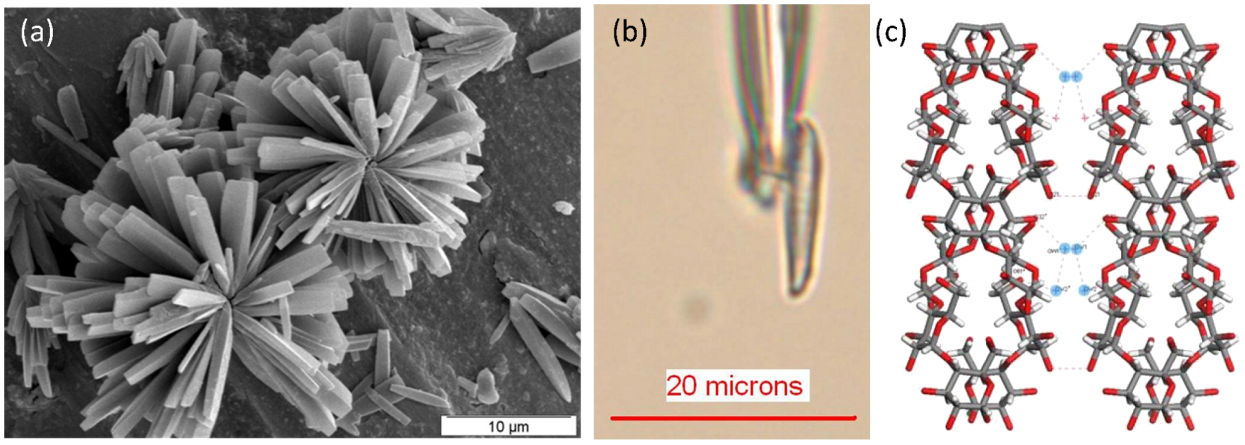


Figure 18: Scanning electron micrograph of high-quality micrometer-sized A-amylose microcrystals grown from short chains of synthetic starch (a) and of a single crystal glued to a borosilicate glass capillary tip (b). (c) Three dimensional representation of the crystal and molecular structure derived from X-ray synchrotron diffraction. (a) and (c) taken from reference [91] with permission from Actualité Chimique (http://www:lactualitechimique.org). (b) (Courtesy, J. L. Putaux; a very similar image of the subject was published in reference [88]).

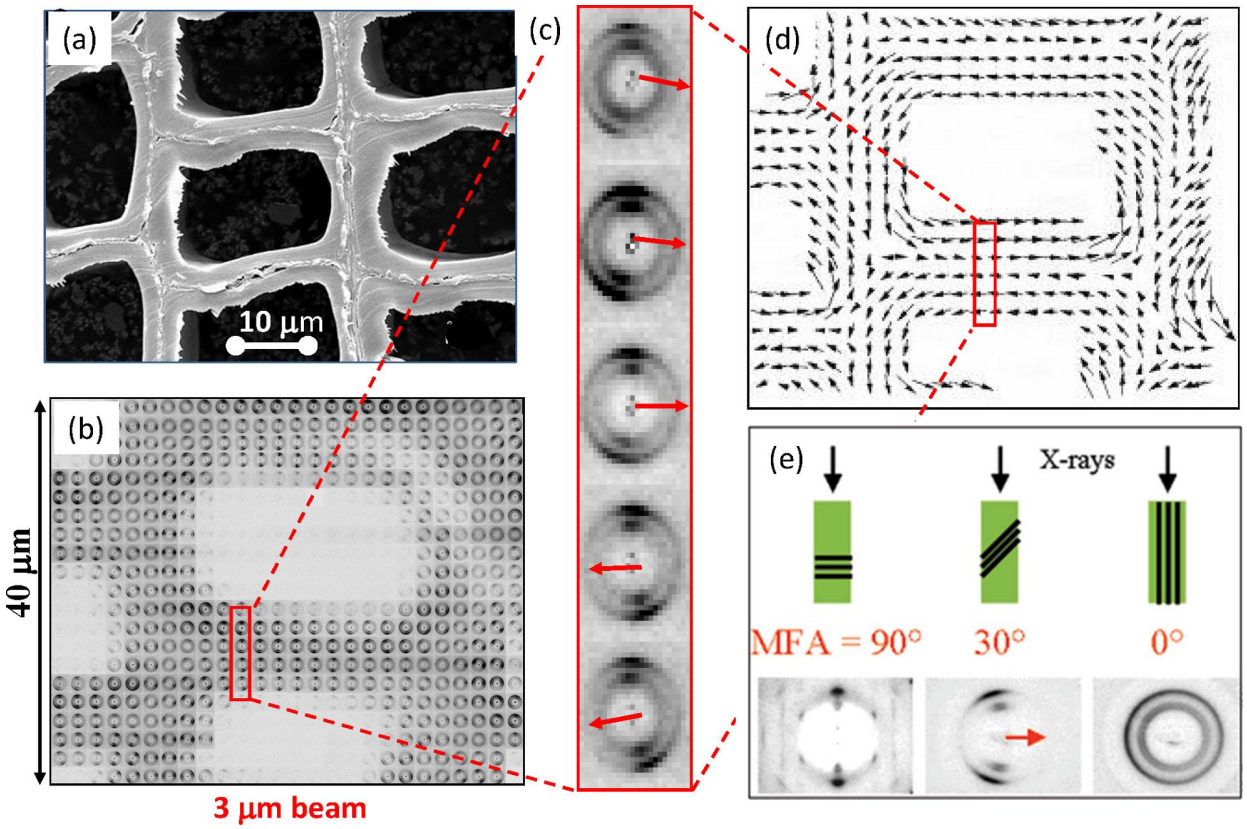


Figure 19: Cartography of distribution and orientation of cellulose in wood using a 3 µm X-ray beam. The scanning of a 10 µm thick wood section, by increments of 2 µm, (a) a collection of "fiber" like diffractograms was collected (b). From each fiber diffractogram, the attribution of the diffraction spots indicates the local orientation of the microfibril axis (c). This is depicted by arrows which indicate a marked local asymmetry in the microfibril (d). The integration of the degrees of disorientation over the full map gives the orientation of the microfibril angle (MFA) along the direction of propagation (e) (adapted from reference [89], and with permission of the International Union of Crystallography, http://journals.iucr.org).

TEMPO-mediated oxidation is one of several methods that can be used to extract nascent crystals of cellulose, or cellulose microfibrils from biomass. For this process to be optimal, some fundamental aspects of the structural and ultrastructural characterization of the cellulosic material have to be ascertained. Indeed, the extraction process should be adapted to the specificity of the various sources (e.g., wood, cotton, jute, bamboo, etc.). For dispersions in aqueous suspension, the structure of cellulose nanofibers (and their aggregates) can be characterized by SAXS. This technique has permitted quite significant insight to be gained about the structure of cellulose from a variety of botanical origins. In the case of wood pulp, cellulose nanofibers displayed a ribbon shape of about one micrometer in length. The cross-sections sizes were found to cluster in two groups with dimensions of 3 nm × 8 nm and 9 nm × 20 nm, respectively. Quite different results were obtained for the structure of microfibril fractions extracted from never-dried delignified spruce wood. In this case, the observed morphology was of the type "nanostrips" that had a characteristic thickness and width of about 0.5 nm and 4 nm, respectively. The thickness is an indication that the nanostrips are made up of only one monolayer of cellular material, which indicated the occurrence of "two-dimensional' crystals that could be further investigated by wide-angle X-ray diffraction [90].

Starch: Depending upon their botanical origin, starch granules display an elliptical shape with dimensions ranging from 0.1 to 100 µm. The advent of micro-focus X-ray diffraction from synchrotron radiation offered the possibility to explore the arrangements of the crystalline domains which are at the origin of the birefringence of the starch granule. Using a 2 µm wide X-ray beam, a complete cartography of the relative orientation on a single granule could be drawn. Two-dimensional fiber diffraction patterns were collected for each domain on a grid of 4 × 4 µm. Information about the nature of the crystalline structure (for allomorphs A, B and C) was obtained confirming the orientation of amylopectin double helices in the crystalline lamellae as well as the location of these domains and their relative orientation with respect to the granule. The most detailed investigation performed on potato starch (B allomorph) indicates that the double helices do not seem to point towards a single focus but rather towards the surface of an inner ellipsoid. Thus, the double helices have a radial orientation, and are perpendicular to the surface of the granule.

At the resolution of these ultra-structural features (10 µm) there is no discontinuity of orientation, i.e., no disclination of orientation. Between 10 µm steps, the change of the direction of the double-helices is gradual, which is consistent with a radial orientation (Figure 20) [92].

Conclusion

The aim of the present article was to describe how synchrotron radiation has benefited the field of structural glycoscience in the studies of complex carbohydrates. The atomic structures of numerous (macro)-molecules have been revealed, from molecular single crystals all the way to the complexity of polysaccharide architectures, throughout the field of protein–carbohydrate interactions. Seemingly, the study of the less well explored area of colloids and glycolipids in their membrane environments can

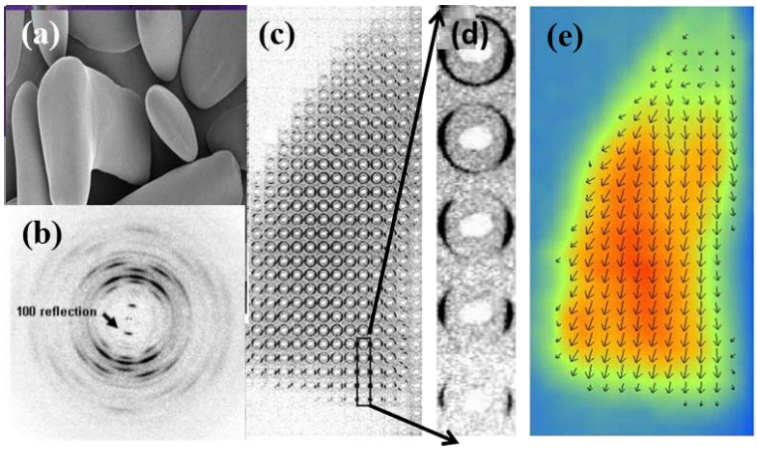


Figure 20: Structural micro-diffraction scanning of a starch granule from *Phajus grandifolius* with dimensions 50 × 200 µm. (a) Scanning electron micrograph. (b) X-ray diffractogram collected from a peripheral region of the starch grain. The width of the azymutal (100) reflection indicates the level of crystallinity. (c,d) The cartography of the crystalline domains collected on a single grain, on a grid having 4 × 4 µm dimensions, using a 1 µm X-ray beam. The total surface explored was greater than 5,000 μm^2. (e) The width and orientation of the (100) reflection on each diffractogram reflects the level of crystallinity of the explored section along with their relative orientation with respect to the fiber axis. The experiment was performed at 100 K to limit the degradation of the grain in the X-ray beam. (Taken from reference [91] with permission from Actualité Chimique http://www.lactualitechimique.org).

be tackled. The increasing speed of data collection times and photon flux are opening the way to time-resolved studies. The application of kinetic crystallography to elucidate glyco-enzymatic mechanisms is still at its infancy. Complementary to instrumental developments, the contribution of organic synthesis will be essential for the development of cage compounds, tailored to initiate light-activated reactions.

The results that have been presented were obtained on third-generation synchrotron sources. More sophisticated fourth-generation X-ray linear sources (X-ray Free Electron Lasers – XFEL) are operating at Stanford (USA), Hamburg (Germany) and in Harima (Japan). The brightness of the X-ray beams are then orders of magnitude greater and with short pulses, down to a few femtoseconds. A world full of novel experiments can be envisaged involving diffraction as well as the possibility to image non-periodic materials. Furthermore, different third-generation sources are planning major upgrades of their machine lattice to produce diffraction limited storage rings (DLSR) that will open new avenues in the science performed at these sources.

Synchrotron radiation offers much more than diffraction experiments and many other experiments can be performed making use of either the spectroscopy or imaging techniques. Spectroscopy techniques allow identification and characterization of molecular substances and their dynamics. Imaging techniques use the light-source beam to obtain pictures with spatial resolution of the sample under study. Integration of the results gathered from such experiments is a requirement to get deeper insight into the structures and mechanisms of vital biological processes in plants, animals and human.

Acknowledgments

The authors are indebted to the many colleagues who provided material to illustrate the wealth of applications and results that have been achieved using synchrotron light over the recent years. Gary Admans (ESRF) is gratefully acknowledge for critical reading of the manuscript. Part of this work was supported by the Cross Disciplinary Program Glyco@Alps, within the framework "Investissements d'avenir" program (ANR-15-IDEX-02). The three-dimensional representations of the proteins were drawn using the PyMol Molecular Graphics Systems, Version 1.8 Schrödinger, LLC.

References

1. Eriksson, M.; van der Veen, J. F.; Quitmann, C. *J. Synchrotron Radiat.* **2014,** *21,* 837–842. doi:10.1107/S1600577514019286
2. Bianconi, A.; Dell'Ariccia, M.; Durham, P. J.; Pendry, B. *Phys. Rev. B: Condens. Matter Mater. Phys.* **1982,** *26,* 6502–6508. doi:10.1103/PhysRevB.26.6502
3. Bragg, W. H.; Bragg, W. L. *Proc. R. Soc. London, Ser. A* **1913,** *88,* 428–438. doi:10.1098/rspa.1913.0040
4. Martinez-Criado, G.; Borfecchia, E.; Mino, L.; Lamberti, C. Micro and Nano X-ray beams. In *Characterization of Semiconductor Heterostructures and Nanostructures;* Lamberti, C.; Agostini, G., Eds.; 2013; pp 361–412. doi:10.1016/b978-0-444-59551-5.00009-1
5. Riekel, C. *Rep. Prog. Phys.* **2000,** *63,* 233. doi:10.1088/0034-4885/63/3/201
6. Kirkpatrick, P.; Baez, A. V. *J. Opt. Soc. Am.* **1948,** *38,* 766–774. doi:10.1364/JOSA.38.000766
7. Snigirev, A.; Kohn, V.; Snigireva, I.; Lengeler, B. *Nature* **1996,** *384,* 49–51. doi:10.1038/384049a0
8. Hendrickson, W. A.; Teeter, M. M. *Nature* **1981,** *290,* 107–113. doi:10.1038/290107a0
9. Karle, J. *Int. J. Quantum Chem.* **1980,** *18,* 357–367. doi:10.1002/qua.560180734
10. Fischer, E. *Ber. Dtsch. Chem. Ges.* **1891,** *24,* 1836–1845. doi:10.1002/cber.189102401311
11. Raymond, S.; Heyraud, A.; Tran Qui, D.; Kvick, A.; Chanzy, H. *Macromolecules* **1995,** *28,* 2096–2100. doi:10.1021/ma00110a051
12. Rencurosi, A.; Mitchell, E. P.; Cioci, G.; Pérez, S.; Pereda-Miranda, R.; Imberty, A. *Angew. Chem., Int. Ed.* **2004,** *43,* 5918–5922. doi:10.1002/anie.200460327
13. Brunelli, M.; Wright, J. P.; Vaughan, G. B. M.; Mora, A. J.; Fitch, A. N. *Angew. Chem., Int. Ed.* **2003,** *42,* 2029–2032. doi:10.1002/anie.200250607
14. Ochsenbein, P.; Kieffer, J.; El Hajji, M. In *European Powder Diffraction Conference, Darmstadt, Germany,* Darmstadt, Germany; 2010.
15. Collings, I.; Watier, Y.; Giffard, M.; Dagogo, S.; Kahn, R.; Bonneté, F.; Wright, J. P.; Fitch, A. N.; Margiolaki, I. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2010,** *D66,* 539–548. doi:10.1107/S0907444910005354
16. Karavassili, F.; Giannopoulou, A. E.; Kotsiliti, E.; Knight, L.; Norrman, M.; Schluckebier, G.; Drube, L.; Fitch, A. N.; Wright, J. P.; Margiolaki, I. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012,** *D68,* 1632–1641. doi:10.1107/S0907444912039339
17. Margiolaki, I.; Giannopoulou, A. E.; Wright, J. P.; Knight, L.; Norrman, M.; Schluckebier, G.; Fitch, A. N.; Von Dreele, R. B. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2013,** *D69,* 978–990. doi:10.1107/S0907444913003867
18. Rose, P. W.; Prlić, A.; Altunkaya, A.; Bi, C.; Bradley, A. R.; Christie, C. H.; Di Costanzo, L.; Duarte, J. M.; Dutta, S.; Feng, Z.; Kramer Green, R.; Goodsell, D. S.; Hudson, B.; Kalro, T.; Lowe, R.; Peisach, E.; Randle, C.; Rose, A. S.; Shao, C.; Tao, Y.-P.; Valasatava, Y.; Voigt, M.; Westbrook, J. D.; Woo, J.; Yang, H.; Young, J. Y.; Zardecki, C.; Berman, H. M.; Burley, S. K. *Nucleic Acids Res.* **2017,** *45,* D271–D281. doi:10.1093/nar/gkw1000
19. Stevens, R. C. *Curr. Opin. Struct. Biol.* **2000,** *10,* 558–563. doi:10.1016/S0959-440X(00)00131-7
20. Cipriani, F.; Röwer, M.; Landret, C.; Zander, U.; Felisaz, F.; Márquez, J. A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2012,** *D68,* 1393–1399. doi:10.1107/S0907444912031459
21. Flot, D.; Mairs, T.; Giraud, T.; Guijarro, M.; Lesourd, M.; Rey, V.; van Brussel, D.; Morawe, C.; Borel, C.; Hignette, O.; Chavanne, J.; Nurizzo, D.; McSweeney, S.; Mitchell, E. *J. Synchrotron Radiat.* **2010,** *17,* 107–118. doi:10.1107/S0909049509041168

22. de Sanctis, D.; Beteva, A.; Caserotto, H.; Dobias, F.; Gabadinho, J.; Giraud, T.; Gobbo, A.; Guijarro, M.; Lentini, M.; Lavault, B.; Mairs, T.; McSweeney, S.; Petitdemange, S.; Rey-Bakaikoa, V.; Surr, J.; Theveneau, P.; Leonard, G. A.; Mueller-Dieckmann, C. *J. Synchrotron Radiat.* **2012,** *19,* 455–461. doi:10.1107/S0909049512009715
23. Audfray, A.; Claudinon, J.; Abounit, S.; Ruvoën-Clouet, N.; Larson, G.; Smith, D. F.; Wimmerová, M.; Le Pendu, J.; Römer, W.; Varrot, A.; Imberty, A. *J. Biol. Chem.* **2012,** *287,* 4335–4347. doi:10.1074/jbc.M111.314831
24. Gabadinho, J.; Beteva, A.; Guijarro, M.; Rey-Bakaikoa, V.; Spruce, D.; Bowler, M. W.; Brockhauser, S.; Flot, D.; Gordon, E. J.; Hall, D. R.; Lavault, B.; McCarthy, A. A.; McCarthy, J.; Mitchell, E.; Monaco, S.; MuellerDieckmann, C.; Nurizzo, D.; Ravelli, R. B. G.; Thibault, X.; Walsh, M. A.; Leonard, G. A.; McSweeney, S. M. *J. Synchrotron Radiat.* **2010,** *17,* 700–707. doi:10.1107/S0909049510020005
25. de Sanctis, D.; Oscarsson, M.; Popov, A.; Svensson, O.; Leonard, G. *Acta Crystallogr., Sect. D* **2016,** *D72,* 413–420. doi:10.1107/S2059798316001042
26. Stepanov, S.; Makarov, O.; Hilgart, M.; Pothineni, S. B.; Urakhchin, A.; Devarapalli, S.; Yoder, D.; Becker, M.; Ogata, C.; Sanishvili, R.; Venugopalan, N.; Smith, J. L.; Fischetti, R. F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2011,** *D67,* 176–188. doi:10.1107/S0907444910053916
27. Delagenière, S.; Brenchereau, P.; Launer, L.; Ashton, A. W.; Leal, R.; Veyrier, S.; Gabadinho, J.; Gordon, E. J.; Jones, S. D.; Levik, K. E.; McSweeney, S. M.; Monaco, S.; Nanao, M.; Spruce, D.; Svensson, O.; Walsh, M. A.; Leonard, G. A. *Bioinformatics* **2011,** *27,* 3186–3192. doi:10.1093/bioinformatics/btr535
28. Monaco, S.; Gordon, E.; Bowler, M. W.; Delagenière, S.; Guijarro, M.; Spruce, D.; Svensson, O.; McSweeney, S. M.; McCarthy, A. A.; Leonard, G.; Nanao, M. H. *J. Appl. Crystallogr.* **2013,** *46,* 804–810. doi:10.1107/S0021889813006195
29. Bowler, M. W.; Nurizzo, D.; Barrett, R.; Beteva, A.; Bodin, M.; Caserotto, H.; Delagenière, S.; Dobias, F.; Flot, D.; Giraud, T.; Guichard, N.; Guijarro, M.; Lentini, M.; Leonard, G. A.; McSweeney, S.; Oskarsson, M.; Schmidt, W.; Snigirev, A.; von Stetten, D.; Surr, J.; Svensson, O.; Theveneau, P.; Mueller-Dieckmann, C. *J. Synchrotron Radiat.* **2015,** *22,* 1540–1547. doi:10.1107/S1600577515016604
30. Higgins, S. J. *Protein Expression: A Practical Approach;* Oxford University Press: Oxford, U.K., 1999; p 304.
31. Garman, S. C.; Wurzburg, B. A.; Tarchevskaya, S. S.; Kinet, J.-P.; Jardetzky, T. S. *Nature* **2000,** *406,* 259–266. doi:10.1038/35018500
32. Lee, J. E.; Fusco, M. L.; Saphire, E. O. *Nat. Protoc.* **2009,** *4,* 592–604. doi:10.1038/nprot.2009.29
33. Agirre, J.; Ariza, A.; Offen, W. A.; Turkenburg, J. P.; Roberts, S. M.; McNicholas, S.; Harris, P. V.; McBrayer, B.; Dohnalek, J.; Cowtan, K. D.; Davies, G. J.; Wilson, K. S. *Acta Crystallogr., Sect. D* **2016,** *D72,* 254–265. doi:10.1107/s2059798315024237
34. Agirre, J.; Iglesias-Fernández, J.; Rovira, C.; Davies, G. J.; Wilson, K. S.; Cowtan, K. D. *Nat. Struct. Biol.* **2015,** *22,* 833–834. doi:10.1038/nsmb.3115
35. Breton, C.; Fournel-Gigleux, S.; Palcic, M. M. *Curr. Opin. Struct. Biol.* **2012,** *22,* 540–549. doi:10.1016/j.sbi.2012.06.007
36. Coutinho, P. M.; Deleury, E.; Davies, G. J.; Henrissat, B. *J. Mol. Biol.* **2003,** *328,* 307–317. doi:10.1016/S0022-2836(03)00307-3
37. Albesa-Jové, D.; Guerin, M. E. *Curr. Opin. Struct. Biol.* **2016,** *40,* 23–32. doi:10.1016/j.sbi.2016.07.007
38. Pedersen, L. C.; Dong, J.; Taniguchi, F.; Kitagawa, H.; Krahn, J. M.; Pedersen, L. G.; Sugahara, K.; Negishi, M. *J. Biol. Chem.* **2003,** *278,* 14420–14428. doi:10.1074/jbc.M210532200
39. Hu, Y.; Chen, L.; Ha, S.; Gross, B.; Falcone, B.; Walker, D.; Mokhtarzadeh, M.; Walker, S. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100,* 845–849. doi:10.1073/pnas.0235749100
40. Rocha, J.; Cicéron, F.; de Sanctis, D.; Lelimousin, M.; Chazalet, V.; Lerouxel, O.; Breton, C. *Plant Cell* **2016,** *28,* 2352–2364. doi:10.1105/tpc.16.00519
41. Lombard, V.; Bernard, T.; Rancurel, C.; Brumer, H.; Coutinho, P. M.; Henrissat, B. *Biochem. J.* **2010,** *432,* 437–444. doi:10.1042/BJ20101185
42. Henrissat, B.; Davies, G. *Curr. Opin. Struct. Biol.* **1997,** *7,* 637–644. doi:10.1016/S0959-440X(97)80072-3
43. Boraston, A. B.; Bolam, D. N.; Gilbert, H. J.; Davies, G. J. *Biochem. J.* **2004,** *382,* 769–781. doi:10.1042/BJ20040892
44. Gilbert, H. J.; Knox, J. P.; Boraston, A. B. *Curr. Opin. Struct. Biol.* **2013,** *23,* 669–677. doi:10.1016/j.sbi.2013.05.005
45. Pérez, S.; Sarkar, A.; Rivet, A.; Breton, A.; Imberty, A. Glyco3D: A Portal for Structural Glycosciences. In *Glycoinformatics;* Lütteke, T.; Frank, M., Eds.; Methods in Molecular Biology, Vol. 1273; SprInger Science + Business Media: New York, 2015; pp 241–258. http://glyco3d.cermav.cnrs.fr
46. Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000,** *103,* 467–479. doi:10.1016/S0092-8674(00)00138-0
47. Hamelryck, T. W.; Loris, R.; Bouckaert, J.; Dao-Thi, M.-H.; Strecker, G.; Imberty, A.; Fernandez, E.; Wyns, L.; Etzler, M. E. *J. Mol. Biol.* **1999,** *286,* 1161–1177. doi:10.1006/jmbi.1998.2534
48. Šulák, O.; Cioci, G.; Delia, M.; Lahmann, M.; Varrot, A.; Imberty, A.; Wimmerová, M. *Structure* **2010,** *18,* 59–72. doi:10.1016/j.str.2009.10.021
49. Blanchard, B.; Nurisso, A.; Hollville, E.; Tétaud, C.; Wiels, J.; Pokorná, M.; Wimmerová, M.; Varrot, A.; Imberty, A. *J. Mol. Biol.* **2008,** *383,* 837–853. doi:10.1016/j.jmb.2008.08.028
50. Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. *J. Mol. Biol.* **1998,** *282,* 1043–1059. doi:10.1006/jmbi.1998.2076
51. Kumar, A.; Sýkorová, P.; Demo, G.; Dobeš, P.; Hyršl, P.; Wimmerová, M. *J. Biol. Chem.* **2016,** *291,* 25032–25049. doi:10.1074/jbc.M115.693473
52. Hester, G.; Kaku, H.; Goldstein, I. J.; Schubert Wright, C. *Nat. Struct. Biol.* **1995,** *2,* 472–479. doi:10.1038/nsb0695-472
53. Sartim, M. A.; Pinheiro, M. P.; de Pádua, R. A.; Sampaio, S. V.; Nonato, M. C. *Toxicon* **2017,** *126,* 59–69. doi:10.1016/j.toxicon.2016.12.007
54. Buts, L.; Loris, R.; De Genst, E.; Oscarson, S.; Lahmann, M.; Messens, J.; Brosens, E.; Wyns, L.; De Greve, H.; Bouckaert, J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2003,** *D59,* 1012–1015. doi:10.1107/S0907444903007170
55. Kostlánová, N.; Mitchell, E. P.; Lortat-Jacob, H.; Oscarson, S.; Lahmann, M.; Gilboa-Garber, N.; Chambat, G.; Wimmerová, M.; Imberty, A. *J. Biol. Chem.* **2005,** *280,* 27839–27849. doi:10.1074/jbc.M505184200
56. Gallego del Sol, F.; Gómez, J.; Hoos, S.; Nagano, C. S.; Cavada, B. S.; England, P.; Calvete, J. J. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2005,** *F61,* 326–331. doi:10.1107/S1744309105004835
57. Cioci, G.; Mitchell, E. P.; Chazalet, V.; Debray, H.; Oscarson, S.; Lahmann, M.; Gautier, C.; Breton, C.; Pérez, S.; Imberty, A. *J. Mol. Biol.* **2006,** *357,* 1575–1591. doi:10.1016/j.jmb.2006.01.066

58. Imberty, A.; Lortat-Jacob, H.; Pérez, S. *Carbohydr. Res.* **2007,** *342,* 430–439. doi:10.1016/j.carres.2006.12.019
59. Li, W.; Johnson, D. J.; Esmon, C. T.; Huntington, J. A. *Nat. Struct. Mol. Biol.* **2004,** *11,* 857–862. doi:10.1038/nsmb811
60. Abramson, J.; Smirnova, I.; Kasho, V.; Verner, G.; Kaback, H. R.; Iwata, S. *Science* **2003,** *301,* 610–615. doi:10.1126/science.1088196
61. Deng, D.; Xu, C.; Sun, P.; Wu, J.; Yan, C.; Hu, M.; Yan, N. *Nature* **2014,** *510,* 121–125. doi:10.1038/nature13306
62. Bianchi, L.; Diez-Sampedro, A. *PLoS One* **2010,** *5,* e10241. doi:10.1371/journal.pone.0010241
63. Cao, Y.; Jin, X.; Levin, E. J.; Huang, H.; Zong, Y.; Quick, M.; Weng, J.; Pan, Y.; Love, J.; Punta, M.; Rost, B.; Hendrickson, W. A.; Javitch, J. A.; Rajashankar, K. R.; Zhou, M. *Nature* **2011,** *473,* 50–54. doi:10.1038/nature09939
64. Dang, S.; Sun, L.; Huang, Y.; Lu, F.; Liu, Y.; Gong, H.; Wang, J.; Yan, N. *Nature* **2010,** *467,* 734–738. doi:10.1038/nature09406
65. Faham, S.; Watanabe, A.; Besserer, G. M.; Cascio, D.; Specht, A.; Hirayama, B. A.; Wright, E. M.; Abramson, J. *Science* **2008,** *321,* 810–814. doi:10.1126/science.1160406
66. Kumar, H.; Kasho, V.; Smirnova, I.; Finer-Moore, J. S.; Kaback, H. R.; Stroud, R. M. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111,* 1784–1788. doi:10.1073/pnas.1324141111
67. McCoy, J. G.; Ren, Z.; Stanevich, V.; Lee, J.; Mitra, S.; Levin, E. J.; Poget, S.; Quick, M.; Im, W.; Zhou, M. *Structure* **2016,** *24,* 956–964. doi:10.1016/j.str.2016.04.003
68. Oldham, M. L.; Chen, J. *Science* **2011,** *332,* 1202–1205. doi:10.1126/science.1200767
69. Sun, L.; Zeng, X.; Yan, C.; Sun, X.; Gong, X.; Rao, Y.; Yan, N. *Nature* **2012,** *490,* 361–366. doi:10.1038/nature11524
70. Wang, J.; Yan, C.; Li, Y.; Hirata, K.; Yamamoto, M.; Yan, N.; Hu, Q. *Cell Res.* **2014,** *24,* 1486–1489. doi:10.1038/cr.2014.144
71. Watanabe, A.; Choe, S.; Chaptal, V.; Rosenberg, J. M.; Wright, E. M.; Grabe, M.; Abramson, J. *Nature* **2010,** *468,* 988–991. doi:10.1038/nature09580
72. Xu, Y.; Tao, Y.; Cheung, L. S.; Fan, C.; Chen, L. Q.; Xu, S.; Perry, K.; Frommer, W. B.; Feng, L. *Nature* **2014,** *515,* 448–452. doi:10.1038/nature13670
73. Nomura, N.; Verdon, G.; Kang, H. J.; Shimamura, T.; Nomura, Y.; Sonoda, Y.; Hussien, S. A.; Qureshi, A. A.; Coincon, M.; Sato, Y.; Abe, H.; Nakada-Nakura, Y.; Hino, T.; Arakawa, T.; Kusano-Arai, O.; Iwanari, H.; Murata, T.; Kobayashi, T.; Hamakubo, T.; Kasahara, M.; Iwata, S.; Drew, D. *Nature* **2015,** *526,* 397–401. doi:10.1038/nature14909
74. Bourgeois, D.; Weik, M. *Crystallogr. Rev.* **2009,** *15,* 87–118. doi:10.1080/08893110802604868
75. Batot, G. Une nouvelle approche pour étudier le mécanisme des glycosyltransférases. Ph.D. Thesis, Université de Grenoble, Grenoble, France, 2013.
76. Jørgensen, R.; Batot, G.; Mannerstedt, K.; Imberty, A.; Breton, C.; Hindsgaul, O.; Royant, A.; Palcic, M. M. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2014,** *F70,* 1015–1021. doi:10.1107/S2053230X1401259X
77. Menchicchi, B.; Fuenzalida, J. P.; Hensel, A.; Swamy, M. J.; David, L.; Rochas, C.; Goycoolea, F. M. *Biomacromolecules* **2015,** *16,* 924–935. doi:10.1021/bm501832y
78. Pernot, P.; Round, A.; Barrett, R.; De Maria Antolinos, A.; Gobbo, A.; Gordon, E.; Huet, J.; Kieffer, J.; Lentini, M.; Mattenet, M.; Morawe, C.; Mueller-Dieckmann, C.; Ohlsson, S.; Schmid, W.; Surr, J.; Theveneau, P.; Zerrad, L.; McSweeney, S. *J. Synchrotron Radiat.* **2013,** *20,* 660–664. doi:10.1107/S0909049513010431
79. Round, A.; Felisaz, F.; Fodinger, L.; Gobbo, A.; Huet, J.; Villard, C.; Blanchet, C. E.; Pernot, P.; McSweeney, S.; Roessle, M.; Svergun, D. I.; Cipriani, F. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015,** *D71,* 67–75. doi:10.1107/S1399004714026959
80. Bizien, T.; Durand, D.; Roblina, P.; Thureau, A.; Vachette, P.; Pérez, J. *Protein Pept. Lett.* **2016,** *23,* 217–231. doi:10.2174/0929866523666160106153655
81. De Maria Antolinos, A.; Pernot, P.; Brennich, M. E.; Kieffer, J.; Bowler, M. W.; Delageniere, S.; Ohlsson, S.; Malbet Monaco, S.; Ashton, A.; Franke, D.; Svergun, D.; McSweeney, S.; Gordon, E.; Round, A. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **2015,** *D71,* 76–85. doi:10.1107/S1399004714019609
82. Daillant, J.; Gibaud, A. *X-ray and Neutron Reflectivity: Principles and Applications;* Springer, 1999.
83. Rondelli, V.; Brocca, P.; Fragneto, G.; Daillant, J.; Tringali, C.; Cantu, C.; Del Favero, E. *Biochim. Biophys. Acta, Biomembr.* **2017,** *1859,* 845–851. doi:10.1016/j.bbamem.2017.01.012
84. Nishiyama, Y.; Sugiyama, J.; Chanzy, H.; Langan, P. *J. Am. Chem. Soc.* **2003,** *125,* 14300–14306. doi:10.1021/ja037055w
85. Nishiyama, Y.; Langan, P.; Chanzy, H. *J. Am. Chem. Soc.* **2002,** *124,* 9074–9082. doi:10.1021/ja0257319
86. Imberty, A.; Chanzy, H.; Pérez, S.; Buléon, A.; Tran, V. *J. Mol. Biol.* **1988,** *201,* 365–378. doi:10.1016/0022-2836(88)90144-1
87. Imberty, A.; Pérez, S. *Biopolymers* **1988,** *27,* 1205–1221. doi:10.1002/bip.360270803
88. Popov, D.; Buléon, A.; Burghammer, M.; Chanzy, H.; Montesanti, N.; Putaux, J.-L.; Potocki-Véronèse, G.; Riekel, C. *Macromolecules* **2009,** *42,* 1167–1174. doi:10.1021/ma801789j
89. Lichtenegger, H.; Müller, M.; Paris, O.; Riekel, C.; Fratzl, P. *J. Appl. Crystallogr.* **1999,** *32,* 1127–1133. doi:10.1107/S0021889899010961
90. Hsiao, B. S. Structure characterization of cellulose nanofibers-microfibrills. *ALBA Synchrotron - Marie Curie Meeting Room;* Barcelona, 2015.
91. Pérez, S.; Burghammer, M. *Actual. Chim.* **2011,** *356–357,* 68–72.
92. Chanzy, H.; Putaux, J.-L.; Dupeyre, D.; Davies, R.; Burghammer, M.; Montanari, S.; Riekel, C. *J. Struct. Biol.* **2006,** *154,* 100–110. doi:10.1016/j.jsb.2005.11.007

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Enzymatic synthesis of glycosides: from natural *O*- and *N*-glycosides to rare *C*- and *S*-glycosides

Jihen Ati, Pierre Lafite and Richard Daniellou*

Review

Open Access

Address:
ICOA UMR CNRS 7311, University of Orléans, rue de Chartres, BP 6759, 45067 Orléans cedex 2, France

Email:
Richard Daniellou* - richard.daniellou@univ-orleans.fr

* Corresponding author

Keywords:
enzyme; glycochemistry; glycoside hydrolase; glycosyltransferase; mechanism

Beilstein J. Org. Chem. **2017,** *13,* 1857–1865.
doi:10.3762/bjoc.13.180

Received: 02 May 2017
Accepted: 17 August 2017
Published: 05 September 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Carbohydrate related enzymes, like glycosyltransferases and glycoside hydrolases, are nowadays more easily accessible and are thought to represent powerful and greener alternatives to conventional chemical glycosylation procedures. The knowledge of their corresponding mechanisms has already allowed the development of efficient biocatalysed syntheses of complex *O*-glycosides. These enzymes can also now be applied to the formation of rare or unnatural glycosidic linkages.

Introduction

The role of glycoconjugates is of prime importance, as they are nowadays well known to mediate many biological processes [1]. As a consequence, in a recently published roadmap for glycosciences in Europe, carbohydrates are expected, both by academics and industrials, to become key players in a near future in tremendous fields such as pharmaceuticals and personalized biomedicine, food, materials and renewable resources, and bioenergy for examples [2]. To achieve this goal, the glycoscientists will have to collaborate strongly to obtain pure and well-defined glycoconjugates. Indeed, even if during the last century, the chemists have engaged great efforts to successfully develop efficient means of synthesis of carbohydrate derivatives, through the use of specific protecting and/or activating groups and the fine control of the resulting anomeric linkage, thus leading now to i) a huge repertoire of stereoselective methods for glycosylation reactions [3] and ii) the premise of few automated oligosaccharide synthesis [4], such glycosylation process still remains highly target-dependent and therefore a challenge in too many cases. Even so, glycochemists were very recently able to chemically synthesize the largest polysaccharide to date: a mycobacterial arabinogalactan of 92 monosaccharide units [5]. However, recent advances especially in the area of molecular biology have allowed the emergence of biocatalytic procedures. Enzymes have proven to be efficient synthetic tools for the eco-compatible synthesis of many classes of compounds. Non-organic solvents, mild experimental condi-

tions, and high regio- or stereospecificity of the biocatalysed reaction have increased the added value of the use of enzymes in transformation processes, from the laboratory bench to the industrial scale [6]. Moreover, genetic modifications of recombinant enzymes are now powerful tools to easily alter the versatility, as well as the properties of the engineered protein. Rational mutagenesis, directed evolution, or even de novo design have dramatically broaden the applicability of enzymes in biocatalysis [7]. In the glycochemistry field, a vast array of carbohydrate-metabolizing enzymes [8] has been used to synthesize glycosides, even using multiple enzymes systems. Glycoside hydrolases (GHs) or glycosyltransferases (GTs) have been focused on in the search for glycosylation tools, and have been extensively studied for genetic engineering [9,10]. The corresponding compounds have proven useful in many applications ranging from glycosylation of natural products to pharmaceutics [11]. Classically, glycosides are linked to the aglycone moiety through an oxygen or a nitrogen atom, although many other kinds of linkages (even if rare) can be found in nature like in glycosylated proteins for example [12]. Herein, we wish to report a short but comprehensive review of the current enzymatic methods described for the synthesis of unusual *C*- and *S*-glycosidic linkages, their mechanisms and the corresponding perspectives.

Review

Glycosyltransferases

Glycosyltransferases (GTs, E.C. 2.4.1.x) catalyse the addition of a glycosyl moiety to an acceptor, using an activated sugar as donor (lipid, nucleotide…) [13]. It is considered that GTs are encoded by 1% of total genes, and over 300 000 representatives of GT superfamily have been classified according to their nucleotide sequence into 103 subfamilies [8]. Depending on the conservation of the anomeric atom stereochemistry of the sugar during GT-catalysed reactions, GTs are also classified as inverting or retaining (Figure 1). Inverting GTs operate via a

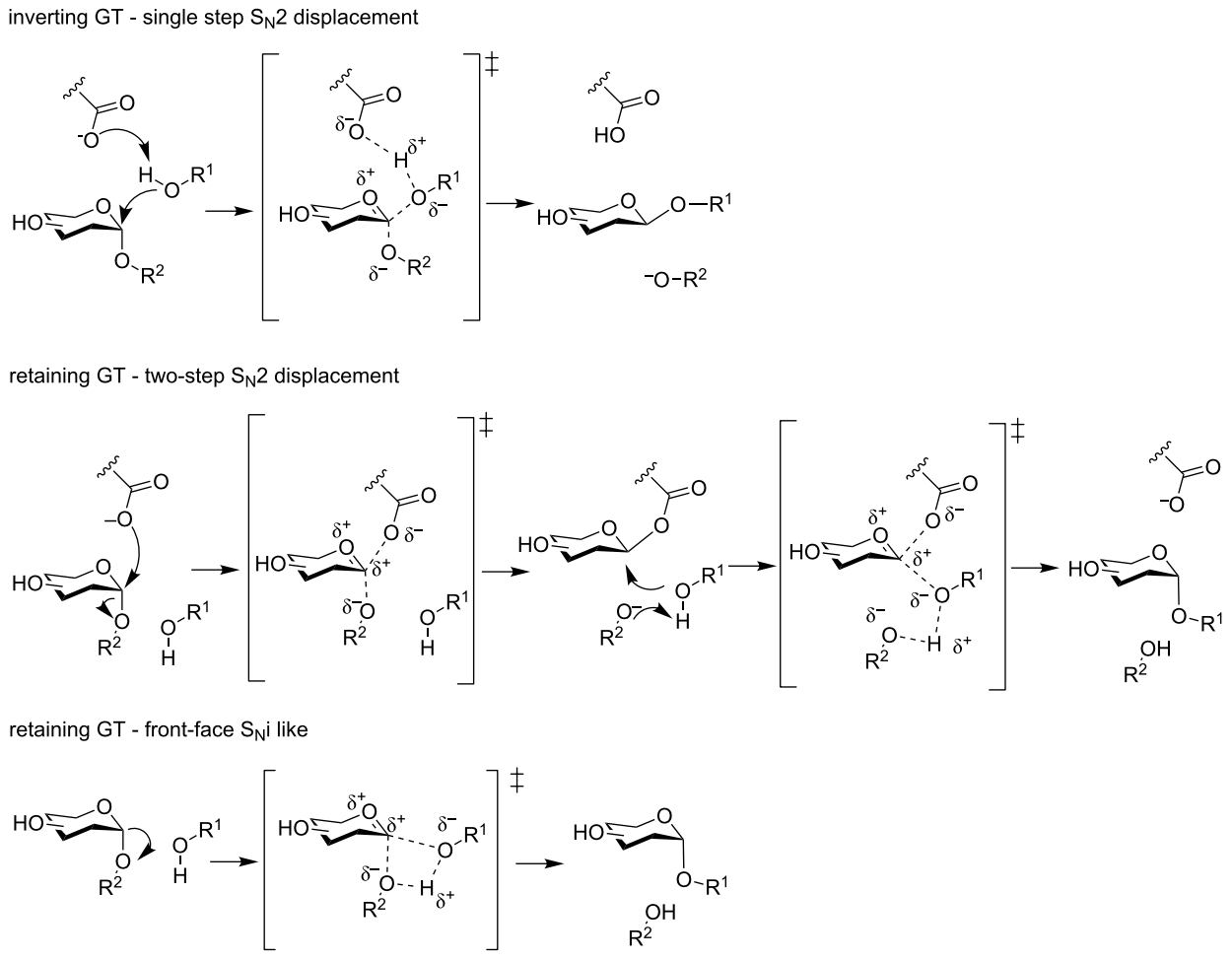


Figure 1: Mechanisms of *O*-GTs-catalysed glycosylation.

S_N2 mechanism in a single displacement step where an acid/base residue enhances the nucleophilicity of the acceptor, via an oxocarbenium-like transition state. Unlike inverting GTs, there is much controversy of the molecular mechanism of retaining GTs [13-15]. Retention of the anomeric carbon stereochemistry can occur either following a two-step displacement S_N2 type mechanism (as in retaining GH), or via a "S_Ni-like" mechanism, that involves a front-side single displacement, both via two distinct oxocarbenium-like transition states [16].

In all three mechanisms, the nucleophilic attack of the acceptor is enhanced by the deprotonation – either by an acid/base residue, like in inverting GT and two-step displacement retaining GT, or the phosphate donor (in an S_Ni-like retaining mechanism). In the case of *O*-GTs, the nucleophile is an alcohol or a phenol (carbohydrates, serine, threonine, …), whereas in *N*-GTs, the nature of the nitrogen-containing group is more diverse (amines, amides, guanidine or even indoles) [12]. *S*- and *C*-GTs follow a similar mechanism with the nucleophilic attack of the acceptor, however, little information is known on the pathways involved in these reactions (S_N2 or S_Ni-like), because of the few examples of such enzymes characterization in literature, when compared to canonical *O*- and *N*-GTs.

S-Glycosyltransferases

Few examples of natural *S*-glycosides have been described in literature [12,17,18]. Historically, glucosinolates have been the first identified *S*-glycosides for 50 years in cruciferous vegetables [19]. Along with the myrosinase GH enzyme, they are part of the "mustard bomb" system as a protective mechanism for plants against insect aggression. Their biosynthetic pathway requires the action of a *S*-GT (UGT74B1) that catalyses the reaction between a thiohydroximate acceptor and UDP-α-D-glucose as sugar donor to yield the corresponding desulfoglucosinolate (Figure 2) [20,21]. UGT74B1 from *A. thaliana* is a versatile enzyme in terms of sugar donor scope, and our group has shown the potency of this enzyme as a biocatalyst for the chemoenzymatic synthesis of non-natural desulfoglycosinolates [22]. More recently, *S*-glycosylated peptides have been identified, and characterized. In bacteria the structures of sublancin [23], glycocin F [24,25], and thurandacin [26] revealed *S*-glycosylation of cysteines. Carbohydrates bound to these bacteriocins are glucose or *N*-acetylglucosamine. For two of these glycopeptides, the corresponding *S*-GTs have been characterized and their versatility for a wide range of sugar donors has been tested [26,27]. More recently, a global protein glycosylation analysis through chemical labelling and mass finger printing have identified many *S*-glycosylation sites on different proteins, with a *N*-acetylglucosamine group bound on cysteines [28].

Other enzymes have been scarcely identified to catalyse the *S*-glycosylation, although their endogenous role is not to generate *S*-glycosides. Brazier-Hicks and colleagues have screened many *A. thaliana* Family 1 GTs with three acceptors, to identify *O*-GT, *N*-GT and *S*-GT enzymatic activities [29]. Among the 99 enzymes tested, 17 were able to use 4-chlorothiophenol as the acceptor. UGT74B1, involved in glucosinolate biosynthesis (see supra), was one of these 17 enzymes. Other studies have identified *S*-GT activities when assaying the catalytic promiscuity of *O*-GT with a wide range of aglycone acceptors (Figure 3). OleD from *Streptomyces antibioticus* has been the first reported *O*-GT to catalyse *S*-glycosylation on thiol acceptors [30]. Genetic engineering of this enzyme has also led to *S*-GT activities on several thiols. UGT73AE1 from *Carthamus tinctorius* was able to transfer glucose on a wide range of acceptors, including a *S*-containing compound, dichlorothiophenol [31]. More recently, *Bc*GT1 from *Bacillus subtilis* was shown to efficiently catalyse the glucosylation of thiol-containing acceptors [32].

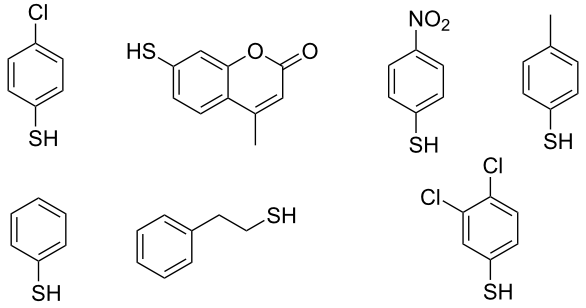


Figure 3: Examples of thiol-containing acceptors used in the chemoenzymatic biosynthesis of *S*-glycosides catalysed by *S*-GT.

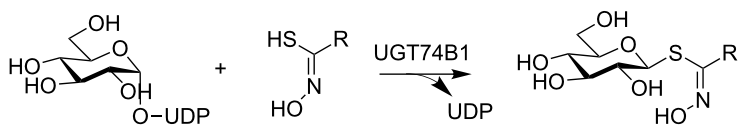


Figure 2: Desulfoglucosinolate biosynthesis by UGT74B1.

C-Glycosyltransferases

For more than 50 years, *C*-glycosides have been identified in plants [33,34] as secondary metabolites. At least 5 families of aromatic aglycones have been reported to be *C*-glycosylated: flavones, xanthones, chromones, anthrones, and gallic acids. Several corresponding plant *C*-GT have been cloned, expressed and characterized, from several crops including maize [35], rice [36,37], wheat [36], buckwheat [38] and other plants such as *Arabidopsis* [39] or *Mangifera indica* [40] (Figure 4). Fungi *C*-glycosyltransferases were also identified in *Streptomyces*, in-

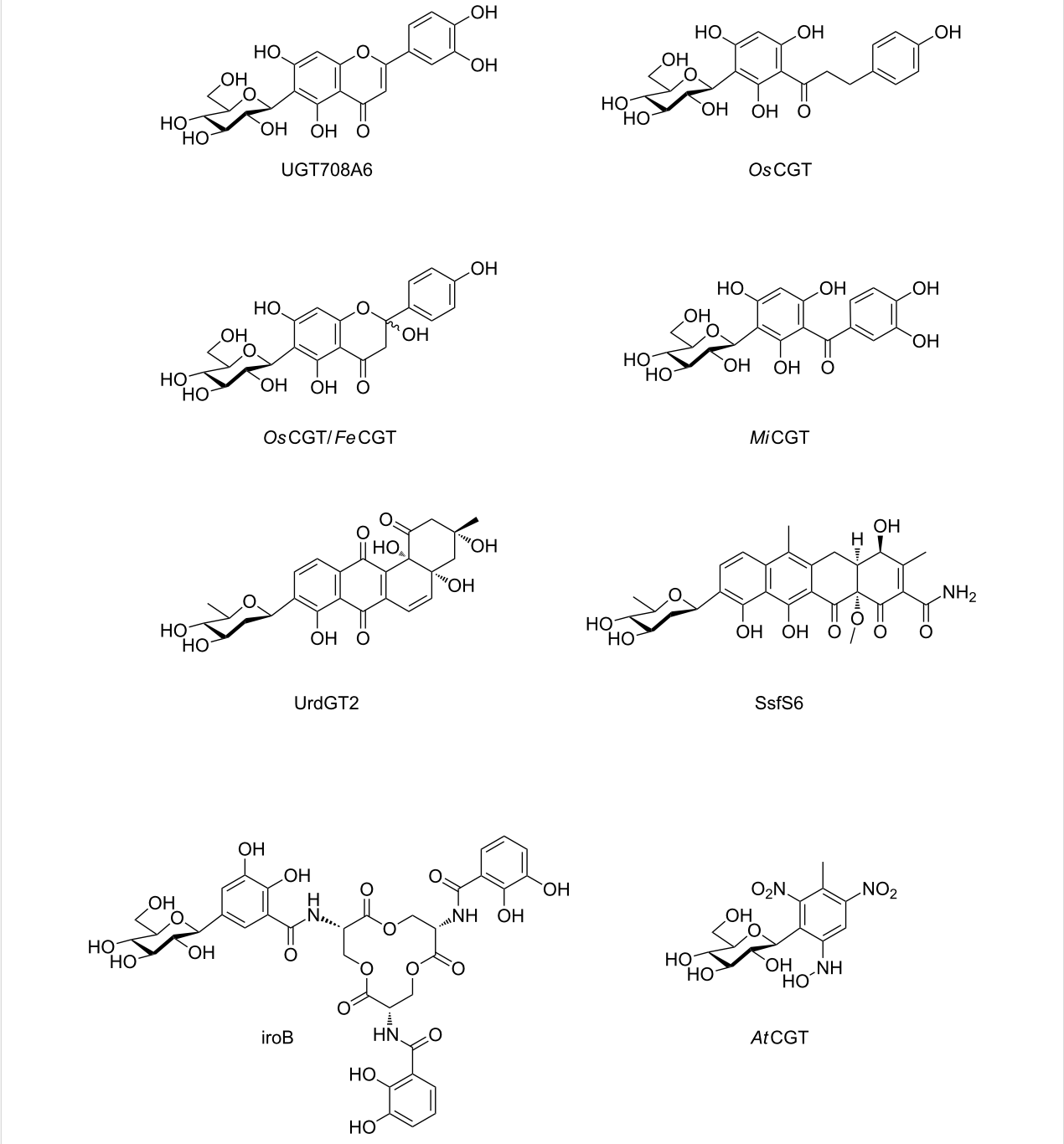


Figure 4: Examples of *C*-glycosylated products biosynthesized by natural *C*-GT. Compounds showed are formed by the action of *C*-GT found in maize (UGT708A6), rice (*Os*CGT), buckwheat (*Fe*CGT), *Mangifera indica* (*Mi*CGT), *Arabidopsis thaliana* (*At*CGT), fungi (UrdGT2 and SsfS6) or bacteria (iroB).

cluding UrdGT [41,42] and SsfS6 [43] that catalyse the transfer of the unusual D-olivosyl carbohydrate moiety on the aglycon acceptor. Bacterial *C*-GTs are the last group identified in *Salmonella enterica* and *Escherichia coli* [44-46] that are involved in the biosynthesis of siderophores, that were shown to be *C*-glycosylated enterobactins. In addition to these naturally occurring *C*-GTs, engineering of *O*-GT to *C*-GT were successfully performed in several studies [37,47,48], and chemoenzymatic syntheses of *C*-glycosides were described in other publications [40,49-51]. In all described *C*-GTs, the aglycone acceptor was found to be a derivative of polyhydroxybenzaldehyde, that exhibit an acidic carbon on the aromatic ring. Depending on the nature of the substrate and the *C*-GT involved in the enzymatic reaction, several regioselectivities were observed. A mechanistic study in 2013 by Gutmann and Nidetzky demonstrated that *C*-glycosylation occurred through a direct nucleophilic attack of an acidic carbon, and showed evidence against an *O*-glycosylation followed by an *O*-to-*C* rearrangement [37]. A last family of *C*-GT are the enzymes involved in *C*-mannosylation of protein tryptophanes [37]. However, if the corresponding *C*-GTs were identified, no mechanistic evidence was reported to date [52].

Glycoside hydrolases

GHs (E.C. 3.2.1.x) are ubiquitous enzymes responsible for the hydrolysis of the carbohydrate moieties in all the living organisms. They are actually classified in the CAZY database under 145 families, which contain more than 435,000 individual proteins [8]. Like the mechanisms described for the GTs, the catalysis of the hydrolytic reaction can occur with inversion or retention of configuration (Figure 5) as first described by Koshland [53]. However, in the case of GHs, the mechanism generally implies the intervention of two amino acid side chains, typically glutamate or aspartate, and goes through oxocarbenium ion-like transition states. Inverting GHs operate via a single step S_N2 displacement using a general acid and a general base assistance from two amino acid side chains located 6 to 11 Å apart. The mechanism of retaining GHs occurs via a two-step S_N2 displacement involving a glycosyl–enzyme intermediate, with the assistance of an acid/base and a nucleophile through two amino acid side chains located 5.5 Å apart.

A particular case of the retaining GHs is the one of the *N*-acetyl-β-hexosaminidases from the families 18, 20, 25, 56, 84 and 85 in which there is no catalytic nucleophile but where the

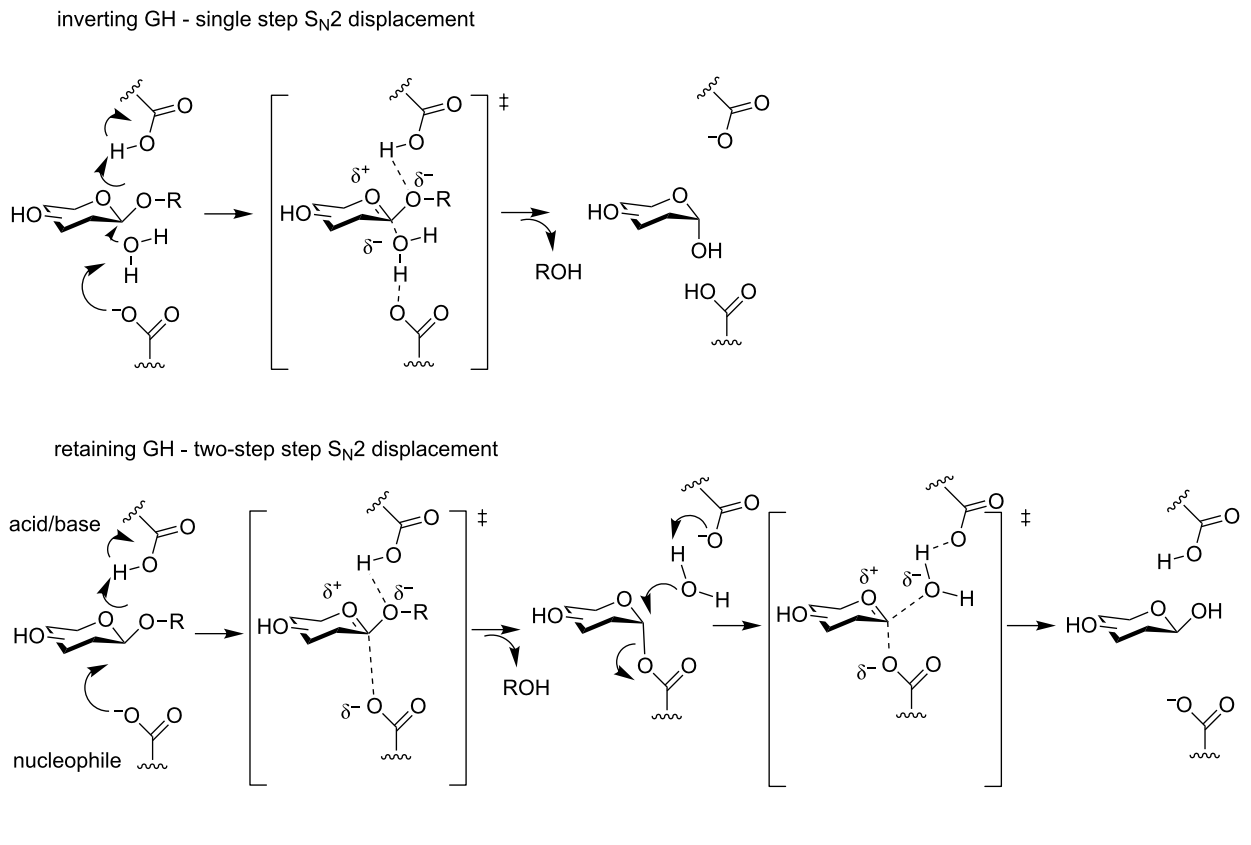


Figure 5: General mechanisms of the *O*-GHs-catalysed hydrolysis.

2-acetamido group of the substrate is acting as an internal nucleophile (Figure 6) [54].

Rarer mechanisms have also been discovered, like for example the one of i) myrosinases in plants that are retaining GHs that lack a general acid and use an exogenous base [55], or ii) some GHs from families 4 and 109 which follow a NAD-dependent hydrolysis [56,57].

Even if the major activity of GHs remains the hydrolysis, since the 1970's, these enzymes have also demonstrated their capacities to catalyse the formation of new glycosidic *O*-linkages, either by reverse hydrolysis or through transglycosylation reactions. First discovered by the team of Bourquelot and Viebel, the reaction of reverse hydrolysis can occur in the presence of GHs when a nucleophile other than water is present in the media [58,59]. Under thermodynamic control, it generally leads to the major formation of the hydrolytic product. More interestingly and in the case of the retaining GHs, the glycosyl–enzyme intermediate can be attacked by another nucleophile than water (like an alcohol) to stereospecifically yield a new glycoside [60]. The reaction is now under kinetic control and the enzyme is named a transglycosidase. The rules that guide the balance between hydrolysis and transglycosylation are still not well understood and controlling this ratio remains a challenge that still need to be solved, even if the use of artificial donors [10], the bioengineering of these biocatalysts [61] and the study of internal water dynamics [62] for examples have permitted important progresses. Consequently, such enzymatic approaches can nowadays efficiently be utilized in particular for the preparation of pure and well-defined complex glycoproteins [63].

Use of external nucleophiles

The identification of the two amino acid side chains in both retaining and inverting GHs is usually performed through site directed mutagenesis of the potent residues [64]. In these cases, the mutated enzymes are no longer able to perform the hydrolysis of the substrates. The use of external and suitable nucleophilic anions such as azide, formate or acetate allows the rescue of the activity and can also represent an efficient methodology of enzymatic synthesis of these particular (but mostly unstable) carbohydrate derivatives. These mutants were also developed as powerful biocatalysts for the synthesis of complex *O*-glycosides through the concepts of glycosynthases or thioglycosynthases [61].

S-Glycoside synthesis

In retaining GHs, the inactivation of the acid/base catalytic residue is of particular interest, and can lead to an original biocatalyst with poor hydrolytic activities but the ability to promote the formation of thioglycosidic linkages (Figure 7). Such mutated enzymes were firstly described by the team of Withers and are named thioligases [65]. Based on the mechanism, here the formation of the glycosyl–enzyme intermediate requires the use of an activated glycosyl donor, such as dinitrophenyl or azide glycosides, and the glycosylation step needs stronger nucleophiles such as thiol derivatives. The choice of the amino acid to mutate the acid/base is of crucial importance as it directly dictates the level of activity [66], but it cannot be predicted nor be a guarantee for success [67].

In general, the reported thiol acceptor is a monosaccharide or a substituted thiophenol (Figure 8). These engineered GHs were already successfully applied to the biocatalysed synthesis of thiodi- or -trisaccharides [68-72], neo-thioglycoprotein [73], or even simple thioglycosides with potent inhibitory properties [74]. More complex biomolecules were also obtained like glycans or glycopolymers [75], or even rarer thiofuranosides [76].

C-Glycosides synthesis

There is no example of *C*-glycoside synthesis promoted by GHs reported in the literature so far, although as depicted by the

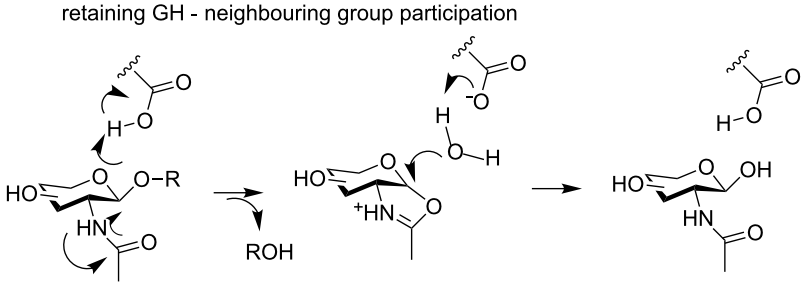


Figure 6: Neighbouring group participation mechanism of retaining the *O*-GHs-catalysed hydrolysis.

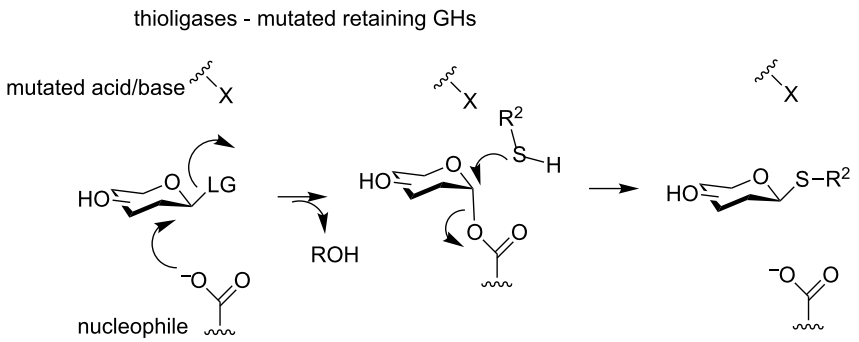


Figure 7: Mechanism of the thioligases.

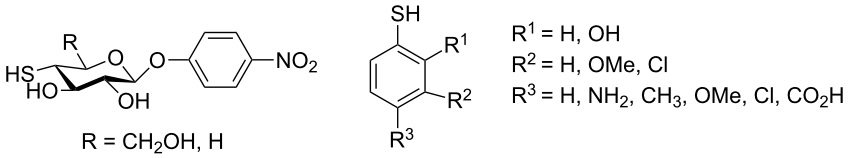


Figure 8: Examples of thiol acceptors utilized with GHs.

mechanism, this kind of biocatalysed reaction can be envisioned.

Conclusion

To conclude, the enzymatic mechanisms that rule the activities of GTs and GHs begin to be well understood by the glycoscientists. Their application to the enzymatic synthesis of a great variety of *O*- and *N*-glycosides are already becoming a routine. In addition the utilization of enzymes so to obtain rarer *C*- and *S*-analogues is an emerging field restricted to few acceptors (Table 1). Still, as demonstrated by the CAZY database, the glycoscientists have nowadays access to a large (and increasing) library of GTs and GHs, and in a near future, they will be able to perform most reactions enzymatically. In addition, the access to large quantities of inexpensive substrates can also be envisioned. In parallel, our knowledge of the enzymatic mechanisms has allowed us to modify and improve the original activities through reasoned site-directed mutagenesis. However, despite major advances, all the rules that finely tune the biocatalysts are still poorly understood and the luck is in too many cases the best road to success. The unlock of the biotechnological bolts in this particular field will certainly occur by our deep understanding of the role of second-sphere amino acids and of protein motions. This will require the generation of huge libraries of mutants and the fast screening of their activities, as well as powerful molecular modelling and crystallization of

Table 1: Summary of rare synthetic activities of carbohydrate-related enzymes.

Activity	Enzyme family	Acceptor	References
S-glycosylation	GTs	desulfoglycosinolates	[20-22]
		cysteine	[26-28]
		aromatics	[29-32]
	GHs	saccharides	[68-73]
		aromatics	[74-76]
C-glycosylation	GTs	aromatics	[35-51]
	GHs	no reference	no reference

proteins. No doubt then that such biocatalysts will represent a competitive tool for glycosylation so to obtain complex *O*-, *S*- or *C*-glycoconjugates of biological interests.

References

1. Varki, A. *Glycobiology* **1993,** *3,* 97–130. doi:10.1093/glycob/3.2.97
2. A Roadmap for Glycoscience in Europe. http://www.glycopedia.eu/news/article/a-roadmap-for-glycoscience-in (accessed April 17, 2017).
3. Bennett, C. S. *Selective Glycosylations: Synthetic Methods and Catalysts;* Wiley VCH Verlag GmbH, 2017. doi:10.1002/9783527696239
4. Hsu, C.-H.; Hung, S.-C.; Wu, C.-Y.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2011,** *50,* 11872–11923. doi:10.1002/anie.201100125
5. Wu, Y.; Xiong, D.-C.; Chen, S.-C.; Wang, Y.-S.; Ye, X.-S. *Nat. Commun.* **2017,** *8,* No. 14851. doi:10.1038/ncomms14851
6. Panke, S.; Held, M.; Wubbolts, M. *Curr. Opin. Biotechnol.* **2004,** *15,* 272–279. doi:10.1016/j.copbio.2004.06.011
7. Desmet, T.; Soetaert, W.; Bojarová, P.; Křen, V.; Dijkhuizen, L.; Eastwick-Field, V.; Schiller, A. *Chem. – Eur. J.* **2012,** *18,* 10786–10801. doi:10.1002/chem.201103069
8. Lombard, V.; Golaconda Ramulu, H.; Drula, E.; Coutinho, P. M.; Henrissat, B. *Nucleic Acids Res.* **2014,** *42,* D490–D495. doi:10.1093/nar/gkt1178
9. Hancock, S. M.; Vaughan, M. D.; Withers, S. G. *Curr. Opin. Chem. Biol.* **2006,** *10,* 509–519. doi:10.1016/j.cbpa.2006.07.015
10. Wang, L.-X.; Huang, W. *Curr. Opin. Chem. Biol.* **2009,** *13,* 592–600. doi:10.1016/j.cbpa.2009.08.014
11. Ernst, B.; Magnani, J. L. *Nat. Rev. Drug Discovery* **2009,** *8,* 661–677. doi:10.1038/nrd2852
12. Lafite, P.; Daniellou, R. *Nat. Prod. Rep.* **2012,** *29,* 729–738. doi:10.1039/c2np20030a
13. Lairson, L. L.; Henrissat, B.; Davies, G. J.; Withers, S. G. *Annu. Rev. Biochem.* **2008,** *77,* 521–555. doi:10.1146/annurev.biochem.76.061005.092322
14. Ardèvol, A.; Rovira, C. *J. Am. Chem. Soc.* **2015,** *137,* 7528–7547. doi:10.1021/jacs.5b01156
15. Rojas-Cervellera, V.; Ardèvol, A.; Boero, M.; Planas, A.; Rovira, C. *Chem. – Eur. J.* **2013,** *19,* 14018–14023. doi:10.1002/chem.201302898
16. Persson, K.; Ly, H. D.; Dieckelmann, M.; Wakarchuk, W. W.; Withers, S. G.; Strynadka, N. C. J. *Nat. Struct. Mol. Biol.* **2001,** *8,* 166–175. doi:10.1038/84168
17. Guillotin, L.; Lafite, P.; Daniellou, R. *Carbohydrate Chemistry;* The Royal Society of Chemistry, 2014; Vol. 40, pp 178–194.
18. Spiro, R. G. *Glycobiology* **2002,** *12,* 43R–56R. doi:10.1093/glycob/12.4.43R
19. Halkier, B. A.; Gershenzon, J. *Annu. Rev. Plant Biol.* **2006,** *57,* 303–333. doi:10.1146/annurev.arplant.57.032905.105228
20. Grubb, C. D.; Zipp, B. J.; Ludwig-Müller, J.; Masuno, M. N.; Molinski, T. F.; Abel, S. *Plant J.* **2004,** *40,* 893–908. doi:10.1111/j.1365-313X.2004.02261.x
21. Kopycki, J.; Wieduwild, E.; Kohlschmidt, J.; Brandt, W.; Stepanova, A. N.; Alonso, J. M.; Pedras, M. S. C.; Abel, S.; Grubb, C. D. *Biochem. J.* **2013,** *450,* 37–46. doi:10.1042/BJ20121403
22. Marroun, S.; Montaut, S.; Marquès, S.; Lafite, P.; Coadou, G.; Rollin, P.; Jousset, G.; Schuler, M.; Tatibouët, A.; Oulyadi, H.; Daniellou, R. *Org. Biomol. Chem.* **2016,** *14,* 6252–6261. doi:10.1039/C6OB01003B
23. Oman, T. J.; Boettcher, J. M.; Wang, H.; Okalibe, X. N.; van der Donk, W. A. *Nat. Chem. Biol.* **2011,** *7,* 78–80. doi:10.1038/nchembio.509
24. Stepper, J.; Shastri, S.; Loo, T. S.; Preston, J. C.; Novak, P.; Man, P.; Moore, C. H.; Havlíček, V.; Patchett, M. L.; Norris, G. E. *FEBS Lett.* **2011,** *585,* 645–650. doi:10.1016/j.febslet.2011.01.023
25. Venugopal, H.; Edwards, P. J. B.; Schwalbe, M.; Claridge, J. K.; Libich, D. S.; Stepper, J.; Loo, T.; Patchett, M. L.; Norris, G. E.; Pascal, S. M. *Biochemistry* **2011,** *50,* 2748–2755. doi:10.1021/bi200217u
26. Wang, H.; Oman, T. J.; Zhang, R.; Garcia De Gonzalo, C. V.; Zhang, Q.; van der Donk, W. A. *J. Am. Chem. Soc.* **2014,** *136,* 84–87. doi:10.1021/ja411159k
27. Wang, H.; van der Donk, W. A. *J. Am. Chem. Soc.* **2011,** *133,* 16394–16397. doi:10.1021/ja2075168
28. Xiao, H.; Wu, R. *Anal. Chem.* **2017,** *89,* 3656–3663. doi:10.1021/acs.analchem.6b05064
29. Brazier-Hicks, M.; Offen, W. A.; Gershater, M. C.; Revett, T. J.; Lim, E.-K.; Bowles, D. J.; Davies, G. J.; Edwards, R. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104,* 20238–20243. doi:10.1073/pnas.0706421104
30. Gantt, R. W.; Goff, R. D.; Williams, G. J.; Thorson, J. S. *Angew. Chem., Int. Ed.* **2008,** *47,* 8889–8892. doi:10.1002/anie.200803508
31. Xie, K.; Chen, R.; Li, J.; Wang, R.; Chen, D.; Dou, X.; Dai, J. *Org. Lett.* **2014,** *16,* 4874–4877. doi:10.1021/ol502380p
32. Chiu, H.-H.; Shen, M.-Y.; Liu, Y.-T.; Fu, Y.-L.; Chiu, Y.-A.; Chen, Y.-H.; Huang, C.-P.; Li, Y.-K. *Appl. Microbiol. Biotechnol.* **2016,** *100,* 4459–4471. doi:10.1007/s00253-015-7270-1
33. Haynes, L. J. *Naturally Occurring C-Glycosyl Compounds;* 1963; pp 227–258.
34. Franz, G.; Grün, M. *Planta Med.* **1983,** *47,* 131–140. doi:10.1055/s-2007-969972
35. Falcone Ferreyra, M. L.; Rodriguez, E.; Casas, M. I.; Labadie, G.; Grotewold, E.; Casati, P. *J. Biol. Chem.* **2013,** *288,* 31678–31688. doi:10.1074/jbc.M113.510040
36. Brazier-Hicks, M.; Evans, K. M.; Gershater, M. C.; Puschmann, H.; Steel, P. G.; Edwards, R. *J. Biol. Chem.* **2009,** *284,* 17926–17934. doi:10.1074/jbc.M109.009258
37. Gutmann, A.; Nidetzky, B. *Pure Appl. Chem.* **2013,** *85,* 1865–1877. doi:10.1351/pac-con-12-11-24
38. Nagatomo, Y.; Usui, S.; Ito, T.; Kato, A.; Shimosaka, M.; Taguchi, G. *Plant J.* **2014,** *80,* 437–448. doi:10.1111/tpj.12645
39. Gandia-Herrero, F.; Lorenz, A.; Larson, T.; Graham, I. A.; Bowles, D. J.; Rylott, E. L.; Bruce, N. C. *Plant J.* **2008,** *56,* 963–974. doi:10.1111/j.1365-313X.2008.03653.x
40. Chen, D.; Chen, R.; Wang, R.; Li, J.; Xie, K.; Bian, C.; Sun, L.; Zhang, X.; Liu, J.; Yang, L.; Ye, F.; Yu, X.; Dai, J. *Angew. Chem.* **2015,** *127,* 12869–12873. doi:10.1002/ange.201506505
41. Dürr, C.; Hoffmeister, D.; Wohlert, S.-E.; Ichinose, K.; Weber, M.; von Mulert, U.; Thorson, J. S.; Bechthold, A. *Angew. Chem., Int. Ed.* **2004,** *43,* 2962–2965. doi:10.1002/anie.200453758
42. Künzel, E.; Faust, B.; Oelkers, C.; Weissbach, U.; Bearden, D. W.; Weitnauer, G.; Westrich, L.; Bechthold, A.; Rohr, J. *J. Am. Chem. Soc.* **1999,** *121,* 11058–11062. doi:10.1021/ja9915347

43. Wang, F.; Zhou, M.; Singh, S.; Yennamalli, R. M.; Bingman, C. A.; Thorson, J. S.; Phillips, G. N. *Proteins: Struct., Funct., Bioinf.* **2013,** *81,* 1277–1282. doi:10.1002/prot.24289
44. Fischbach, M. A.; Lin, H.; Liu, D. R.; Walsh, C. T. *Proc. Natl. Acad. Sci. U. S. A.* **2005,** *102,* 571–576. doi:10.1073/pnas.0408463102
45. Bister, B.; Bischoff, D.; Nicholson, G. J.; Valdebenito, M.; Schneider, K.; Winkelmann, G.; Hantke, K.; Süssmuth, R. D. *BioMetals* **2004,** *17,* 471–481. doi:10.1023/B:BIOM.0000029432.69418.6a
46. Foshag, D.; Campbell, C.; Pawelek, P. D. *Biochim. Biophys. Acta, Proteins Proteomics* **2014,** *1844,* 1619–1630. doi:10.1016/j.bbapap.2014.06.010
47. Gutmann, A.; Nidetzky, B. *Angew. Chem., Int. Ed.* **2012,** *51,* 12879–12883. doi:10.1002/anie.201206141
48. Härle, J.; Günther, S.; Lauinger, B.; Weber, M.; Kammerer, B.; Zechel, D. L.; Luzhetskyy, A.; Bechthold, A. *Chem. Biol.* **2011,** *18,* 520–530. doi:10.1016/j.chembiol.2011.02.013
49. Bungaruang, L.; Gutmann, A.; Nidetzky, B. *Adv. Synth. Catal.* **2013,** *355,* 2757–2763. doi:10.1002/adsc.201300251
50. Gutmann, A.; Bungaruang, L.; Weber, H.; Leypold, M.; Breinbauer, R.; Nidetzky, B. *Green Chem.* **2014,** *16,* 4417–4425. doi:10.1039/C4GC00960F
51. Chen, D.; Sun, L.; Chen, R.; Xie, K.; Yang, L.; Dai, J. *Chem. – Eur. J.* **2016,** *22,* 5873–5877. doi:10.1002/chem.201600411
52. Buettner, F. F. R.; Ashikov, A.; Tiemann, B.; Lehle, L.; Bakker, H. *Mol. Cell* **2013,** *50,* 295–302. doi:10.1016/j.molcel.2013.03.003
53. Koshland, D. E., Jr. *Biol. Rev. Cambridge Philos. Soc.* **1953,** *28,* 416–436. doi:10.1111/j.1469-185X.1953.tb01386.x
54. Bojarová, P.; Křen, V. *Trends Biotechnol.* **2009,** *27,* 199–209. doi:10.1016/j.tibtech.2008.12.003
55. Burmeister, W. P.; Cottaz, S.; Rollin, P.; Vasella, A.; Henrissat, B. *J. Biol. Chem.* **2000,** *275,* 39385–39393. doi:10.1074/jbc.M006796200
56. Yip, V. L. Y.; Varrot, A.; Davies, G. J.; Rajan, S. S.; Yang, X.; Thompson, J.; Anderson, W. F.; Withers, S. G. *J. Am. Chem. Soc.* **2004,** *126,* 8354–8355. doi:10.1021/ja047632w
57. Rajan, S. S.; Yang, X.; Collart, F.; Yip, V. L. Y.; Withers, S. G.; Varrot, A.; Thompson, J.; Davies, G. J.; Anderson, W. F. *Structure* **2004,** *12,* 1619–1629. doi:10.1016/j.str.2004.06.020
58. Bourquelot, E.; Bridel, M. *Ann. Chim. Phys.* **1913,** *29,* 145–218.
59. Viebel, S. *Enzymologia* **1936,** *1,* 124–132.
60. Toone, E. J.; Simon, E. S.; Bednarski, M. D.; Whitesides, G. M. *Tetrahedron* **1989,** *45,* 5365–5422. doi:10.1016/S0040-4020(01)89487-4
61. Jahn, M.; Withers, S. G. *Biocatal. Biotransform.* **2003,** *21,* 159–166. doi:10.1080/10242420310001614351
62. David, B.; Iragues, R.; Jouanneau, D.; Daligault, F.; Czjzek, M.; Sanejouand, Y.-H.; Tellier, C. *ACS Catal.* **2017,** *7,* 3357–3367. doi:10.1021/acscatal.7b00348
63. Priyanka, P.; Parsons, T. B.; Miller, A.; Platt, F. M.; Fairbanks, A. J. *Angew. Chem., Int. Ed.* **2016,** *55,* 5058–5061. doi:10.1002/anie.201600817 and references herein.
64. Ly, H. D.; Withers, S. G. *Annu. Rev. Biochem.* **1999,** *68,* 487–522. doi:10.1146/annurev.biochem.68.1.487
65. Jahn, M.; Marles, J.; Warren, R. A. J.; Withers, S. G. *Angew. Chem., Int. Ed.* **2003,** *42,* 352–354. doi:10.1002/anie.200390114
66. Müllegger, J.; Jahn, M.; Chen, H.-M.; Warren, R. A. J.; Withers, S. G. *Protein Eng., Des. Sel.* **2005,** *18,* 33–40. doi:10.1093/protein/gzi003
67. Guillotin, L.; Richet, N.; Lafite, P.; Daniellou, R. *Biochimie* **2017,** *137,* 190–196. doi:10.1016/j.biochi.2017.03.020
68. Stick, R. V.; Stubbs, K. A. *Tetrahedron: Asymmetry* **2005,** *16,* 321–335. doi:10.1016/j.tetasy.2004.12.004
69. Kim, Y.-W.; Chen, H.; Kim, J. H.; Withers, S. G. *FEBS Lett.* **2006,** *580,* 4377–4381. doi:10.1016/j.febslet.2006.06.095
70. Kim, Y.-W.; Lovering, A. L.; Chen, H.; Kantner, T.; McIntosh, L. P.; Strynadka, N. C. J.; Withers, S. G. *J. Am. Chem. Soc.* **2006,** *128,* 2202–2203. doi:10.1021/ja057904a
71. Kim, Y.-W.; Chen, H.-M.; Kim, J. H.; Müllegger, J.; Mahuran, D.; Withers, S. G. *ChemBioChem* **2007,** *8,* 1495–1499. doi:10.1002/cbic.200700263
72. Tshililo, N. O.; Strazzulli, A.; Cobucci-Ponzano, B.; Maurelli, L.; Iacono, R.; Bedini, E.; Corsaro, M. M.; Strauss, E.; Moracci, M. *Adv. Synth. Catal.* **2017,** *359,* 663–676. doi:10.1002/adsc.201601091
73. Müllegger, J.; Chen, H. M.; Warren, R. A. J.; Withers, S. G. *Angew. Chem., Int. Ed.* **2006,** *45,* 2585–2588. doi:10.1002/anie.200503900
74. Chen, H.-M.; Withers, S. G. *Carbohydr. Res.* **2010,** *345,* 2596–2604. doi:10.1016/j.carres.2010.10.001
75. Armstrong, Z.; Withers, S. G. *Biopolymers* **2013,** *99,* 666–674. doi:10.1002/bip.22335
76. Almendros, M.; Danalev, D.; François-Heude, M.; Loyer, P.; Legentil, L.; Nugier-Chauvin, C.; Daniellou, R.; Ferrières, V. *Org. Biomol. Chem.* **2011,** *9,* 8371–8378. doi:10.1039/c1ob06227a

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Intramolecular glycosylation

Xiao G. Jia and Alexei V. Demchenko*

Review

Open Access

Address:
Department of Chemistry and Biochemistry, University of Missouri – St. Louis, One University Blvd., 434 Benton Hall (MC27), St. Louis, MO 63121, USA

Email:
Alexei V. Demchenko* - demchenkoa@umsl.edu

* Corresponding author

Keywords:
carbohydrates; glycosylation; intramolecular reactions; oligosaccharides

Beilstein J. Org. Chem. **2017,** *13,* 2028–2048.
doi:10.3762/bjoc.13.201

Received: 04 May 2017
Accepted: 13 September 2017
Published: 29 September 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

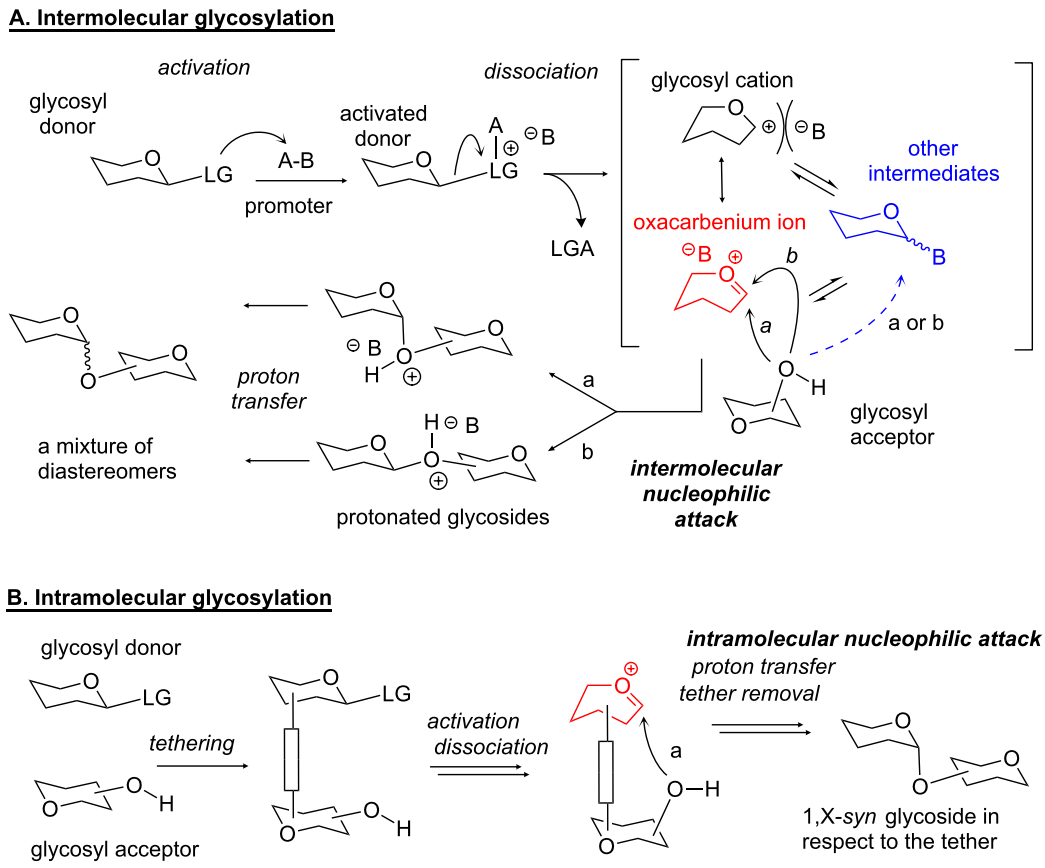
Abstract

Carbohydrate oligomers remain challenging targets for chemists due to the requirement for elaborate protecting and leaving group manipulations, functionalization, tedious purification, and sophisticated characterization. Achieving high stereocontrol in glycosylation reactions is arguably the major hurdle that chemists experience. This review article overviews methods for intramolecular glycosylation reactions wherein the facial stereoselectivity is achieved by tethering of the glycosyl donor and acceptor counterparts.

Introduction

With recent advances in glycomics [1,2], we now know that half of the proteins in the human body are glycosylated [3], and cells display a multitude of glycostructures [4]. Since glycan and glycoconjugate biomarkers are present in all body fluids, they offer a fantastic opportunity for diagnostics. Changes in the level of glycans, as well as changes in glycosylation and branching patterns, can indicate the presence and progression of a disease [5-9]. With a better understanding of functions of carbohydrates, the quest for reliable synthetic methods has launched, thus elevating the priority for improving our synthetic competences. The development of new methods for stereocontrolled glycosylation [10-14] in application to the expeditious synthesis of oligosaccharides represents a vibrant worldwide effort [15-32]. Nevertheless, despite extensive studies that have emerged since the very first experiments performed by Arthur Michael and Emil Fischer in the late 1800's, the glycosylation reaction remains challenging to chemists.

Enzymatic glycosylation reactions are highly stereoselective [33]. However, the stereocontrol of chemical glycosylation reactions remains cumbersome despite of significant advances. Common intermolecular glycosylation reactions in the absence of a participating auxiliary typically proceed with poor stereoselectivity. In these systems, there are no forces that are able to direct the glycosyl acceptor attack on the activated glycosyl donor that exists as a flattened oxacarbenium intermediate (Scheme 1a). Early attempts to achieve some stereocontrol of glycosylations were mainly dedicated to the development of participating groups and optimization of the reaction conditions. More recently, the research emphasis is switching towards



Scheme 1: The mechanistic outline of the intermolecular (a) and intramolecular (b) glycosylation reactions.

understanding of other, more fundamental factors and aspects of glycosylation. Extensive studies dedicated to conformation, configuration, stereoelectronics of the starting material, and key reaction intermediates have emerged [34-37].

Beside these attempts, an area of the intramolecular glycosylation has also been developed with an idea of providing higher efficiency of glycosylation reactions by bringing the reaction counterparts in a close proximity to each other. In many variations of this general concept, the intramolecular approach also allows for achieving better stereocontrol in comparison to that of an intermolecular reaction. This is usually credited to the facial selectivity for the glycosyl acceptor attack restricted by the tethering (Scheme 1b). However, the execution of this concept requires additional steps for the preparation of the tethered donor–acceptor combinations, and in some cases post-glycosylational modifications are also required. As a result, glycosylation that is already a four-step process (activation, dissociation, nucleophilic attack, proton transfer, Scheme 1a) has to be supplemented with additional manipulations that could lead to the decrease in over-all efficiency and yields. Hence, intramolecular glycosylations have a particular relevance to special cases of glycosylation or particularly challenging targets, such as 1,2-*cis* glycosides, where other, more direct methods fail to provide acceptable results.

Presented herein is an overview of methods that have been developed to achieve higher efficiency and/or better stereoselection by tethering the donor and acceptor counterparts, reactions that are commonly referred to as intramolecular glycosylations. A number of approaches for connecting the reaction counterparts, glycosyl donor and acceptor together, have been developed to provide the enhanced facial selectivity for the acceptor attack [38-41]. Beyond early intramolecular glycosylations achieved via the orthoester rearrangement by Lindberg [42] and Kochetkov [43], as well as the decarboxylation of glycosyl carbonates by Ishido [44], Barresi and Hindsgaul [45] are often credited for the invention of the intramolecular glycosylation in 1991. However, it is a pioneering albeit less known research by Kusumoto et al. in 1986 [46] that actually started the developments in this area. Of this general idea for the intramolecular glycosylation, three different concepts have been invented: a “molecular clamp” approach, intramolecular aglycone delivery (IAD), and leaving group-based methods (approaches A–C, Figure 1). This review will discuss recent developments in the field of intramolecular glycosylations with the

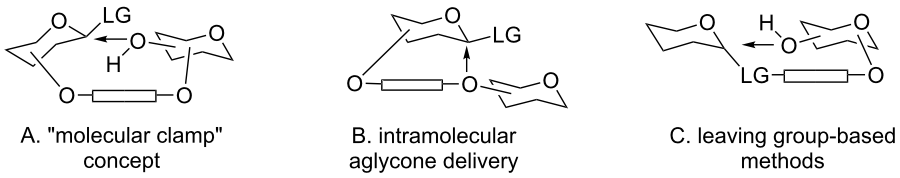


Figure 1: Three general concepts for intramolecular glycosylation reactions.

main emphasis on the developments of the past decade. A similar overview, albeit with the emphasis on molecular clamping, was presented as an introduction to the doctoral dissertation by Jia [47]. For previous developments in this area the reader should refer to a number of comprehensive overviews of intramolecular glycosylations in general [38-40] and IAD in particular [41,48-50].

Review

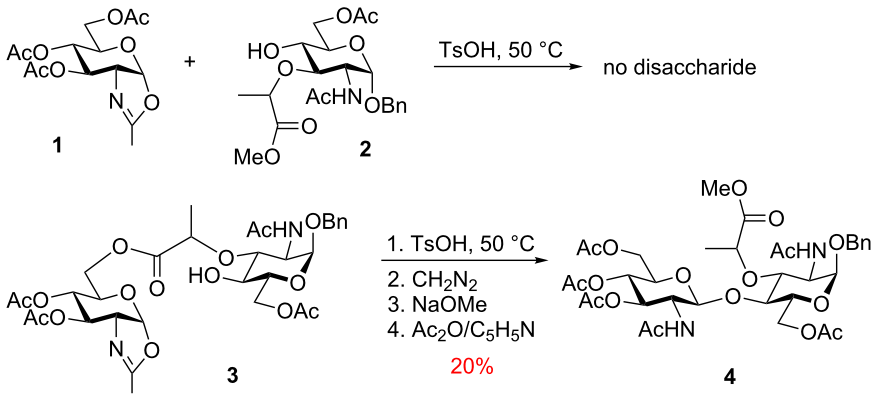
Molecular clamping method

Early developments

The "molecular clamp" concept (approach A, Figure 1) represents the first general concept for a intramolecular glycosylation strategy. The attachment of the glycosyl donor and acceptor via a tether takes place away from the reactive centers. These attachment strategies clearly distinguish the molecular clamp method from other intramolecular concepts wherein the attachment involves one of the reactive sites, acceptor hydroxy group in IAD or the leaving group of the donor. "Molecular clamping" was introduced by Kusumoto et al. [46], however, this term was coined by the same group much later [51]. We adopt this term to generally refer to this concept, which in other applications was also named "intramolecular glycosylation of prearranged glycosides" by Ziegler [52,53], "template-directed cyclo-glycosylation" by Valverde et al. [54], "remote glycosidation" by Takahashi [55] and "templated oligosaccharide synthesis" by Demchenko [56].

Initially introduced by Kusumoto et al. in 1986 [46], the molecular clamping clearly demonstrated the advantages that intramolecular glycosylations can offer. The first attempt to obtain a target disaccharide quipped with muramic acid from donor **1** and acceptor **2** failed (Scheme 2). The authors rationalized that *"... a novel device was required to facilitate the coupling. We thus tried to connect the two components prior to the glycosidation reaction with an ester linkage which can be formed more readily than a glycosidic bond. ... The glycosylation reaction then becomes an intramolecular process and hence could be expected to proceed more easily."* The authors then refer to a known phenomenon in the field of peptide chemistry *"where two components to be coupled had been brought close together by auxiliary groups."*

With this general idea in mind, and after *"examination of molecular models"* the authors created compound **3** that was tethered via the muramic acid moiety to the C-6 position of the donor that in their opinion was *"sterically most favorable for the formation of β(1→4) glycoside."* Indeed, after sequential glycosylation in the presence of TsOH at 50 °C, methanolysis, and per-acetylation, disaccharide **4** was isolated in 20% yield. The authors then very reasonably concluded that "*Consequently, the presence of the ester linkage which kept the two sugar moieties in close proximity to each other certainly favored the formation of the desired glycoside bond in the above experiment. Thus, this is the first example of the so-called*



Scheme 2: First intramolecular glycosylation using the molecular clamping.

"entropic activation" in glycosidation reaction." The authors have also projected that the *"entropic activation demonstrated in this work seems to have wide applicability..."* and disclosed their attempts to link the reaction counterparts with dicarboxylic acids. This served as an ultimate perspective on future developments in the field, but about a decade had passed before Ziegler resurrected this concept.

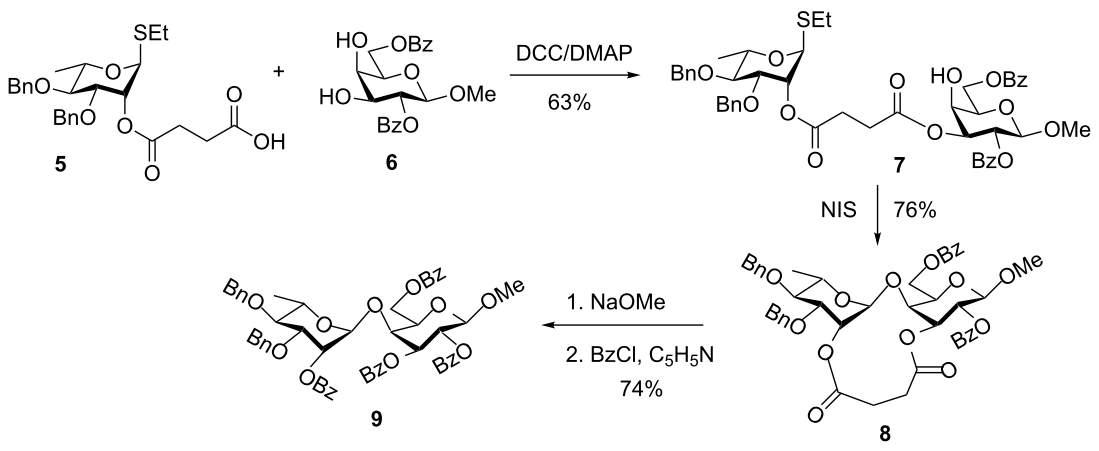
Flexible succinoyl and related tethers

Ziegler and co-workers investigated the use of a flexible succinoyl linker to link the glycosyl donor and acceptor counterpart. This reaction was named "intramolecular glycosylation of prearranged glycosides" [52,53]. Like in all "molecular clamp" applications, the tethering of the reaction counterparts takes place at positions not directly involving glycosylation sites: acceptor hydroxy group, like in the IAD or the donor leaving group, like in the leaving group-based approaches. In accordance with Ziegler's execution of this concept shown in Scheme 3, glycosyl donor **5** equipped with the succinoyl group at C-2 was coupled to the diol galactosyl acceptor **6** in the presence of DCC and DMAP. The resulting tether compound **7** was obtained in 63% yield. The intramolecular glycosylation of the latter gave cyclic compound **8** in 76% yield, which was sequentially deacylated and per-benzoylated to afford disaccharide **9** in 74% as a pure 1,2-*trans* isomer [52]. Expansion of this approach to other positions and sugar series showed that the stereoselectivity could be relaxed, and seemed to be dependent of the donor–acceptor match–mismatch. Thus, when succinoyl was attached to the 6-OH of the galactosyl acceptor, equal amounts of α- and β-anomers were obtained. Also, when a glucosyl acceptor was employed, mainly the 1,2-*cis*-linked product was obtained.

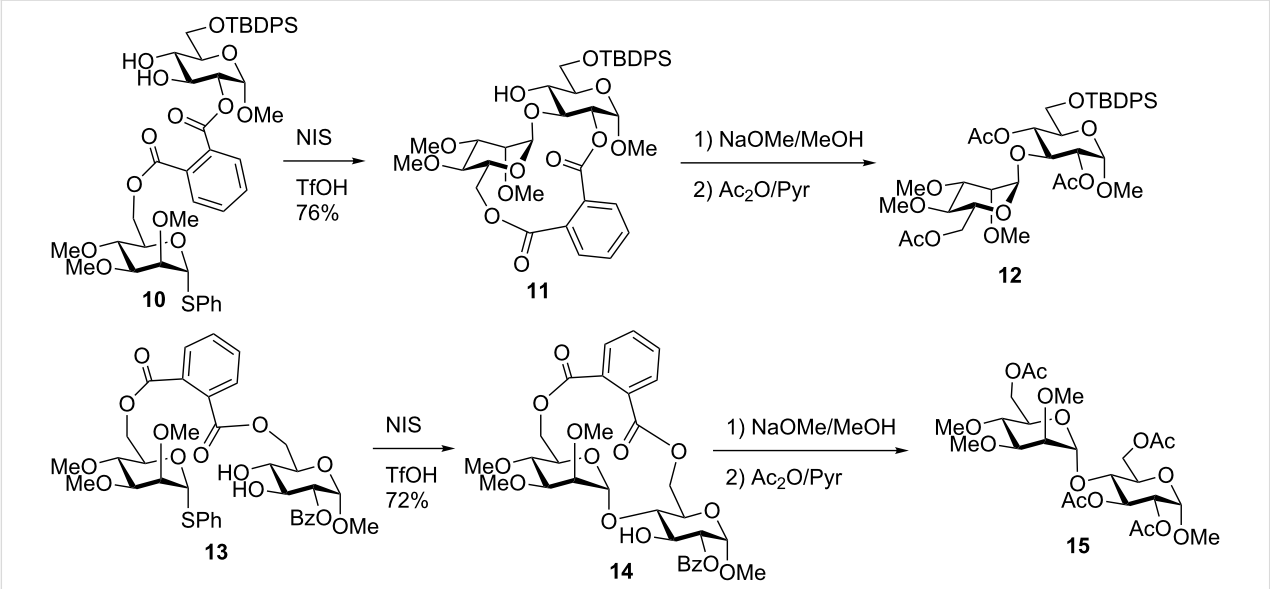
Valverde et al. also investigated succinoyl tethers [54], but their studies were mainly focusing on phthaloyl and non-symmetrical linkers described below. Among other flexible linkers investigated are carbonate [57], as well as oxalic [57], malonic [53,57,58], and glutaric [59] dicarboxylic acids. However, like in the case of succinoyl linkers, higher flexibility led to more relaxed stereoselectivity. Further variations upon this method involved the modification of the macrocycle ring size, torsional rigidity of the spacer, position of the attachment to both donor and acceptor, relative configuration of hydroxy groups, and the length of the linker [58,60-72]. Among early examples, xylylene and phthalimido linker showed very high efficiency, and will be highlighted below. Another early development discussed below is the peptide-templated synthesis. Beyond these influential early studies that led to further developments, this topic was comprehensively overviewed and for early developments the reader should refer to the original references and excellent comprehensive overviews of the topic [38,40]. It is a commonly accepted fact that the outcome of many glycosylations that fall under the general molecular clamp concept can be unpredictable. Therefore, practically every approach developed under this category was extensively studied and applied to a variety of sugar series and targets [58,73,74].

Phthaloyl and related tethers

Phthaloyl tethering was also introduced by Ziegler [53] and practically concomitantly by Valverde et al. [54] as "template-directed cyclo-glycosylation." In the latter application, glycosyl donor precursors were reacted with phthalic anhydride to afford the corresponding esters. The activation with thionyl chloride was used for tethering the donors to the glycosyl acceptor counterpart and the regioselectivity was controlled using tin-mediated coupling under microwave irradiation. The tethered compound **10** was then glycosylated in the presence of NIS/TfOH to afford compound **11** (Scheme 4). The tether was removed with NaOMe and the product was globally acetylated to afford **12** as an α-(1→3)-linked isomer. The regioselectivity in this case was



Scheme 3: Succinoyl as a flexible linker for intramolecular glycosylation of prearranged glycosides.



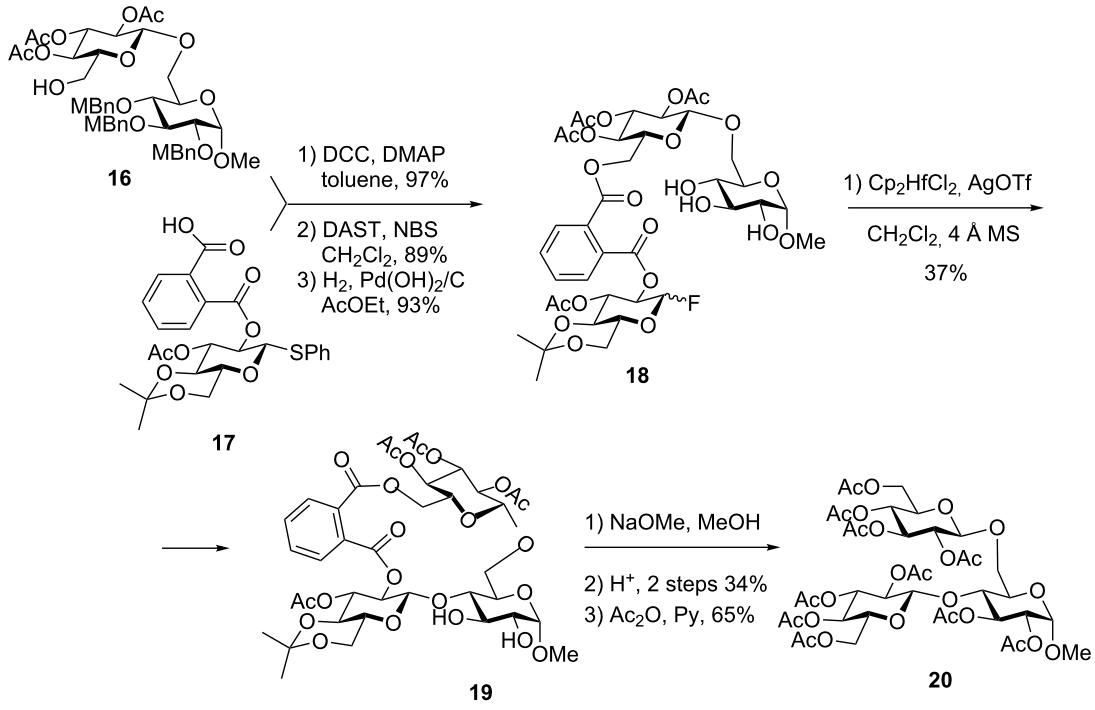
Scheme 4: Template-directed cyclo-glycosylation using a phthaloyl linker.

driven by the phthaloyl tether attachment to the neighboring C-2 position. In contrast, 6,6'-linked donor–acceptor pair **13** led to the formation of the (1→4)-linked regioisomer **15** [64]. Apparently, the rigid phthaloyl tether helps to achieve high regioselectivity because the anomeric center of the activated donor cannot easily reach out for hydroxy groups at remote positions.

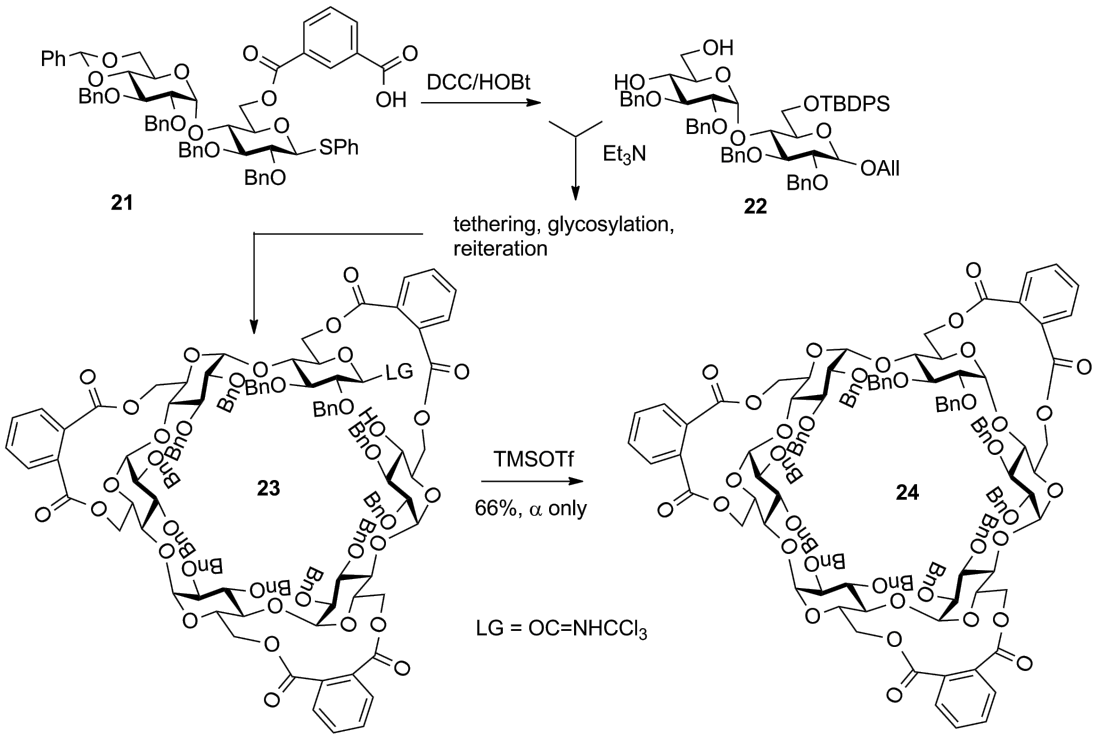
In other applications, such as in the glucosyl donor series, this application was less effective. For instance, relaxed regioselectivity was observed in cases when the phthaloyl linker was attached to the primary position of the acceptor [64]. Also, relaxed stereoselectivity was observed in case of glucosyl donors equipped with a non-participating group at C-2. Valverde at al. also investigated isophthalic tether, derived from benzene-1,3-dicarboxylic acid, and observed improved stereoselectivity in a number of applications [65]. The phthalimido tethering was further extended to a number of useful applications including the synthesis of branched structures by Takahashi and cyclodextrins by Fukase discussed below.

Thus, Takahashi et al. considered both flexible succinoyl and the rigid phthaloyl tether, but based on the outcome of the computational studies of relative conformations and energies chose the latter linker [55]. To apply the remote glycosidation methodology to the synthesis of the 4,6-branched trisaccharide, phthaloylated thioglycoside **17** was coupled with the 6-hydroxy group of the acceptor precursor **16** in the presence of DCC and DMAP (Scheme 5). The tethering was accomplished in 97% yield and the resulting conjugate was converted into glycosyl fluoride by the treatment with DAST and NBS in 89% yield. Finally, selective cleavage of *p*-methylbenzyl ethers was accomplished with H_2 over $Pd(OH)_2/C$ to provide donor–accepter conjugate **18** in 93% yield. Subsequent remote glycosidation of **18** was conducted in the presence of Cp_2HfC1_2 and AgOTf in CH_2C1_2 under reflux. The cyclized product **19** was obtained in 37% yield, the tether was removed with NaOMe, and the resulting free hydroxy groups were acetylated to afford the branched trisaccharide **20**.

The chemical synthesis of cyclodextrins is very challenging: controlling α-gluco stereoselectivity, and especially the final cyclization, represent a great challenge. For example, in Ogawa's synthesis of α-cyclodextrin the chain assembly was non-stereoselective and the cyclization was achieved in only 21% yield [75]. Kusumoto et al. clearly demonstrated the advantage of the molecular clamping in application to the synthesis of α-cyclodextrin (Scheme 6) [51]. The tethering was used to improve the selectivity during the stepwise chain elongation via the coupling of maltose building blocks **21** and **22**, as well as the efficiency of macrocyclization. The macrolactonization using the phthaloyl group clamp was accomplished using DCC and DMAP in refluxing 1,2-dichloroethane. A fairly high dilution (0.04 M) allowed to achieve the formation of the cyclic ester in 79% yield. This impressive yield was explained by the ability of the phthaloyl clamping groups to present the oligosaccharide chain in a favorable conformation for cyclization. After hydrolyzing the anomeric protecting group, several conditions were tried to close the ring and glycosylation with the trichloroacetimidoyl leaving group in **23** activated with trimethylsilyl triflate gave the desired α-linked product **24** in 66% yield [51].



Scheme 5: Phthaloyl linker-mediated synthesis of branched oligosaccharides via remote glycosidation.



Scheme 6: Molecular clamping with the phthaloyl linker in the synthesis of α-cyclodextrin.

Xylylene tether

Generally during glycosylation, it has been found that the more rigid the spacers, and smaller macrocycle formed, the more selective the reaction [63,69]. As an example of this approach, a rigid xylylene linker introduced by Schmidt [68], was successfully applied to the intramolecular synthesis of 1,2-*cis* glycosides with complete selectivity (Scheme 7) [69]. Thus, thioglycoside **25** is first alkylated at C-3 position. The resulting intermediate **26** is then used as the alkylating reagent to create a tether to acceptor **27** using tin-mediated primary alkylation to afford the tethered pair **28**.

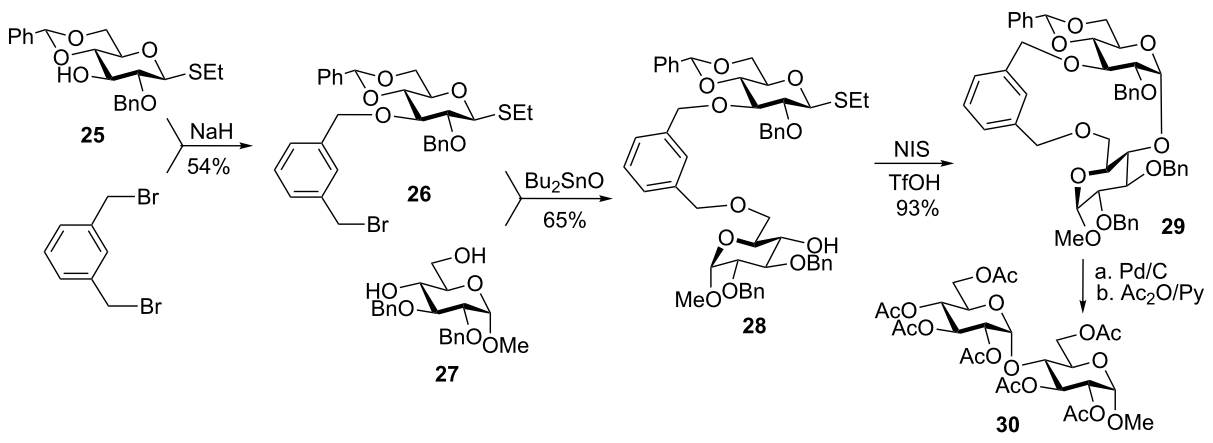
The latter is then intramolecularly glycosylated in the presence of NIS/TfOH in 93% yield and complete stereoselectivity. The resulting cyclic compound **29** is then subjected to concomitant xylylene tether removal and debenzylation followed by global acetylation to afford product **30**.

The extension of this approach to convergent oligosaccharide synthesis and reiterative sequencing in presented in Scheme 8. Thus, maltose and lactose disaccharide building blocks were linked via the xylylene tether, and the resulting compound **31** was glycosylated in the presence of NIS/TfOH to afford tetrasaccharide **32** in 78% as a pure β-diastereomer [70]. Schmidt demonstrated the usefulness of xylylene tethers in application to the iterative synthesis of maltotriose [70]. In this application, the xylylene tether was used to link two glucose derivatives via the 3'- and the 6-positions to create a tethered combination **33** (Scheme 8). NIS/TfOH was then applied to glycosylate the two sugar units to give disaccharide **34** in 84% yield (α/β = 85:15). Subsequent selective deprotection of the 6'-position, introduction of the new donor moiety **35** followed by liberating the hydroxy group at C-4' gave the tethered donor–acceptor combination **36**. After the NIS/TfOH-promoted glycosylation the desired trisaccharide **37** was obtained in 75% yield as a pure α-linked diastereomer. The per-acetylated maltotriose target was obtained after palladium-catalyzed hydrogenation that affected the removal of the template and all benzyl protecting groups followed by acetylation of the resulting hydroxy groups.

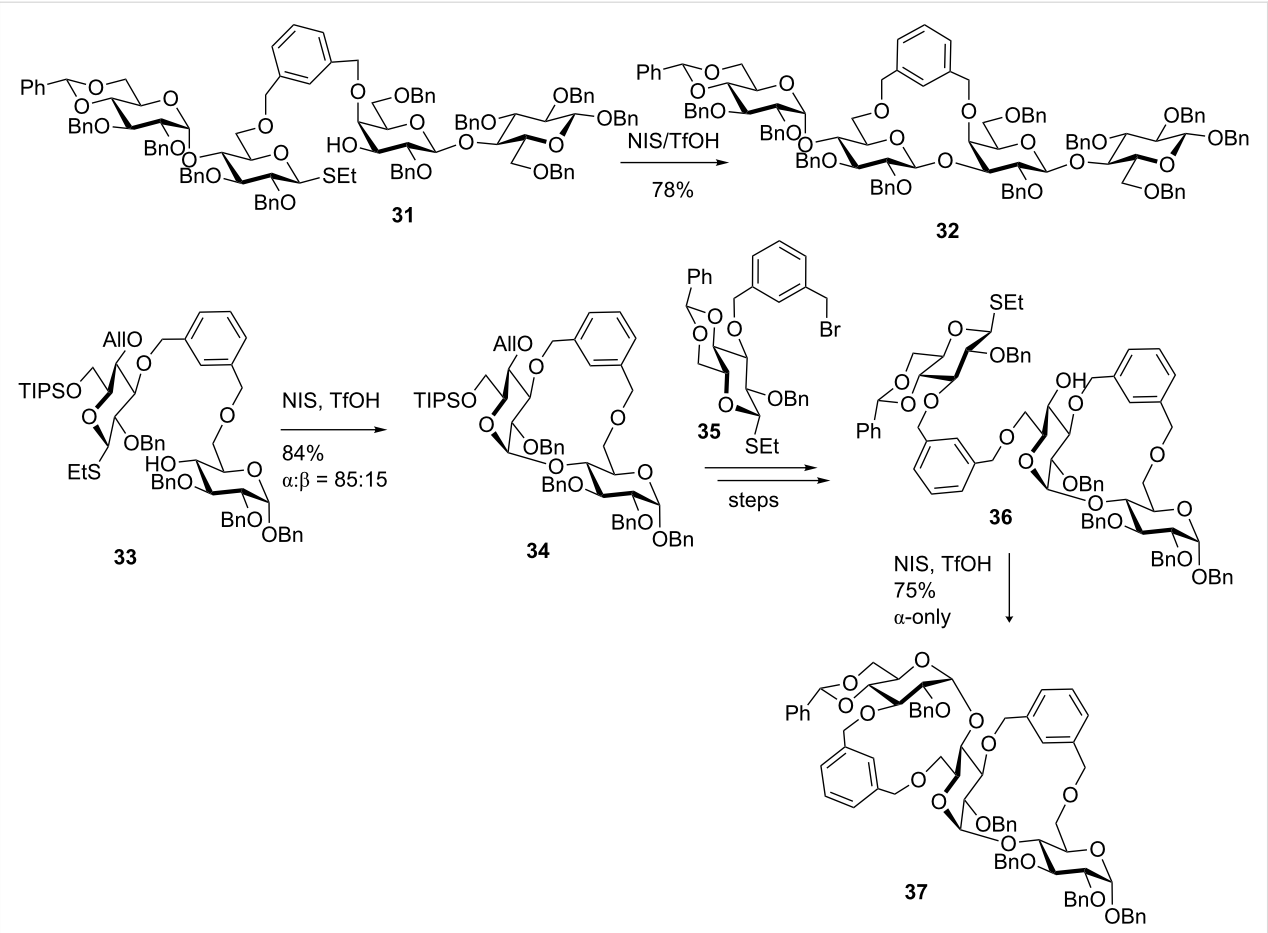
Peptide tether/template

Short peptide chains have also been investigated as templates for glycosylation. The general underpinning idea is to streamline the oligosaccharide synthesis and purification by using well developed peptide coupling reactions with or without the use of solid phase methods. To execute this concept, Fairbanks et al. investigated a number of peptide chains with various amino acids as templates (Scheme 9) [76,77]. Using DCC-mediated coupling reactions asparagine was attached both to a mannose donor and a trihydroxymannose acceptor, and the central amino acid unit(s) was varied. Intramolecular glycosylation was carried out with NIS/TfOH, resulting in a mixture of disaccharide products showing slight regioselectivity bias towards the formation of (1→3) linkages.

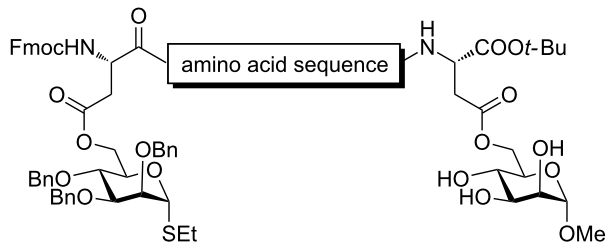
The stereoselectivity of these linkages can vary, but it was typically very relaxed perhaps due to a fairly low rigidity of this type of a template. Hence, further development of this methodology focused on solid-supported peptide templates [78]. For instance, Warriner and co-workers investigated a solid supported peptide sequence that was connected to the 6-hydroxy groups of the sugar units using carbonate linkages (Scheme 10) [79]. The hydroxyproline (Hyp, (2*S*,4*R*)-4-hydroxypyrrolidine-2-carboxylic acid) linked glycosyl donor and acceptor system failed to provide the product of the intramolecular glycosylation, probably due to steric interactions. A glycine residue spacer was found necessary to separate the two rigid Hyp bound counterparts. Thus, glycosylation of conjugate **38** in the presence of NIS and TMSOTf resulted in the formation of the (1→4)-linked disaccharide **40** in 80% yield with high α-selec-



Scheme 7: *m*-Xylylene as a rigid tether for intramolecular glycosylation.



Scheme 8: Oligosaccharide synthesis using rigid xylylene linkers.



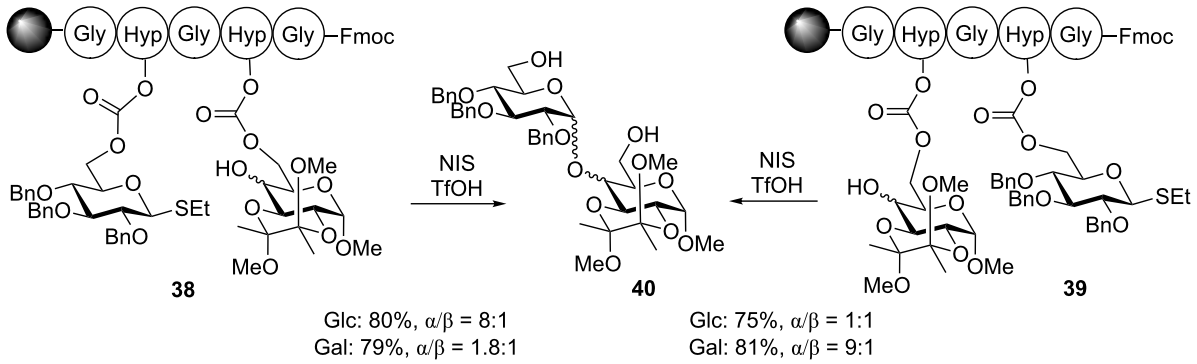
amino acid	α(1->2)	β(1->2)	α(1->3)	β(1->3)
none	–	13%	11%	23%
Gly	–	–	21%	20%
Phe	–	–	13%	18%
Pro	14%	16%	–	19%
GlyGly	–	–	–	56%

Scheme 9: Stereo- and regiochemical outcome of peptide-based linkers.

tivity (α/β = 8:1). Interestingly, when the donor and acceptor positions on the peptide were reversed, such as conjugate **39**, glycosidation of this compound produced disaccharide **40** in 75% yield albeit the stereoselectivity was entirely lost (α/β = 1:1). Galactosyl acceptors also showed a dramatic effect of the relative position of the donor and acceptor on the peptide sequence. Intriguingly, the stereoselectivity outcome was reversed (1.8:1 and 9:1) in comparison to glucosyl acceptors. When a similar concept was applied to mannosyl acceptor low 2:1 stereoselectivity was obtained regardless of the relative positioning of the reaction counterparts. This peptide-based templating was extended to the synthesis of a small library of disaccharides.

Non-symmetrical and other tethers

Non-symmetrical templates have also been developed with a general idea of achieving differentially cleavable attachments that could provide more flexibility in the synthesis of longer oligosaccharides [62,65]. Some representative examples of this general concept include benzyl–silicon tether [72], which is a

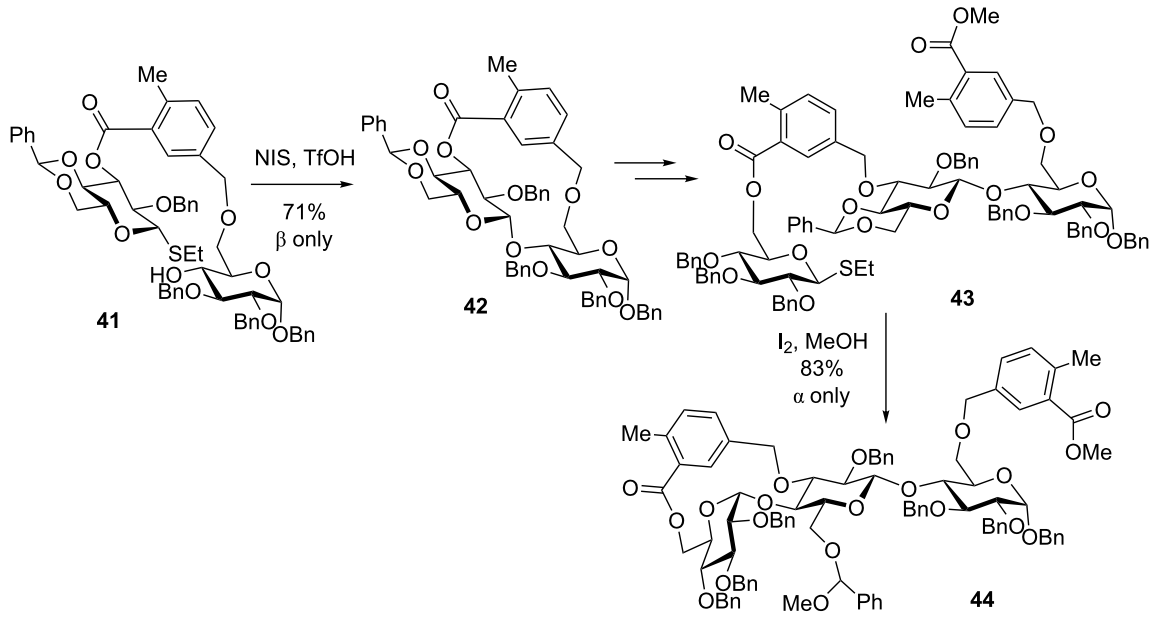


Scheme 10: Positioning effect of donor and acceptor in peptide templated synthesis.

hybrid approach to xylylene and a regular silicon [59] type of tethering. Another example of a non-symmetrical tethering strategy is benzyl–benzoyl hybrid tethering [72] that elaborated on xylylene and phthaloyl tethering approaches discussed above. Thus, this strategy was used in the synthesis of a trisaccharide through reiterative template-assisted synthesis (Scheme 11). Compound **41**, wherein the donor and acceptor counterparts were subjected to tethering via this rigid hybrid linker, was subjected to the NIS/TfOH-promoted glycosylation. The tether in the resulting disaccharide **42** could then be selectively opened with NaOMe. This leads to liberating only one hydroxy group (at C-3") that could be used for tethering with a glycosyl donor using a similar tethering concept to afford compound **43**.

The second glycosylation reaction is conducted in the presence of iodine in methanol. These conditions allow to cleave benzylidene groups concomitantly with the activation of the leaving group. As a result, the formation of the 14-membered ring is observed and compound **44** obtained in 83% yield with complete α-stereoselectivity. The ester part of the template is then cleaved with sodium methoxide in methanol revealing the 6"-hydroxy group that can be used for subsequent transformations [72].

In a recent attempt to simplify the synthesis of the non-symmetrical tethers, a highly trendy triazole-forming click chemistry was combined with rigid spacers by the Schmidt group. α,α'-Dibromo *ortho*- and *meta*-xylene-derived rigid spacers were used in this application, and this approach allowed to in-



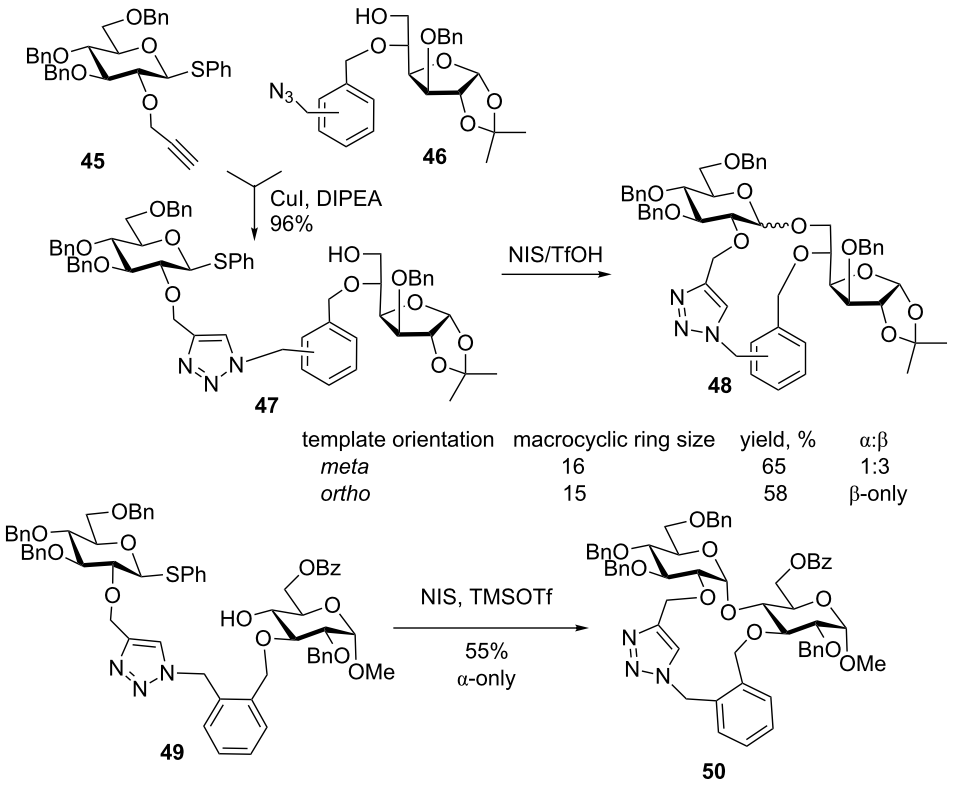
Scheme 11: Synthesis of a trisaccharide using a non-symmetrical tether strategy.

vestigate the size of the macrocycle formed during the glycosylation (Scheme 12) [80,81]. Thioglycoside donor **45** containing a 2-*O*-propargyl group and acceptor **46** with an azide-containing protecting group were connected using a click reaction to afford the tethered intermediate **47**. Upon treatment with NIS/TfOH, disaccharide **48** was obtained with complete β-selectivity when the *ortho*-xylyl group (15-membered ring) was used, versus α/β = 1:3 selectivity in the case of the *meta*-xylene linked counterpart [80]. As in the previous example with the xylylene-derived linker, the triazole linker was removed under standard hydrogenation conditions followed by global acetylation. The results obtained with the 6-hydroxyglucopyranosyl acceptor were somewhat mixed [81]. Attaching the template at various positions of the acceptor to achieve either 16- or 17-membered macrocycles resulted in high yields of 90% and 82%, respectively. However, the stereoselectivity of the reactions was modest, α/β = 3:1 and 1:2, respectively.

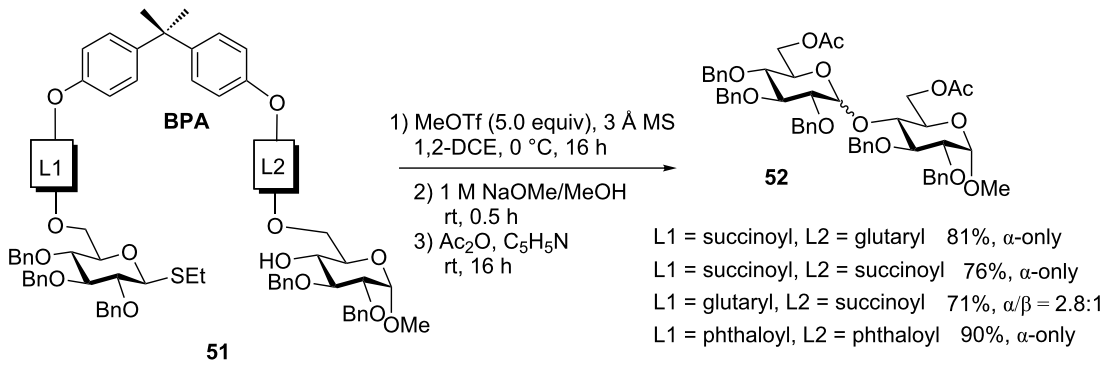
With the observation that selectivity can be influenced by the size of the macrocycle formed as a result of the intramolecular glycosylation, a tethered system linked via the O-3 position with the acceptor **49** was obtained (Scheme 12). Following the NIS/TMSOTf-promoted glycosylation, macrocycle **50** was formed in 55% yield with exclusive α-stereoselectivity. Interestingly, when a similar template was attached to the O-2 position followed by glycosylation with the 3-hydroxy group, the reaction proceeded with high β-selectivity. With the varying anomeric stereoselectivities and yields, it was hypothesized that the benzylic methylene group may be responsible for the increased rotational freedom between the triazoyl and benzyl moieties. Investigations with *o*-azidobenzyl protecting groups were used to reduce the degrees of freedom and also to form smaller ring sizes [81].

Templated oligosaccharide synthesis

Recently, Demchenko and co-workers introduced templated oligosaccharide synthesis, wherein bisphenol A (BPA) was used as the template and succinoyl, glutaryl or phthaloyl linkers were used to tether glycosyl donors and acceptors together [56,82]. The templated synthesis also falls into the general molecular clamping method. High stereoselectivity could be achieved with both flexible and rigid linkers (L1 and L2, Scheme 13). However, the use of the rigid BPA template core appears to be the key to ensure the high stereoselectivity because with flexible peptide core, no difference in stereoselectivity was detected. Thus, if linker L1 is shorter than L2, succinoyl vs glutaryl, respectively (or the same length, succinoyl) in compound **51**, the glycosyl acceptor counterpart is delivered from the bottom face



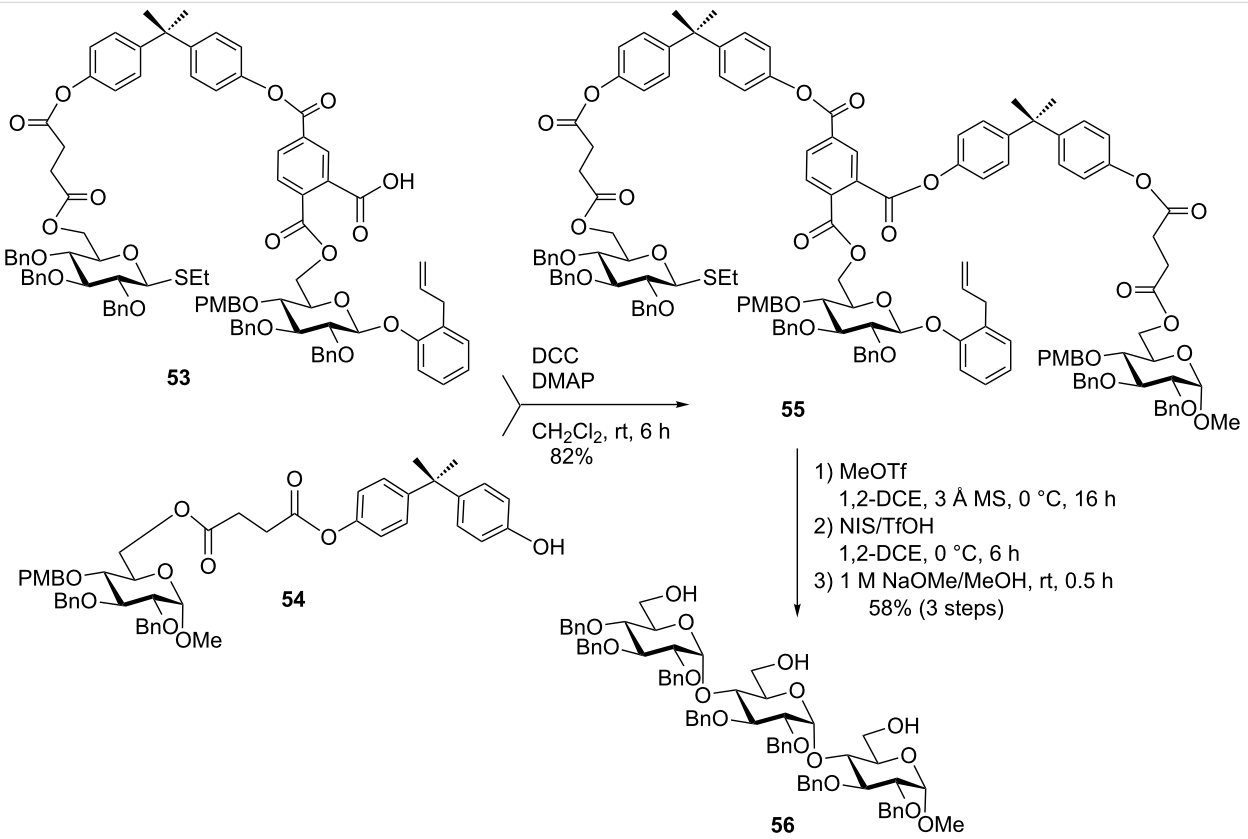
Scheme 12: Effect of ring on glycosylation with a furanose.



Scheme 13: Rigid BPA template with various linkers.

of the activated donor. These reactions produced the corresponding disaccharide **52** in 76–81% yields and complete α-stereoselectivity. Conversely, if linker L1 is longer than L2, glutaryl vs succinoyl, respectively, the stereoselectivity is lost (α/β = 2.8:1). Interestingly, the template effect is stronger than that of a participating solvent acetonitrile that was unable to favor β-anomers, like in intramolecular glycosylations. Instead, complete β-selectivity could be achieved using glycosyl donors equipped with the participating group at C-2.

A further mechanistic study of this work led to the appreciation of phthaloyl linkers leading to better yields, albeit complete α-selectivity [82]. To demonstrate the utility of the method a trisaccharide was synthesized using trimellitic anhydride as a precursor for the bridging linker (Scheme 14) [56]. The more flexible succinoyl linkers showed a clear advantage over more rigid phthaloyl linkers in terms of stereoselectivity and yields. Thus, a tethered donor-central unit conjugate **53** was coupled with the BPA-conjugated glycosyl acceptor **54** using DCC/DMAP-medi-



Scheme 14: The templated synthesis of maltotriose in complete stereoselectivity.

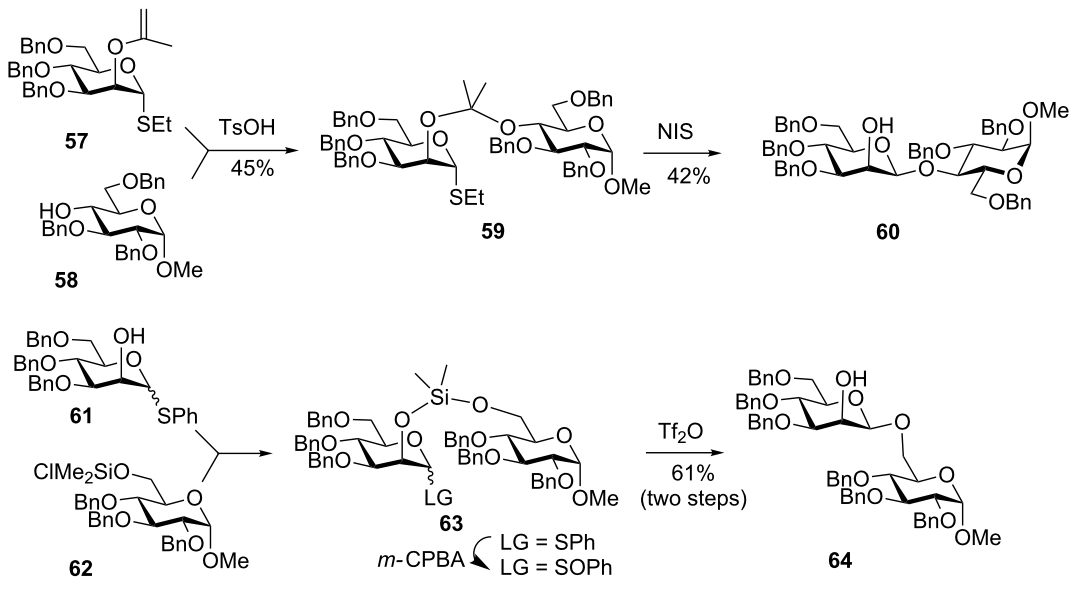
ated coupling reaction to obtain the templated conjugate of three monosaccharide units **55** in 82% yield. The selective activation of the *S*-ethyl leaving group in compound **55** was achieved with MeOTf and the glycosylation of the central building block took place with concomitant removal of the *p*-methoxybenzyl (PMB) group. The *o*-allylphenyl leaving group was activated with NIS/TfOH, and again the PMB group of the acceptor was removed during the glycosylation step. The resulting maltotriose **56** was then released from the template by reaction with NaOMe in MeOH [56].

Intramolecular aglycone delivery (IAD)

This approach was invented by Barresi and Hindsgaul [45] who named it intramolecular aglycone delivery (aglycon in the original literature) and it is commonly abbreviated as IAD (approach B, Figure 1). The distinctive characteristic of the IAD methods, and its major difference from other intramolecular approaches is the glycosyl donor which is tethered directly via the hydroxy group of the glycosyl acceptor to be glycosylated. In all other approaches, the acceptor is linked away from the hydroxy group that is to be glycosylated. The tethering site at the glycosyl donor can be either the neighboring C-2 position or a remote position. Barresi and Hindsgaul employed the activation of the thioethyl leaving group with *N*-iodosuccinimide, which resulted in excellent stereoselectivity for the synthesis of challenging β-mannoside [45,83]. Overall, this is a two-step process: first, formation of the intermolecular ketal between the donor and acceptor counterpart, and then glycosylation directly on the ketal oxygen of the glycosyl acceptor is performed. This was accomplished by the treatment of 2-isopropenylmannose **57** in the presence of TsOH (Scheme 15) to obtain mixed ketal **59**. The second step involved glycosidation in the presence of NIS that produced disaccharide **60** in 42% yield and complete β-selectivity. Despite fair yields during both the ketal formation and glycosylation stage, this excellent idea gave rise to the development of procedures that helped to evolve the IAD method into a very effective methodology. In particular, the implementation of silyl, allyl, and more recently, naphthylmethyl tethers helped to achieve significantly higher yields in comparison to those reported in the original protocols. Since the IAD has been overviewed multiple times [41,48-50], presented herein are only the basics as well as the key recent developments of this.

Stork and Bols independently demonstrated that silicon bridge-mediated aglycone delivery helps to enhance the yields while maintaining excellent stereocontrol [84,85]. For example, the Stork group used chlorodimethylsilyl protected acceptor precursor **62** for conjugation to the 2-hydroxy group of donor **61** as shown in Scheme 15. The thiophenyl leaving group of the tethered compound **63** was then oxidized into the corresponding sulfoxide with *m*-CPBA. The latter was glycosidated in the presence of Tf_2O to afford disaccharide **64** in complete stereoselectivity and a good yield of 61% over two steps (73% from the sulfoxide intermediate). This dimethylsilyl linker strategy was also applied towards the synthesis of α-glucosides by Bols [85].

Subsequently, the Bols group expanded the scope of the IAD method by investigating long-range tethering [39,85-89]. In this application the tether attachment was placed away from the



Scheme 15: First examples of the IAD.

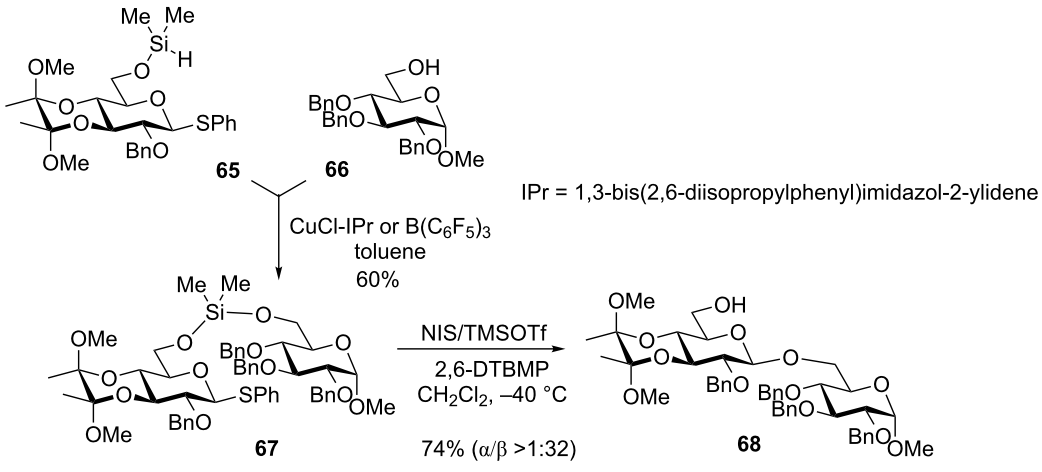
anomeric center offering a more flexible mode for obtaining either 1,2-*cis* or 1,2-*trans* linkages depending on the placement of the tether. While complete stereoselectivities were obtained with a ribofuranosyl donor tethered at C-5, application of the long range IAD towards glucopyranosides was less successful. Among a variety of attachment points, only tethering from the C-4 position showed some promise favoring the formation of the 1,4-*syn* products. Unfortunately, the IAD from the C-3 position afforded a mixture of diastereomeric glycosides, whereas tethering from the C-6 position gave predominantly the 1,6-anhydro product.

Following upon the early studies by Stork and Bols, Montgomery et al. further expanded the idea of the long range IAD via silicon tethering [90]. In the most recent report, they hypothesized that the conformational restriction of the pyranose should position the C-6 oxygen of the donor away from the developing oxacarbenium intermediate, thereby circumventing the formation of the cyclized product [91]. This was achieved by protecting the 3,4-*trans*-diol with a cyclic bis-ketal. Primary aliphatic alcohols underwent glycosylation very readily with donor **65** affording glycosides in excellent yields with high β-selectivity (>1:32). With primary glycosyl acceptors, such as **66** (Scheme 16), yields were slightly diminished due to the formation of the homocoupling products. Secondary alcohol acceptors were even less efficient showing a high substrate specificity of this approach. Other donor series including 2-azido and 2-deoxy sugars were investigated and provided similar results. This method was also applied towards the delivery of acceptors from the neighboring C-2 position [91]. This approach tolerated a much wider range of acceptors and showed excellent stereoselectivity with secondary acceptors providing high yields and complete stereoselectivities: α- for glucosides and β- for mannosides.

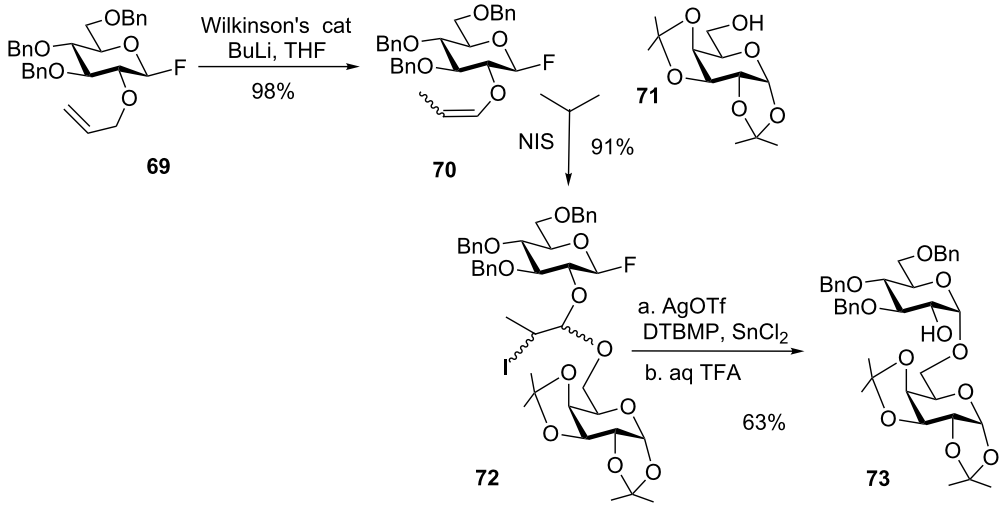
Another direction in the development of the IAD method emerged with the introduction of the allyl-mediated strategy by Fairbanks and co-workers who achieved improved yields and complete stereoselectivity in α-glucosylations and β-mannosylations [92]. In accordance with the linking strategy, the vinyl ether **70** was obtained in 98% yield from the corresponding 2-*O*-allyl ether **69** by the treatment with Wilkinson's catalyst and BuLi (Scheme 17) [93]. Subsequent NIS-mediated tethering of **70** and acceptor **71** gave the tethered donor–acceptor pair **72**. The latter was then intramolecularly glycosylated in the presence of silver triflate, tin(II) chloride, and 2,6-di-*tert*-butyl-4-methylpyridine (DTBMP). Finally, the tether was cleaved off using TFA to give pure 1,2-*cis* glycoside **73** in 63% yield over two steps.

An alternative linker was developed by Ito and Ogawa who implemented DDQ-mediated oxidative transformation of the *p*-methoxybenzyl (PMB) protecting group at the C-2 position of the donor into a tethering mixed acetal with a hydroxy group of the acceptor [94]. The early studies have successfully applied this PMB-based IAD method to the synthesis of a variety of oligosaccharides and glycoconjugates containing challenging β-mannosides [95,96]. A very impressive application of the IAD in polymer-supported reactions has also emerged [97]. Interestingly, the PMB tether was although used as the linker for the attachment to the polymer support. Bertozzi et al. investigated a similar concept based on 3,4-dimethoxybenzylidene tethering that was found superior in application to the synthesis of α,α-linked trehalose derivatives [98,99].

A major improvement of this approach has emerged with the implementation of a 2-napthylmethyl group as a tether group into this strategy [100]. This adjustment has allowed a greater range of hindered glycosyl acceptors to be tethered and glyco-

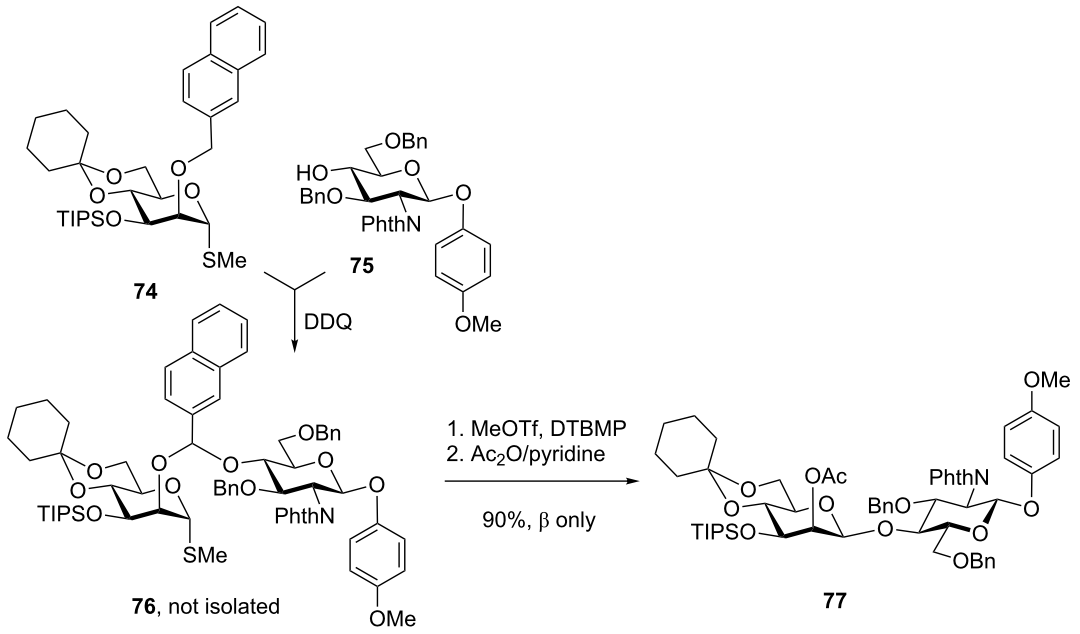


Scheme 16: Long range IAD via dimethylsilane.



Scheme 17: Allyl-mediated tethering strategy in the IAD.

sylated in high yields and stereoselectivity. The versatility of this approach lies in that it generally provides significantly higher yields in comparison to practically all previously developed IAD approaches. A representative example depicted in Scheme 18 shows the synthesis of disaccharide **77**, which clearly demonstrates that in terms of the over-all yields. This approach can even compete with direct intermolecular glycosylations while providing excellent stereoselectivity. Thus, mixed acetal **76** can be readily formed in 2 h by the addition of DDQ to a mixture of donor **74** and acceptor **75**. Without further purification, the latter mixture can be glycosylated in the presence of MeOTf and DTBMP followed by acetylation to give disaccharide **77** in an excellent yield of 90% and complete β-selectivity [100]. Initially investigated for the synthesis of β-mannosides, α-glucosides, and β-arabinofuranosides [100], this approach was extended to the synthesis of β-rhamnosides [101] and many other challenging linkages and targets [41,102-108].

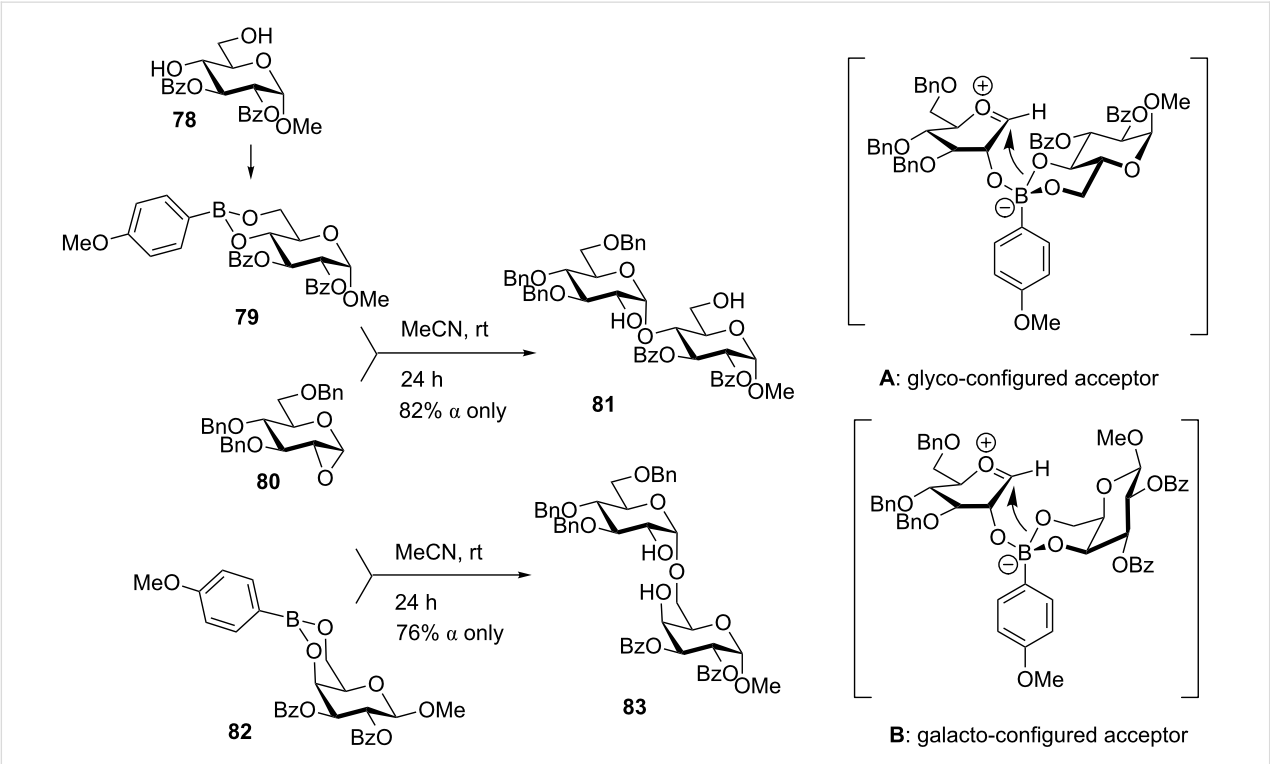


Scheme 18: IAD using tethering via the 2-naphthylmethyl group.

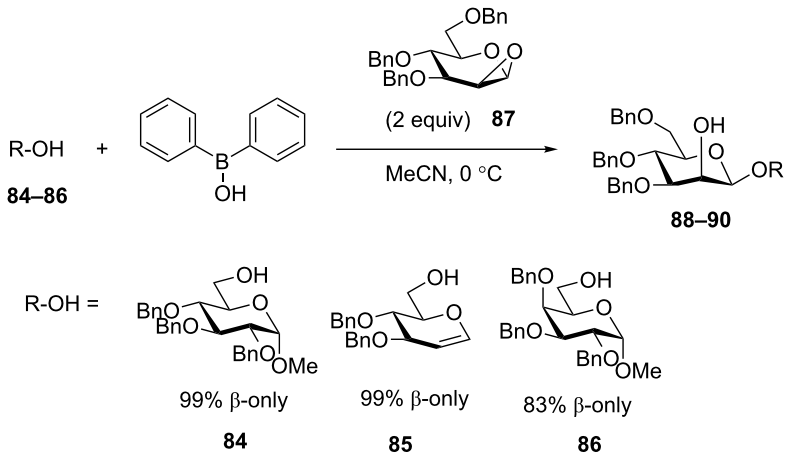
Arylboronic esters have recently been probed by Toshima and co-workers as a successful linkage for the IAD method [109]. The arylboronic sugar derivatives, such as **79**, can be easily obtained from the corresponding 4,6-diol **78** and a arylboronic acid in toluene at reflux (Scheme 19). Boronic ester **79** was then reacted with 1,2-anhydro donor **80**. It was assumed that this reaction proceeds via the oxacarbenium ion tethered to a tetracoordinated boronate ester. The subsequent glycosylation then proceeds regioselectively from the less-hindered boron–oxygen bond (see intermediate **A**). In this case, where gluco-configured acceptor **78** was used the (1→4)-linked product **81** was formed exclusively in 82% yield with high α-selectivity. Similarly, when mannose, glucosamine, and glucal were used as glycosyl acceptors, the 1→4 linkage was formed exclusively with high α-selectivity in 92%, 77%, and 72% yield, respectively. Conversely, the galacto-configured boronic ester acceptor **82** was used, the α-(1→6)-linked product **83** was formed in 70% yield. Again, the regioselectivity of glycosylation is driven by the less-hindered boron–oxygen bond, which is from C-6 face in the case of galactose (intermediate **B,** Scheme 19). In the case of other acceptors: a 3,4-diol of galactose gave the α-(1→4) linkage predominantly (65%) while a 2,3-diol of mannoside led to the α-(1→3)-linked disaccharide in 70% yield.

This method has recently found a valuable extension to the synthesis of β-mannosides [110]. Thus, diphenylborinic acid-derived glycosyl acceptors **84–86** were reacted with 1,2-anhydromannosyl donor **87** (Scheme 20). The tethered oxacarbenium ion intermediate then directs the nucleophilic attack intramolecularly to the β-face of the mannosyl donor. As a result, disaccharides **88–90** were obtained in 83–99% yields and exclusive β-manno stereoselectivity. Advantages of this methodology have been tested in application to the synthesis of a tetrasaccharide repeating unit of lipopolysaccharide derived from *E. coli* O75 [111].

Demchenko and co-workers introduced the use of the picolinyl group at the neighboring C-2 position of glycosyl donors as an arming participating group [112,113]. These glycosylations provided complete 1,2-*trans* stereoselectivity, *anti* with respect to the orientation of the picolinyl group. When the picolinyl ether or picoloyl ester group was placed at remote positions, glycosylations occurred *syn* with respect to the orientation of the picolinyl/picoloyl group [114]. The stereoselectivity was explained by the occurrence of the hydrogen bonding between the hydroxy group of glycosyl acceptor (NuH) and the nitrogen atom of the picolinyl/picoloyl group. Subsequently, the glycosyl acceptor is delivered towards the oxacarbenium ion from the same face (*syn*) as the picolinyl/picoloyl group (Figure 2). This method, named H-bond-mediated aglycone delivery (HAD), has been applied towards the synthesis of α-glucosides [114-116], α-galactosides, β-rhamnosides [114], and β-mannosides [117].



Scheme 19: Origin of selectivity in boronic ester mediated IAD.



Scheme 20: Arylborinic acid approach to the synthesis of β-mannosides.

The latter approach was extended to the synthesis of a β-mannotrisaccharide, wherein complete β-manno selectivity was obtained at room temperature [117]. A useful extension of this method to glycosyl donors with switchable selectivity has also been disclosed by the Demchenko group [118,119].

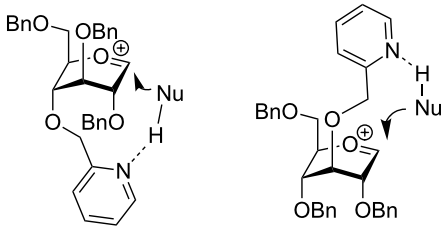
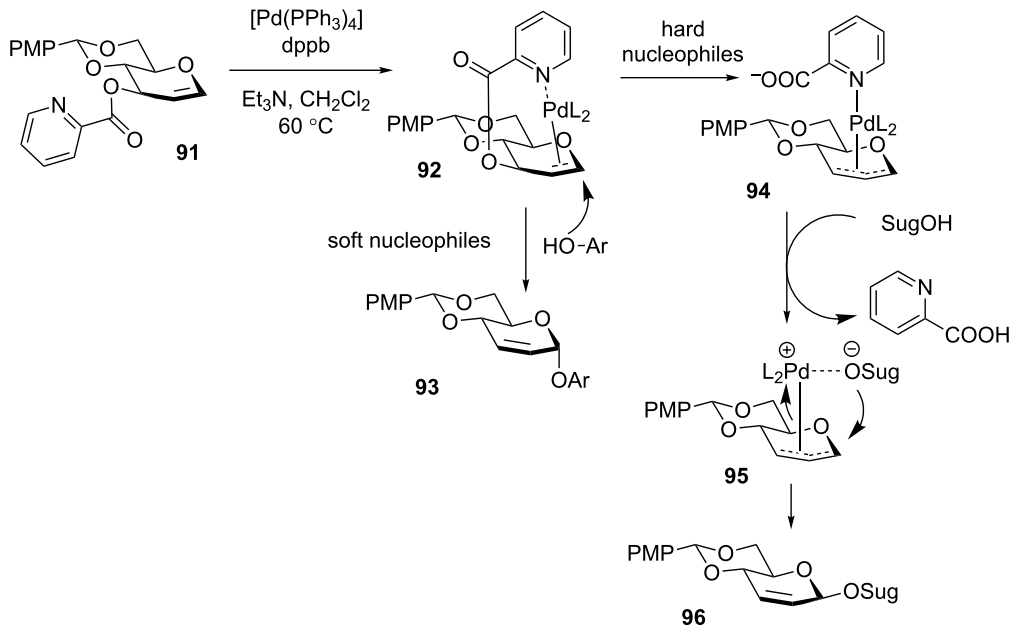


Figure 2: Facial selectivity during HAD.

Not being strictly intramolecular, the HAD method led to a variety of other delivery methods, which included the realm of organometallics. For instance, Liu and co-workers have developed the use of catalytic palladium to control the stereoselectivity in glycosylations via a palladium π-allyl intermediate. Previously, the application of this technique to glycosidic bond formation has been hampered by the difficulty in the formation of the palladium π-allyl intermediates and their poor reactivity in the electron-rich glycal systems [120]. To overcome this challenge the Liu group explored the application of palladium π-allyl intermediates to *O*-glycosylation through the use of a picoloyl group to direct palladium binding at the C-3 position [121]. Glycosylation results are indicative of two reaction pathways with differing in the selectivity outcome based on the hard/soft properties of the nucleophiles. In both pathways, the first step involves picoloyl group-directed coordination of palladium from the top β-face of the 1,2-dehydro donor **91** to form intermediate **92** (Scheme 21). With softer nucleophiles, such as phenol (ArOH), the nucleophilic attack is directed away from the steric bulk of the palladium to give α-glycosides **93**. When the acceptor is a hard nucleophile, such as a sugar alcohol (SugOH), the picoloyl group is displaced to generate the π-allyl complex **94**. The harder nucleophiles then tend to coordinate to palladium via intermediate **95**, followed by intramolecular nucleophilic delivery to form β-anomer **96**. Both primary and secondary sugar acceptors worked well providing disaccharides with high β-selectivity and good yields. Overall, compounds **93** and **96**, obtained as a result of this interesting reaction, represent products of the Ferrier rearrangement, 2,3-dehydro derivatives.

Leaving group-based methods

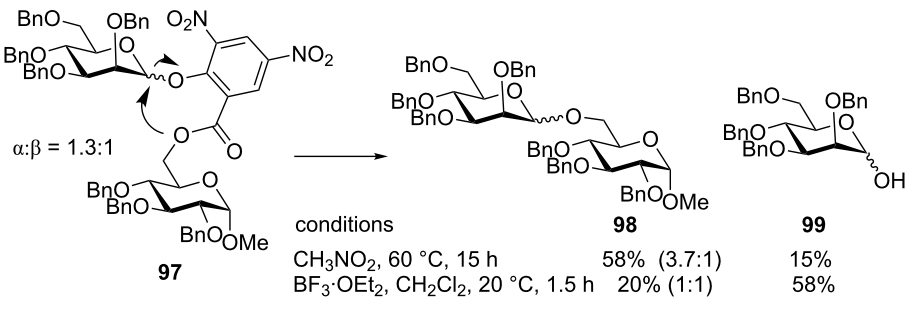
This overview continues with the discussion of the leaving group-based tethering concept (approach C, Figure 1). As the name of the concept implies, the glycosyl acceptor is linked (away from the glycosylation site) to the leaving group of the glycosyl donor. The first examples of this type of intramolecular glycosylation was based on the 1,2-orthoester rearrangement by Lindberg [42] and Kochetkov [43], as well as the decarboxylation of glycosyl carbonates by Ishido [44]. Intramolecular glycosylations where the glycosyl acceptor was purposefully attached directly to the leaving group of the glycosyl donor have been introduced by the Schmidt group [122]. The applicability of these techniques is still relatively unexplored, yet, it has been proposed that these reactions tend to be intermolecular rather than intramolecular [123,124]. Subsequent studies involved the exploration of various reaction conditions [125,126], and the investigation of other leaving groups [123,124,127].



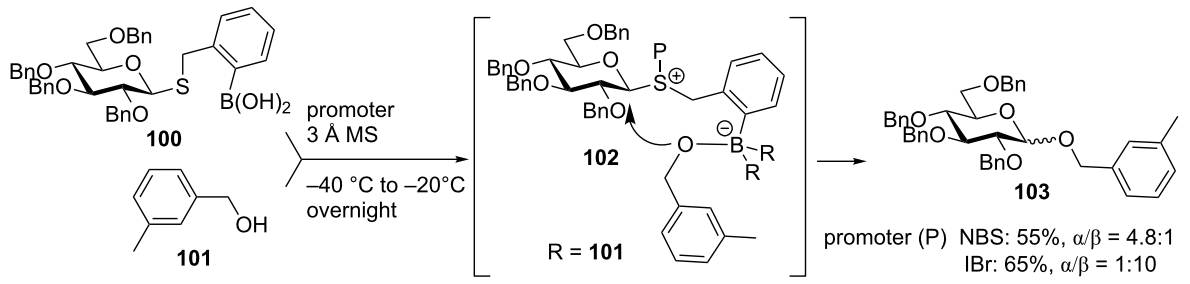
Scheme 21: Possible mechanisms to explain α and β selectivity in palladium mediated IAD.

For instance, Jensen et al. developed the methyl 3,5-dinitrosalicylate (DISAL) anomeric leaving group that could be used as a platform for linking the glycosyl acceptor in place of the methyl ester [128]. Glycosylation of conjugate **97** wherein glycosyl acceptor was linked via an ester bond at the *ortho*-position of the DISAL leaving group of the donor gave best results under elevated temperatures. Thus, mannoside **98** was obtained in 58% yield with modest stereoselectivity (Scheme 22). The yields are hampered by the competing formation of the hemiacetal product **99**. Crossover experiments with 1,2:5,6-di-*O*-isopropylidene-α-D-glucofuranose acceptor showed only disaccharides resulting from the intramolecular glycosylation. However, when crossover experiments with cyclohexanol were conducted, the intermolecularly formed cyclohexyl glycoside was found to be the major product (5.2 to 1) compared to the intramolecular glycosylation product. The addition of Lewis acids helps to reduce the reaction time and the temperature required, but also increases the formation of hydrolysis products and reduces overall stereoselectivity.

Recently, Liu et al. explored the use of *ortho*-dihydroxyboryl-substituted benzyl thioglycosides as a delivery method for the leaving group-based intramolecular glycosylation [129]. They hypothesized that if boronic acid-derived donor **100** is activated in the presence of glycosyl acceptor **101,** the boronic ester **102** would form as the key intermediate. Upon dissociation of the anomeric C–S bond of the sulfonium intermediate **102**, an oxygen nucleophile on the boronate ester would attack the C-1 center on the opposite side resulting in **103** with good stereoselection (Scheme 23). Initial trials with 3-methylbenzyl alcohol showed good selectivity (α/β = 4.8:1) when boronic acid and NBS were employed. Control experiments with a thio-



Scheme 22: DISAL as the leaving group that favors the intramolecular glycosylation pathway.



Scheme 23: Boronic acid as a directing group in the leaving group-based glycosylation method.

phenyl or a thiobenzyl leaving group showed lower stereoselectivities and a slight reduction in yields. The addition of triflic acid or silver triflate resulted in a significant reduction of stereoselectivity, so further trials were done in the absence of metal or acid reagents. Surprisingly, when IBr was used as a promoter the selectivity reversed resulting in the formation of glycoside **103** in 65% yield and high β-stereoselectivity (α/β = 1:10). The selectivity also reverses when the reaction is carried out in the presence of a coordinating solvent, for example, a similar reaction performed in acetonitrile delivers glycoside **103** in 51% yield (α/β = 1:4). When using less than three equivalents of acceptor to donor ratio, the yield drastically drops giving evidence that the borate intermediate plays an important role in the stereoselection.

Conclusion

Intramolecular glycosylation has seen dramatic advancements in the past two decades. New tethers, templates and conditions have advanced the synthesis of challenging glycosidic linkages. A more streamlined synthesis of starting materials has also made these methodologies more attractive for use in more complicated multistep syntheses. Despite the advancements made, there are still no definitive rules on why small changes may affect the stereochemical outcomes so dramatically. There is a greater need to study the underlying concepts and rules governing the use of tethers and templates and how to apply them to new systems and targets.

Acknowledgements

The authors are indebted to the National Institute of General Medical Sciences (GM111835 and GM120673) for support of their work on the development of new methods and strategies for oligosaccharide synthesis.

References

1. Cummings, R. D.; Pierce, J. M. *Chem. Biol.* **2014,** *21,* 1. doi:10.1016/j.chembiol.2013.12.010
2. Russo, L.; Cipolla, L. *Chem. – Eur. J.* **2016,** *22,* 13380. doi:10.1002/chem.201602156
3. Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Bertozzi, C. R.; Stanley, P.; Hart, G. W.; Etzler, M. E. *Essentials of Glycobiology,* 2nd ed.; CSH Laboratory Press: New York, 2009.
4. Alonzi, D. S.; Neville, D. C. A.; Lachmann, R. H.; Dwek, R. A.; Butters, T. D. *Biochem. J.* **2008,** *409,* 571. doi:10.1042/BJ20070748
5. Lebrilla, C. B.; An, H. J. *Mol. BioSyst.* **2009,** *5,* 17. doi:10.1039/B811781K
6. Adamczyk, B.; Tharmalingam, T.; Rudd, P. M. *Biochim. Biophys. Acta* **2012,** *1280,* 1347. doi:10.1016/j.bbagen.2011.12.001
7. Chaubard, J.-L.; Krishnamurthy, C.; Yi, W.; Smith, D. F.; Hsieh-Wilson, L. C. *J. Am. Chem. Soc.* **2012,** *134,* 4489. doi:10.1021/ja211312u
8. Fujitani, N.; Furukawa, J.-i.; Araki, K.; Fujioka, T.; Takegawa, Y.; Piao, J.; Nishioka, T.; Tamura, T.; Nikaido, T.; Ito, M.; Nakamura, Y.; Shinohara, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 2105. doi:10.1073/pnas.1214233110
9. Ruhaak, L. R.; Miyamoto, S.; Lebrilla, C. B. *Mol. Cell. Proteomics* **2013,** *12,* 846. doi:10.1074/mcp.R112.026799
10. Zhu, X.; Schmidt, R. R. *Angew. Chem., Int. Ed.* **2009,** *48,* 1900. doi:10.1002/anie.200802036
11. McKay, M. J.; Nguyen, H. M. *ACS Catal.* **2012,** *2,* 1563. doi:10.1021/cs3002513
12. Nigudkar, S. S.; Demchenko, A. V. *Chem. Sci.* **2015,** *6,* 2687. doi:10.1039/C5SC00280J
13. Das, R.; Mukhopadhyay, B. *ChemistryOpen* **2016,** *5,* 401. doi:10.1002/open.201600043
14. Li, X.; Zhu, J. *Eur. J. Org. Chem.* **2016,** 4724. doi:10.1002/ejoc.201600484
15. Boons, G.-J. *Tetrahedron* **1996,** *52,* 1095. doi:10.1016/0040-4020(95)00897-7
16. Fraser-Reid, B.; Madsen, R.; Campbell, A. S.; Roberts, C. S.; Merritt, R. J. In *Bioorganic Chemistry: Carbohydrates;* Hecht, S. M., Ed.; Oxford University Press: New York - Oxford, 1999; p 89.
17. Davis, B. G. *J. Chem. Soc., Perkin Trans. 1* **2000,** 2137. doi:10.1039/a809774g
18. Seeberger, P. H.; Haase, W.-C. *Chem. Rev.* **2000,** *100,* 4349. doi:10.1021/cr9903104
19. Tanaka, H.; Yamada, H.; Takahashi, T. *Trends Glycosci. Glycotechnol.* **2007,** *19,* 183. doi:10.4052/tigg.19.183
20. Seeberger, P. H. *Chem. Soc. Rev.* **2008,** *37,* 19. doi:10.1039/B511197H
21. Boltje, T. J.; Buskas, T.; Boons, G.-J. *Nat. Chem.* **2009,** *1,* 611. doi:10.1038/nchem.399

22. Parameswar, A. R.; Demchenko, A. V. In *Progress in the synthesis of complex carbohydrate chains of plant and microbial polysaccharides;* Nifantiev, N. E., Ed.; Transworld Res. Network: Kerala, 2009; pp 463 ff.
23. Smoot, J. T.; Demchenko, A. V. *Adv. Carbohydr. Chem. Biochem.* **2009,** *62,* 161. doi:10.1016/S0065-2318(09)00005-5
24. Fraser-Reid, B.; López, J. C. *Topics in current chemistry;* Springer-Verlag: Berlin-Heidelberg, 2011; Vol. 301.
25. Kaeothip, S.; Demchenko, A. V. *Carbohydr. Res.* **2011,** *346,* 1371. doi:10.1016/j.carres.2011.05.004
26. Tanaka, K.; Fukase, K. In *Solid-Phase Organic Synthesis;* Toy, P. H.; Lam, Y., Eds.; John Wiley & Sons, Inc.: Hoboken, 2012; pp 489 ff.
27. Crotti, S.; Adamo, R. *Curr. Org. Synth.* **2013,** *10,* 501. doi:10.2174/15701794113109990056
28. Yang, L.; Qin, Q.; Ye, X.-S. *Asian J. Org. Chem.* **2013,** *2,* 30. doi:10.1002/ajoc.201200136
29. Yasomanee, J. P.; Demchenko, A. V. *Trends Glycosci. Glycotechnol.* **2013,** *25,* 13. doi:10.4052/tigg.25.13
30. Bennett, C. S. *Org. Biomol. Chem.* **2014,** *12,* 1686. doi:10.1039/c3ob42343c
31. Bouhall, S. K.; Sucheck, S. J. *J. Carbohydr. Chem.* **2014,** *33,* 347. doi:10.1080/07328303.2014.931964
32. Pistorio, S. G.; Stine, K. J.; Demchenko, A. V. In *Carbohydrate Chemistry: State-of-the-art and challenges for drug development;* Cipolla, L., Ed.; Imperial College Press: London, 2016; pp 247 ff. doi:10.1142/9781783267200_0010
33. Muthana, S.; Cao, H.; Chen, X. *Curr. Opin. Chem. Biol.* **2009,** *13,* 573. doi:10.1016/j.cbpa.2009.09.013
34. Mydock, L. K.; Demchenko, A. V. *Org. Biomol. Chem.* **2010,** *8,* 497. doi:10.1039/B916088D
35. Crich, D. *Acc. Chem. Res.* **2010,** *43,* 1144. doi:10.1021/ar100035r
36. Frihed, T. G.; Bols, M.; Pedersen, C. M. *Chem. Rev.* **2015,** *115,* 4963. doi:10.1021/cr500434x
37. Bohé, L.; Crich, D. *Carbohydr. Res.* **2015,** *403,* 48. doi:10.1016/j.carres.2014.06.020
38. Jung, K.-H.; Müller, M.; Schmidt, R. R. *Chem. Rev.* **2000,** *100,* 4423. doi:10.1021/cr990307k
39. Madsen, J.; Bols, M. In *Carbohydrates in Chemistry and Biology;* Ernst, B.; Hart, G. W.; Sinay, P., Eds.; Wiley-VCH: Weinheim, New York, 2000; Vol. 1, p 449. doi:10.1002/9783527618255.ch18
40. Ziegler, T. In *Handbook of Chemical Glycosylation;* Demchenko, A. V., Ed.; Wiley-VCH Verlag GmbH & Co: Weinheim, Germany, 2008; pp 469 ff.
41. Ishiwata, A.; Lee, Y. J.; Ito, Y. *Org. Biomol. Chem.* **2010,** *8,* 3596. doi:10.1039/c004281a
42. Höök, J. E.; Lindberg, B. *Acta Chem. Scand.* **1968,** *22,* 2157. doi:10.3891/acta.chem.scand.22-2157
43. Malysheva, N. N.; Torgov, V. I.; Klimov, E. M.; Kochetkov, N. K. *Izv. Akad. Nauk SSSR, Ser. Khim.* **1974,** 2153.
44. Inaba, S.; Yamada, M.; Yoshino, T.; Ishido, Y. *J. Am. Chem. Soc.* **1973,** *95,* 2062. doi:10.1021/ja00787a084
45. Barresi, F.; Hindsgaul, O. *J. Am. Chem. Soc.* **1991,** *113,* 9376. doi:10.1021/ja00024a057
46. Kusumoto, S.; Imoto, M.; Ogiku, T.; Shiba, T. *Bull. Chem. Soc. Jpn.* **1986,** *59,* 1419. doi:10.1246/bcsj.59.1419
47. Jia, X. G. Ph.D. Thesis, University of Missouri, St. Louis, MO, 2016.
48. Fairbanks, A. J. *Synlett* **2003,** 1945. doi:10.1055/s-2003-42056
49. Cumpstey, I. *Carbohydr. Res.* **2008,** *343,* 1553. doi:10.1016/j.carres.2008.04.031
50. Bols, M.; Skrydstrup, T. *Chem. Rev.* **1995,** *95,* 1253. doi:10.1021/cr00037a006
51. Wakao, M.; Fukase, K.; Kusumoto, S. *J. Org. Chem.* **2002,** *67,* 8182. doi:10.1021/jo025887r
52. Ziegler, T.; Lau, R. *Tetrahedron Lett.* **1995,** *36,* 1417. doi:10.1016/0040-4039(95)00010-A
53. Lau, R.; Schüle, G.; Schwanaberg, U.; Ziegler, T. *Liebigs Ann.* **1995,** 1745. doi:10.1002/jlac.1995199510245
54. Valverde, S.; Gómez, A. M.; Hernández, A.; Herradón, B.; López, J. C. *J. Chem. Soc., Chem. Commun.* **1995,** 2005. doi:10.1039/C39950002005
55. Yamada, H.; Imamura, K.; Takahashi, T. *Tetrahedron Lett.* **1997,** *38,* 391. doi:10.1016/S0040-4039(96)02196-X
56. Pornsuriyasak, P.; Jia, X. G.; Kaeothip, S.; Demchenko, A. V. *Org. Lett.* **2016,** *18,* 2316. doi:10.1021/acs.orglett.6b01102
57. Lemanski, G.; Ziegler, T. *Tetrahedron* **2000,** *56,* 563. doi:10.1016/S0040-4020(99)01053-4
58. Lemanski, G.; Ziegler, T. *Helv. Chim. Acta* **2000,** *83,* 2655. doi:10.1002/1522-2675(20001004)83:10<2655::AID-HLCA2655>3.0.CO;2-U
59. Wakao, M.; Fukase, K.; Kusumoto, S. *Synlett* **1999,** 1911. doi:10.1055/s-1999-2997
60. Ziegler, T.; Ritter, A.; Hürttlen, J. *Tetrahedron Lett.* **1997,** *38,* 3715. doi:10.1016/S0040-4039(97)00729-6
61. Ziegler, T.; Lemanski, G. *Eur. J. Org. Chem.* **1998,** 163. doi:10.1002/(SICI)1099-0690(199801)1998:1<163::AID-EJOC163>3.0.CO;2-I
62. Ziegler, T.; Lemanski, G.; Hürttlen, J. *Tetrahedron Lett.* **2001,** *42,* 569. doi:10.1016/S0040-4039(00)02032-3
63. Claude, P.; Lehmann, C.; Ziegler, T. *Beilstein J. Org. Chem.* **2011,** *7,* 1609. doi:10.3762/bjoc.7.189
64. Valverde, S.; Gómez, A. M.; López, J. C.; Herradón, B. *Tetrahedron Lett.* **1996,** *37,* 1105. doi:10.1016/0040-4039(95)02316-X
65. Valverde, S.; García, M.; Gómez, A. M.; López, J. C. *Synlett* **2000,** 22. doi:10.1055/s-2000-6456
66. Valverde, S.; García, M.; Gómez, A. M.; López, J. C. *Chem. Commun.* **2000,** 813. doi:10.1039/b000771o
67. Cid, M. B.; Valverde, S.; López, J. C.; Gómez, A. M. *Synlett* **2005,** 1095. doi:10.1055/s-2005-865228
68. Huchel, U.; Schmidt, R. R. *Tetrahedron Lett.* **1998,** *39,* 7693. doi:10.1016/S0040-4039(98)01664-5
69. Müller, M.; Huchel, U.; Geyer, A.; Schmidt, R. R. *J. Org. Chem.* **1999,** *64,* 6190. doi:10.1021/jo990132e
70. Müller, M.; Schmidt, R. R. *Eur. J. Org. Chem.* **2001,** 2055. doi:10.1002/1099-0690(200106)2001:11<2055::AID-EJOC2055>3.0.CO;2-N
71. Abdel-Rahman, A. A.-H.; El Ashry, E. S. H.; Schmidt, R. R. *Carbohydr. Res.* **2002,** *337,* 195. doi:10.1016/S0008-6215(01)00306-8
72. Paul, S.; Müller, M.; Schmidt, R. R. *Eur. J. Org. Chem.* **2003,** 128. doi:10.1002/1099-0690(200301)2003:1<128::AID-EJOC128>3.0.CO;2-S
73. Schüle, G.; Ziegler, T. *Tetrahedron* **1996,** *52,* 2925. doi:10.1016/0040-4020(95)01093-9
74. Lemanski, G.; Ziegler, T. *Eur. J. Org. Chem.* **2006,** 2618. doi:10.1002/ejoc.200600078
75. Ogawa, T.; Takahashi, Y. *Carbohydr. Res.* **1985,** *138,* c5. doi:10.1016/0008-6215(85)85239-3
76. Tennant-Eyles, R. J.; Fairbanks, A. J.; Davis, B. G. *Chem. Commun.* **1999,** 1037. doi:10.1039/a901916b

77. Tennant-Eyles, R. J.; Davis, B. G.; Fairbanks, A. J. *Tetrahedron: Asymmetry* **2000,** *11,* 231. doi:10.1016/S0957-4166(99)00494-2
78. Tennant-Eyles, R. J.; Davis, B. G.; Fairbanks, A. J. *Tetrahedron: Asymmetry* **2003,** *14,* 1201. doi:10.1016/S0957-4166(03)00118-6
79. Greenwell, D. R.; Ibnouzaki, A. F.; Warriner, S. L. *Angew. Chem., Int. Ed.* **2002,** *41,* 1215. doi:10.1002/1521-3773(20020402)41:7<1215::AID-ANIE1215>3.0.CO;2-V
80. Tiwari, V. K.; Kumar, A.; Schmidt, R. R. *Eur. J. Org. Chem.* **2012,** 2945. doi:10.1002/ejoc.201101815
81. Kumar, A.; Geng, Y.; Schmidt, R. R. *Eur. J. Org. Chem.* **2012,** 6846. doi:10.1002/ejoc.201201076
82. Jia, X. G.; Pornsuriyasak, P.; Demchenko, A. V. *J. Org. Chem.* **2016,** *81,* 12232. doi:10.1021/acs.joc.6b02151
83. Barresi, F.; Hindsgaul, O. *Can. J. Chem.* **1994,** *72,* 1447. doi:10.1139/v94-181
84. Stork, G.; Kim, G. *J. Am. Chem. Soc.* **1992,** *114,* 1087. doi:10.1021/ja00029a047
85. Bols, M. *J. Chem. Soc., Chem. Commun.* **1992,** 913. doi:10.1039/C39920000913
86. Bols, M. *Acta Chem. Scand.* **1993,** *47,* 829. doi:10.3891/acta.chem.scand.47-0829
87. Bols, M. *J. Chem. Soc., Chem. Commun.* **1993,** 791. doi:10.1039/C39930000791
88. Bols, M. *Tetrahedron* **1993,** *49,* 10049. doi:10.1016/S0040-4020(01)80200-3
89. Bols, M.; Hansen, H. C. *Chem. Lett.* **1994,** *23,* 1049. doi:10.1246/cl.1994.1049
90. Buchan, Z. A.; Bader, S. J.; Montgomery, J. *Angew. Chem., Int. Ed.* **2009,** *48,* 4840. doi:10.1002/anie.200901666
91. Walk, J. T.; Buchan, Z. A.; Montgomery, J. *Chem. Sci.* **2015,** *6,* 3448. doi:10.1039/C5SC00810G
92. Seward, C. M. P.; Cumpstey, I.; Aloui, M.; Ennis, S. C.; Redgrave, A. J.; Fairbanks, A. J. *Chem. Commun.* **2000,** 1409. doi:10.1039/b004522p
93. Cumpstey, I.; Fairbanks, A. J.; Redgrave, A. J. *Org. Lett.* **2001,** *3,* 2371. doi:10.1021/ol016175a
94. Ito, Y.; Ogawa, T. *Angew. Chem., Int. Ed. Engl.* **1994,** *33,* 1765. doi:10.1002/anie.199417651
95. Dan, A.; Ito, Y.; Ogawa, T. *J. Org. Chem.* **1995,** *60,* 4680. doi:10.1021/jo00120a002
96. Dan, A.; Lergenmüller, M.; Amano, M.; Nakahara, Y.; Ogawa, T.; Ito, Y. *Chem. – Eur. J.* **1998,** *4,* 2182. doi:10.1002/(SICI)1521-3765(19981102)4:11<2182::AID-CHEM2182>3.0.CO;2-U
97. Ito, Y.; Ogawa, T. *J. Am. Chem. Soc.* **1997,** *119,* 5562. doi:10.1021/ja964093p
98. Pratt, M. R.; Leigh, C. D.; Bertozzi, C. R. *Org. Lett.* **2003,** *5,* 3185. doi:10.1021/ol034836t
99. Leigh, C. D.; Bertozzi, C. R. *J. Org. Chem.* **2008,** *73,* 1008. doi:10.1021/jo702032c
100. Ishiwata, A.; Munemura, Y.; Ito, Y. *Eur. J. Org. Chem.* **2008,** 4250. doi:10.1002/ejoc.200800249
101. Yong, J. L.; Ishiwata, A.; Ito, Y. *J. Am. Chem. Soc.* **2008,** *130,* 6330. doi:10.1021/ja801574q
102. Ishiwata, A.; Ito, Y. *Trends Glycosci. Glycotechnol.* **2009,** *21,* 266. doi:10.4052/tigg.21.266
103. Ishiwata, A.; Kaeothip, S.; Takeda, Y.; Ito, Y. *Angew. Chem., Int. Ed.* **2014,** *53,* 9812. doi:10.1002/anie.201404904
104. Ishiwata, A.; Ito, Y. *J. Am. Chem. Soc.* **2011,** *133,* 2275. doi:10.1021/ja109932t
105. Ishiwata, A.; Sakurai, A.; Nishimiya, Y.; Tsuda, S.; Ito, Y. *J. Am. Chem. Soc.* **2011,** *133,* 19524. doi:10.1021/ja208528c
106. Ishiwata, A.; Ito, Y. *Yuki Gosei Kagaku Kyokaishi* **2012,** *70,* 382. doi:10.5059/yukigoseikyokaishi.70.382
107. Kaeothip, S.; Ishiwata, A.; Ito, Y. *Org. Biomol. Chem.* **2013,** *11,* 5892. doi:10.1039/c3ob41212a
108. Ishiwata, A.; Ito, Y. *Glycochemical Synthesis: Strategies and Applications;* John Wiley & Sons: Hoboken, 2015; pp 361 ff.
109. Nakagawa, A.; Tanaka, M.; Hanamura, S.; Takahashi, D.; Toshima, K. *Angew. Chem., Int. Ed.* **2015,** *54,* 10935. doi:10.1002/anie.201504182
110. Tanaka, M.; Nashida, J.; Takahashi, D.; Toshima, K. *Org. Lett.* **2016,** *18,* 2288. doi:10.1021/acs.orglett.6b00926
111. Nishi, N.; Nashida, J.; Kaji, E.; Takahashi, D.; Toshima, K. *Chem. Commun.* **2017,** *53,* 3018. doi:10.1039/C7CC00269F
112. Smoot, J. T.; Pornsuriyasak, P.; Demchenko, A. V. *Angew. Chem., Int. Ed.* **2005,** *44,* 7123. doi:10.1002/anie.200502694
113. Smoot, J. T.; Demchenko, A. V. *J. Org. Chem.* **2008,** *73,* 8838. doi:10.1021/jo801551r
114. Yasomanee, J. P.; Demchenko, A. V. *J. Am. Chem. Soc.* **2012,** *134,* 20097. doi:10.1021/ja307355n
115. Yasomanee, J. P.; Demchenko, A. V. *Chem. – Eur. J.* **2015,** *21,* 6572. doi:10.1002/chem.201406589
116. Yasomanee, J. P.; Demchenko, A. V. *Angew. Chem., Int. Ed.* **2014,** *53,* 10453. doi:10.1002/anie.201405084
117. Pistorio, S. G.; Yasomanee, J. P.; Demchenko, A. V. *Org. Lett.* **2014,** *16,* 716. doi:10.1021/ol403396j
118. Yasomanee, J. P.; Parameswar, A. R.; Pornsuriyasak, P.; Rath, N. P.; Demchenko, A. V. *Org. Biomol. Chem.* **2016,** *14,* 3159. doi:10.1039/C6OB00107F
119. Kayastha, A. K.; Jia, X. G.; Yasomanee, J. P.; Demchenko, A. V. *Org. Lett.* **2015,** *17,* 4448. doi:10.1021/acs.orglett.5b02110
120. Xiang, S.; He, J.; Tan, Y. J.; Liu, X.-W. *J. Org. Chem.* **2014,** *79,* 11473. doi:10.1021/jo502078c
121. Xiang, S.; Hoang, K. L. M.; He, J.; Tan, Y. J.; Liu, X.-W. *Angew. Chem., Int. Ed.* **2015,** *54,* 604. doi:10.1002/anie.201408739
122. Behrendt, M. E.; Schmidt, R. R. *Tetrahedron Lett.* **1993,** *34,* 6733. doi:10.1016/S0040-4039(00)61687-8
123. Scheffler, G.; Schmidt, R. R. *Tetrahedron Lett.* **1997,** *38,* 2943. doi:10.1016/S0040-4039(97)00561-3
124. Scheffler, G.; Behrendt, M. E.; Schmidt, R. R. *Eur. J. Org. Chem.* **2000,** 3527. doi:10.1002/1099-0690(200011)2000:21<3527::AID-EJOC3527>3.0.CO;2-P
125. Azumaya, I.; Kotani, M.; Ikegami, S. *Synlett* **2004,** 959. doi:10.1055/s-2004-822885
126. Azumaya, I.; Niwa, T.; Kotani, M.; Iimori, T.; Ikegami, S. *Tetrahedron Lett.* **1999,** *40,* 4683. doi:10.1016/S0040-4039(99)00843-6
127. Scheffler, G.; Schmidt, R. R. *J. Org. Chem.* **1999,** *64,* 1319. doi:10.1021/jo971778e
128. Laursen, J. B.; Petersen, L.; Jensen, K. J. *Org. Lett.* **2001,** *3,* 687. doi:10.1021/ol006988j
129. Liu, X.; Zhang, B.; Gu, X.; Chen, G.; Chen, L.; Wang, X.; Xiong, B.; You, Q.-D.; Chen, Y.-L.; Shen, J. *Carbohydr. Res.* **2014,** *398,* 45. doi:10.1016/j.carres.2014.05.010

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Preactivation-based chemoselective glycosylations: A powerful strategy for oligosaccharide assembly

Weizhun Yang[‡1], Bo Yang[‡1], Sherif Ramadan[1,2] and Xuefei Huang[*1,3]

Review

Open Access

Address:
[1]Department of Chemistry, Michigan State University, 578 South Shaw Lane, East Lansing, MI 48824, USA, [2]Chemistry Department, Faculty of Science, Benha University, Benha, Qaliobiya 13518, Egypt and [3]Department of Biomedical Engineering, Michigan State University, East Lansing, MI 48824, USA

Email:
Xuefei Huang[*] - xuefei@chemistry.msu.edu

* Corresponding author ‡ Equal contributors

Keywords:
chemoselectivity; glycosides; preactivation; synthesis

Beilstein J. Org. Chem. **2017,** *13,* 2094–2114.
doi:10.3762/bjoc.13.207

Received: 01 May 2017
Accepted: 14 September 2017
Published: 09 October 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Most glycosylation reactions are performed by mixing the glycosyl donor and acceptor together followed by the addition of a promoter. While many oligosaccharides have been synthesized successfully using this premixed strategy, extensive protective group manipulation and aglycon adjustment often need to be performed on oligosaccharide intermediates, which lower the overall synthetic efficiency. Preactivation-based glycosylation refers to strategies where the glycosyl donor is activated by a promoter in the absence of an acceptor. The subsequent acceptor addition then leads to the formation of the glycoside product. As donor activation and glycosylation are carried out in two distinct steps, unique chemoselectivities can be obtained. Successful glycosylation can be performed independent of anomeric reactivities of the building blocks. In addition, one-pot protocols have been developed that have enabled multiple-step glycosylations in the same reaction flask without the need for intermediate purification. Complex glycans containing both 1,2-*cis* and 1,2-*trans* linkages, branched oligosaccharides, uronic acids, sialic acids, modifications such as sulfate esters and deoxy glycosides have been successfully synthesized. The preactivation-based chemoselective glycosylation is a powerful strategy for oligosaccharide assembly complementing the more traditional premixed method.

Review

Introduction

Carbohydrates are widely present in nature and many of them are involved in important physiological and pathological events, such as anticoagulation, inflammation and pathogen infection [1,2]. In order to explore their biological functions, oligosaccharides with high purity are needed [3]. However, this is hampered by the limited availability of complex glycans from nature. Thus, chemical synthesis is a powerful approach to provide much needed samples to enable biological studies [4].

Traditional carbohydrate synthesis is commonly carried out from the reducing end to the non-reducing end with a glycosyl donor premixed with an acceptor. Upon the addition of a promoter to the reaction mixture, the donor is activated to glycosylate the acceptor yielding a disaccharide, which is subsequently deprotected to expose a free hydroxy group (Scheme 1a). The newly generated acceptor can be coupled with another donor and this process is repeated until the desired oligosaccharide structure is assembled. Although many oligosaccharides have been successfully produced through this approach, the traditional oligosaccharide synthesis requires multiple synthetic manipulations on oligosaccharide intermediates, which lowers the overall synthetic efficiency.

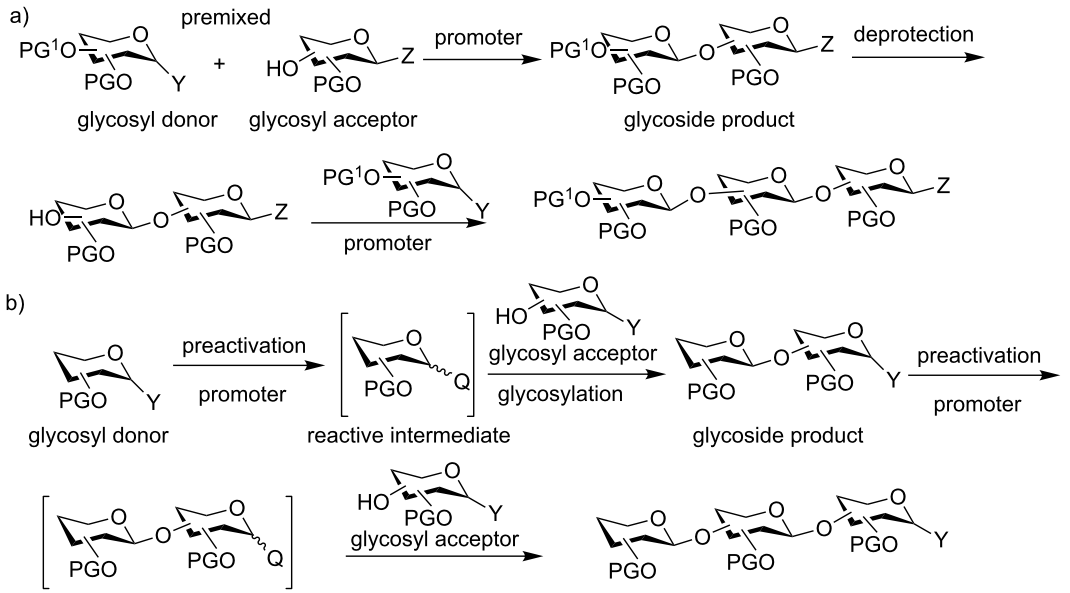
To expedite the oligosaccharide assembly process, many innovative strategies have been developed [5], such as active-latent activation [6-8], orthogonal glycosylation [9,10], reactivity-based armed-disarmed glycosylation [11-14], fluorine-supported glycosylation [15,16] and automated solid-phase synthesis [17]. All of these methods use the donor/acceptor premixed approach and preferential activation of the donor is achieved by the higher anomeric reactivity of the donor towards the promoter compared to the acceptor. In comparison, the preactivation-based iterative glycosylation is unique, where a glycosyl donor is preactivated in the absence of an acceptor to produce a reactive intermediate (Scheme 1b) [18-21]. Upon complete donor activation, the acceptor is added to the reaction mixture, which nucleophilically attacks the intermediate forming the desired glycosidic product [22-24].

With the preactivation protocol, the donor activation and acceptor glycosylation occur in two distinctive steps. As a result, a unique chemoselectivity can be achieved with preactivation. Glycosyl donors and acceptors with the same aglycon leaving group can be used enabling an iterative glycosylation, simplifying the overall synthetic design.

For a preactivation based glycosylation reaction to be successful the intermediate formed upon preactivation must be stable prior to the addition of the acceptor and yet reactive enough to quickly react with the acceptor during the glycosylation step without the need for another exogenous promoter or separation of the intermediate. Various types of glycosyl building blocks and promoter systems have been developed for preactivation. This review will be divided according to the type of glycosyl donors that can undergo a preactivation-based chemoselective glycosylation with an emphasis on thioglycosides due to their wide applicability.

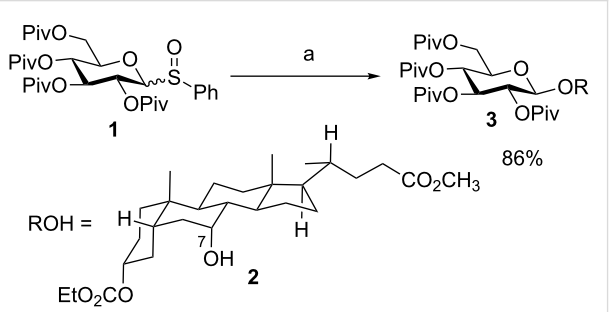
Preactivation of glycosyl sulfoxides: early success of preactivation

One of the earliest preactivation-based glycosylation reactions utilized glycosyl sulfoxide donors for glycosylation of unreactive substrates such as steroid derivative **2** by the Kahne group [25]. The axial C-7 hydroxy group in **2** is sterically hindered due to unfavorable 1,3-diaxial interactions. The traditional premixed glycosylation gave only low yields (<30%) of the products [26]. In contrast, when glycosyl sulfoxide donor **1** was preactivated with Tf_2O at −78 °C, followed by the addition of



Scheme 1: a) Traditional glycosylation typically employs the premixed approach with both the donor and the acceptor mixed together, before the promoter is added; b) the preactivation based glycosylation strategy activates the glycosyl donor in the absence of the acceptor, which temporally separates the donor activation step from acceptor glycosylation.

sterol **2** and 2,6-di-*tert*-butyl-4-methylpyridine (DTBMP) as an acid scavenger, the desired compound **3** was obtained in an excellent 86% yield (Scheme 2). While this method has not been applied to glycosyl sulfoxide as the acceptor for iterative glycosylation, this early example demonstrated the power of preactivation. Subsequently, a wide range of glycosyl donors have been explored.



Scheme 2: Glycosylation of an unreactive substrate. Reagents and conditions: (a) Tf_2O, −78 °C, CH_2Cl_2 (DCM), then **2**, DTBMP.

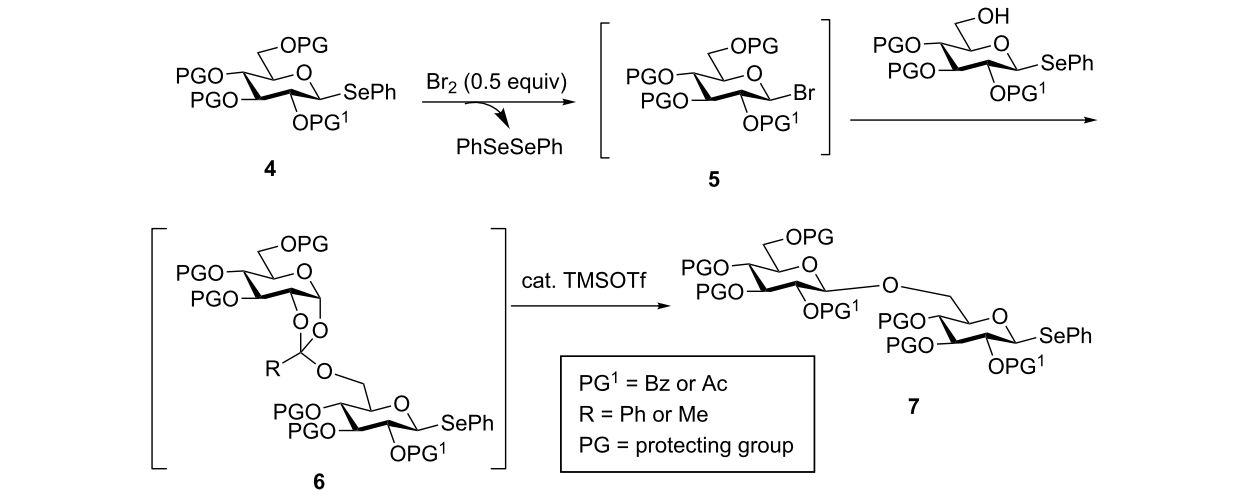
β-Glycosyl bromide-mediated iterative gycosylation of selenoglycosides

Yoshida and co-workers developed a preactivation-based glycosylation approach using selenoglycosides via the intermediacy of β-glycosyl bromides (Scheme 3) [27,28]. Upon the addition of 0.5 equiv of bromine, half of the selenoglycoside donor **4** would be activated to presumably form glycosyl bromide **5** accompanied by the generation of PhSeBr. PhSeBr could react with the remaining donor **4** for quantitative activation of **4**. The addition of the acceptor to the reaction mixture upon donor preactivation afforded orthoester **6**. The orthoester **6** was rearranged in situ with trimethylsilyl trifluoromethanesulfonate (TMSOTf) to disaccharide **7**, which could be subjected to bromine-promoted glycosylation for further chain elongation. As an example, preactivation of a monosaccharide **8** with bromine was followed by the addition of a bifunctional disaccharide building block **10** and subsequent TMSOTf-promoted orthoester rearrangement, producing trisaccharide selenoglycoside **11** in 90% yield (Scheme 4). Following the same reaction protocol trisaccharide **11** and glycosylated acceptor **9** lead to tetrasaccharide **12,** which was further extended to heptasaccharide **13**. This method has also been applied to generate a library of phytoalexin elicitor-active oligoglucosides [28].

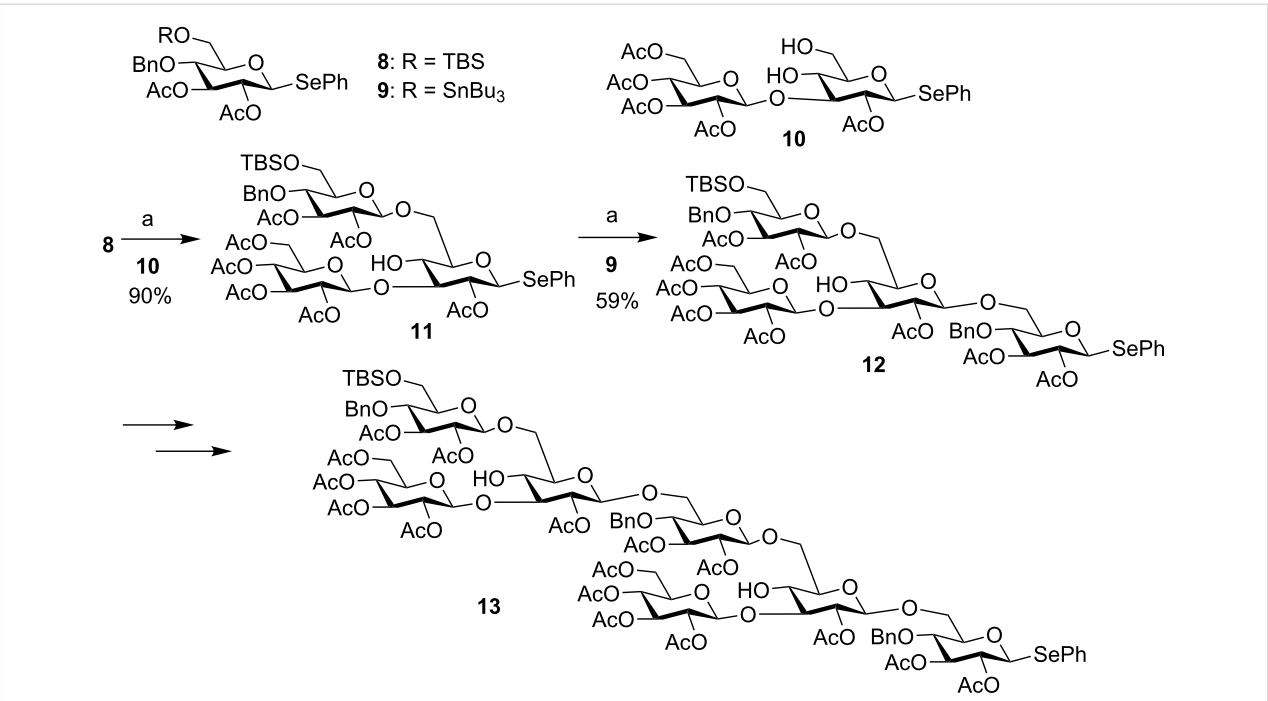
A limitation of this glycosyl bromide-mediated selenoglycoside iterative glycosylation is that it is restricted to the formation of 1,2-*trans*-glycosyl linkages. Furthermore, an additional isomerization step is needed to transform the orthoester to the desired glycoside.

Preactivation-based iterative glycosylation of 2-pyridyl glycosides

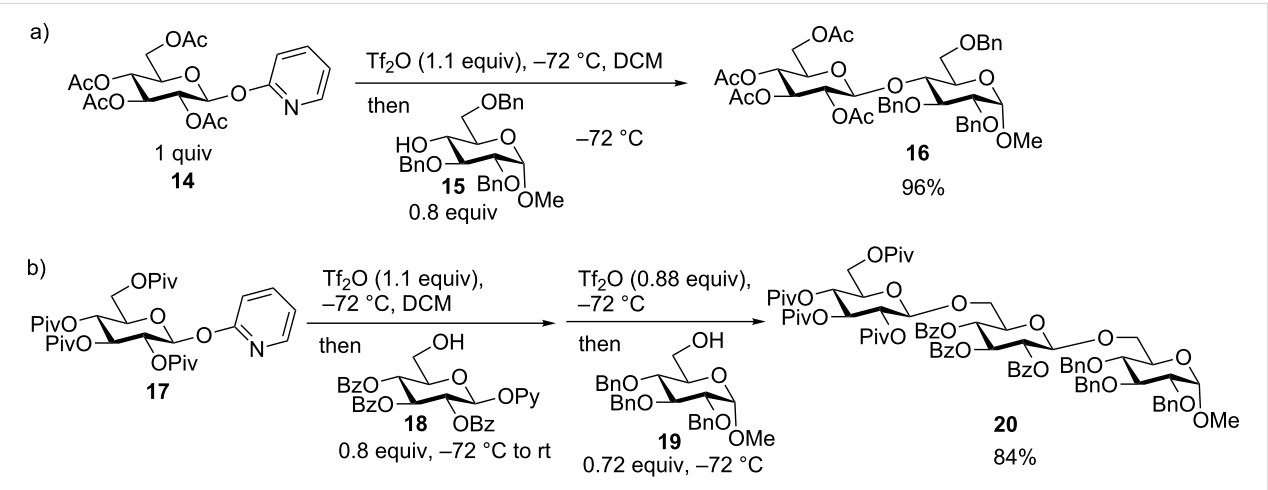
O-Unprotected 2-pyridyl glycosyl donors have been utilized in oligosaccharide synthesis [29]. The Ye group reported a preactivation protocol using protected 2-pyridyl donors [30]. The preactivation of 2-pyridyl glycoside **14** was performed using Tf_2O as the promoter, which was followed by the addition of acceptor **15** generating disaccharide **16** in 96% yield (Scheme 5a). The acceptor could also bear a 2-pyridyl aglycon such as acceptor **18**. The preactivation-based glycosylation of donor **17** with acceptor **18** led to a disaccharide intermediate, which was then subjected to another round of Tf_2O-mediated glycosylation leading to trisaccharide **20** in one pot (Scheme 5b). As compounds **16** and **20** have relatively simple structures, the scope of this 2-pyridyl glycosylation method will



Scheme 3: Bromoglycoside-mediated glycosylation.



Scheme 4: Glycosyl bromide-mediated selenoglycosyl donor-based iterative glycosylation. Reagents and conditions: (a) Br_2, (0.5 equiv), 0 °C, CH_2Cl_2; then **10** or **9**, rt; then TMSOTf (0.1 equiv), 0 °C.

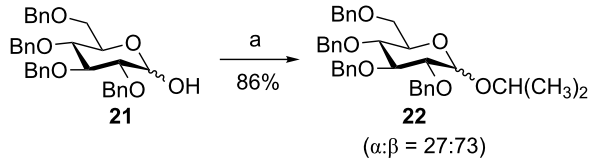


Scheme 5: Preactivation-based glycosylation using 2-pyridyl glycosyl donors.

need to be established in the total synthesis of more complex oligosaccharides.

Chemoselective dehydrative glycosylation with glycosyl hemiacetals

Most glycosylation reactions require a functionalization of the anomeric position of a glycosyl donor followed by the reaction with a promoter to induce the irreversible transfer of the donor to an acceptor [31-35]. The displacement of the anomeric hydroxy group of a glycosyl hemiacetal by an acceptor for dehydrative glycosylation is an interesting alternative as glycosyl hemiacetals are often undesired side products in glycosylation reactions due to the competitive reaction with trace amounts of water present in the reaction mixture. The Gin group established a preactivation glycosylation procedure using glycosyl hemiacetals [36]. As an example, the hemiacetal donor **21** was preactivated with Tf_2O and diphenyl sulfoxide (Ph_2SO) at −40 °C. This was followed by the addition of the acceptor isopropyl alcohol, affording glycoside **22** in 86% yield (α:β = 27:73, Scheme 6). This glycosylation strategy can be applied to



Scheme 6: Chemoselective dehydrative glycosylation. Reagents and conditions: (a) Ph_2SO, Tf_2O, 2-chloropyridine, then $(CH_3)_2CHOH$, −40 °C.

a variety of glycosyl acceptors, including oxygen, sulfur, carbon and nitrogen nucleophiles (Figure 1) [36]. Even the unreactive *N*-(trimethylsilyl)trimethylacetamide could be efficiently glycosylated to afford the corresponding glycosyl amide **26**.

Two possible reaction pathways have been proposed for this dehydrative glycosylation (Scheme 7) [37]. Upon mixing diphenyl sulfoxide and triflic anhydride, diphenyl sulfide bis(triflate) (**27**) is formed in situ (Scheme 7a). In pathway 1, hemiacetal **28** could attack the sulfonium center of diphenyl sulfide bis(triflate) (**27**) to give the glycosyl oxosulfonium intermediate **29**, which subsequently glycosylated the acceptor to yield the product **30** (Scheme 7b). Alternatively, in pathway 2, hemiacetal **28** could attack the sulfonyl center of diphenyl sulfide bis(triflate) (**27**) to give the glycosyl triflate intermediate **31**, followed by glycosylation to give **30** (Scheme 7c). To distinguish between these two possibilities, an ^{18}O-labeling study was carried out by subjecting ^{18}O-labeled hemiacetal **28**

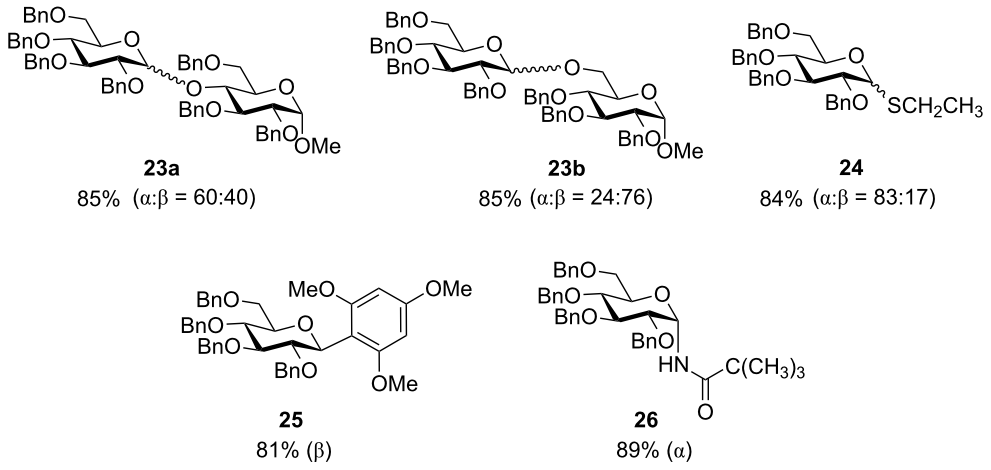
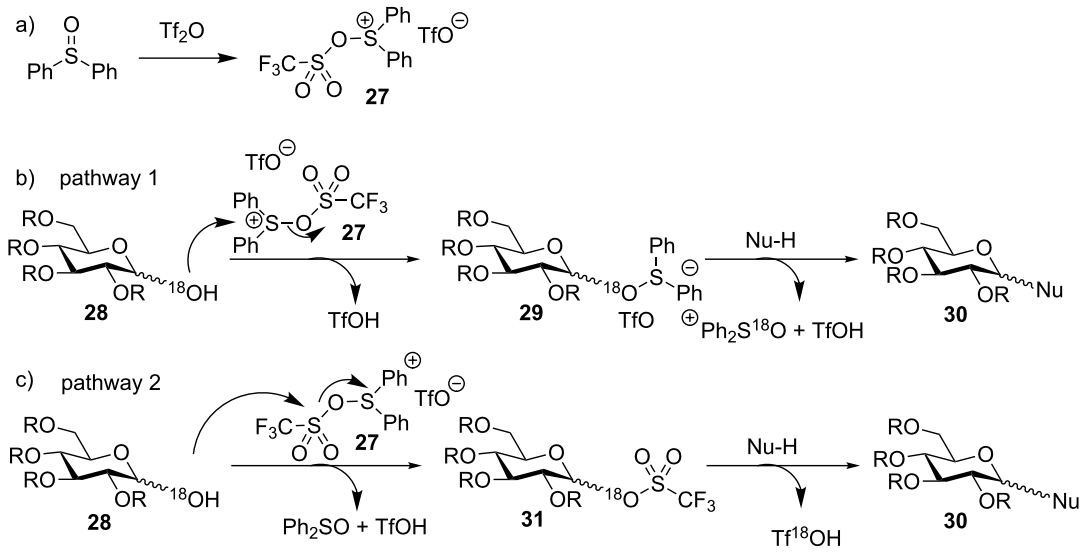


Figure 1: Representative structures of products formed by the preactivation-based dehydrative glycosylation of glycosyl hemiacetal.



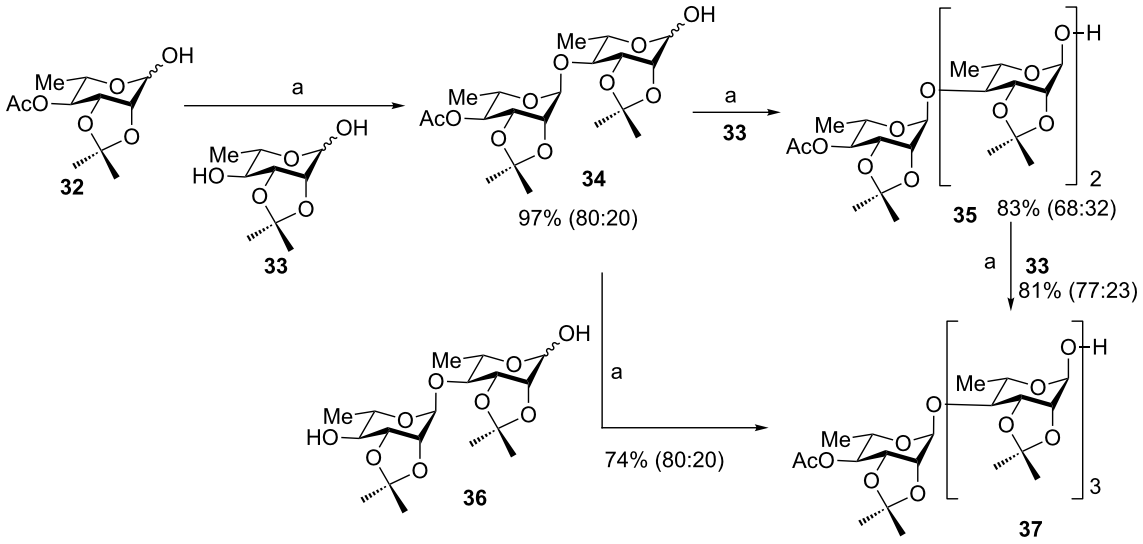
Scheme 7: Possible mechanism for the dehydrative glycosylation. (a) Formation of diphenyl sulfide bis(triflate) (**27**) as the promoter for glycosyl hemiacetal activation; (b) pathway 1 and (c) pathway 2 as potential mechanisms for glycosyl hemiacetal activation.

to the glycosylation conditions. Indeed, ^{18}O-labeled diphenyl sulfoxide was detected in the reaction mixture as the main ^{18}O-labeled compound, which suggested pathway 1 was the major reaction mechanism.

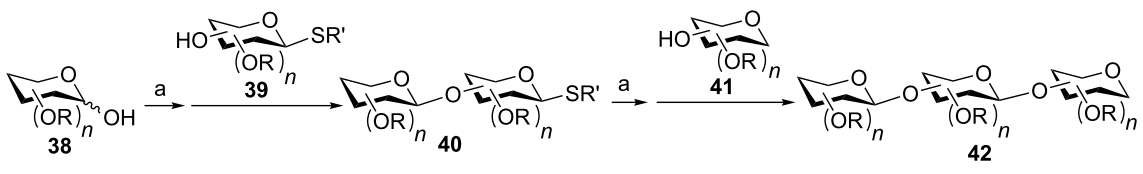
The hemiacetal donor can be utilized in iterative glycosylation (Scheme 8) [21]. Donor **32** was preactivated by Ph_2SO and Tf_2O, followed by the addition of glycosyl hemiacetal **33** with one of its hydroxy groups free available as the acceptor producing disaccharide **34**. The regioselectivity is presumably due to the higher nucleophilicity of the alkyl hydroxy group than that of the hemiacetal hydroxy group. This process can be repeated for chain elongation without the need for any protective group manipulation or aglycon adjustment. Using this method, the 1,4-α-linked tetrasaccharide **37** was prepared in good overall yield.

Inspired by Gin's work, van der Marvel and co-workers developed a sequential glycosylation strategy by combining hemiacetal and thioglycosyl building blocks as illustrated in Scheme 9 [38]. The hemiacetal donor **38** was preactivated with Ph_2SO and Tf_2O, and reacted with a bifunctional thioglycosyl acceptor **39** to form disaccharide **40**. Interestingly, thioglycoside **40** could also be activated by Ph_2SO/Tf_2O. The subsequent addition of acceptor **41** to the reaction mixture furnished trisaccharide **42**. This approach was applied to the synthesis of hyaluronic acid (HA) oligomers [39]. The sequential reaction of building blocks **43**, **44** and **46** led to HA trisaccharide **47** (Scheme 10). The modest overall yield of 26% for the two glycosylation reactions was attributed to the formation of orthoester and oxazolidine side products due to the basic reaction conditions, which were needed to neutralize the acid formed during glycosylation and to avoid the cleavage of the acid-labile benzylidene protective group.

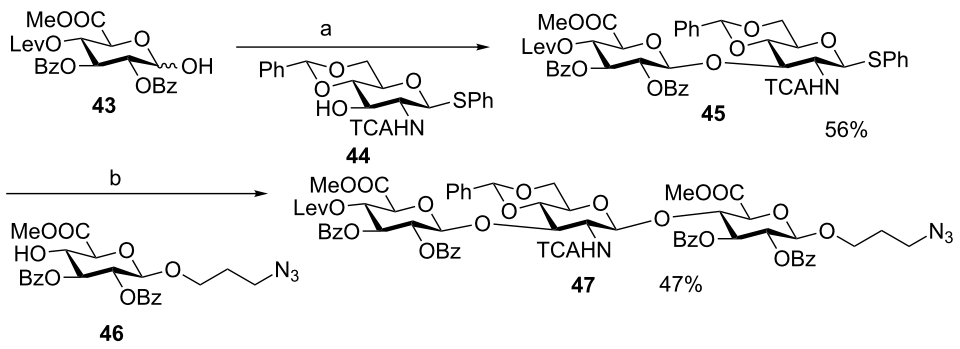
The van der Marel group further applied their strategy to the synthesis of heparin (HP) and heparan sulfate (HS), which are more complex members of the glycosaminoglycan family [40]. A pentasaccharide **48** was chosen as the synthetic target (Figure 2). A major challenge of HP and HS synthesis lies in the coupling of an azido glucoside with a uronic acid in an α-selective fashion. A variety of azido hemiacetal glucoside



Scheme 8: Chemoselective iterative dehydrative glycosylation. Reagents and conditions: (a) Ph_2SO, Tf_2O, 2,4,6-tri-*tert*-butylpyrimidine (TTBP), −78 °C to −40 °C; then acceptor.



Scheme 9: Chemoselective iterative dehydrative glycosylation. Reagents and conditions: (a) Ph_2SO, Tf_2O, −40 °C.



Scheme 10: Chemical synthesis of a hyaluronic acid (HA) trimer **47**. Reagents and conditions: (a) Ph_2SO, TTBP, CH_2Cl_2, −60 °C, then Tf_2O, **44**, −60 °C to 0 °C; (b) Ph_2SO, TTBP, CH_2Cl_2, −60 °C, then Tf_2O, **46**, −60 °C to 0 °C.

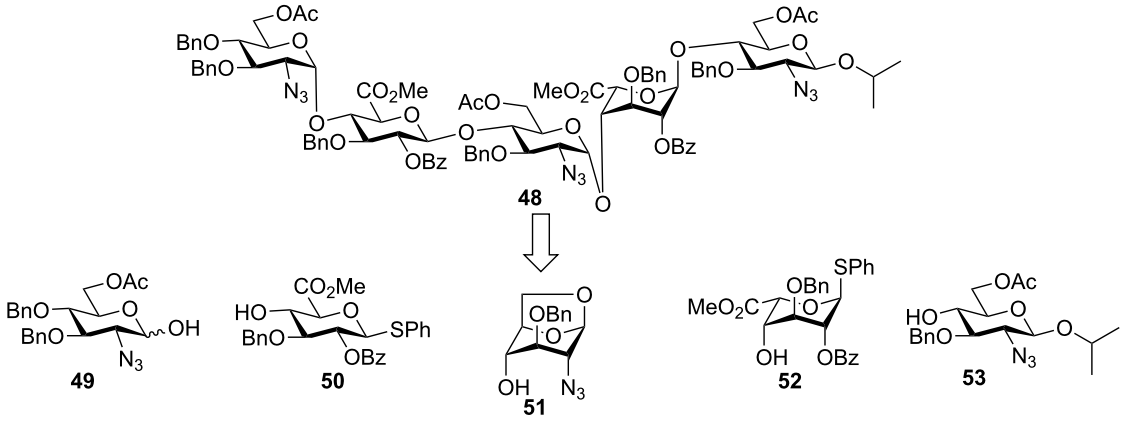
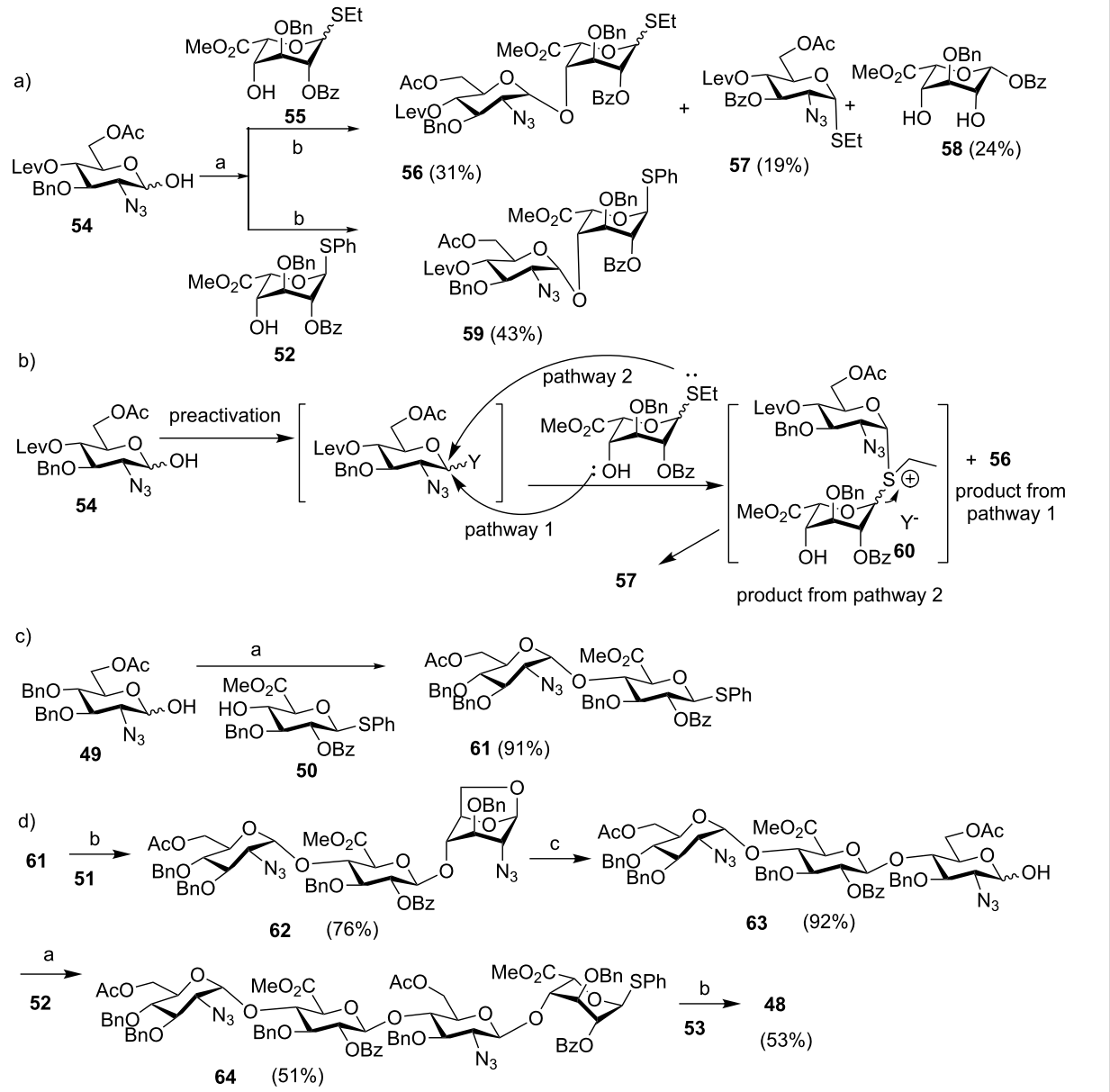


Figure 2: Retrosynthetic analysis of pentasaccharide **48**.

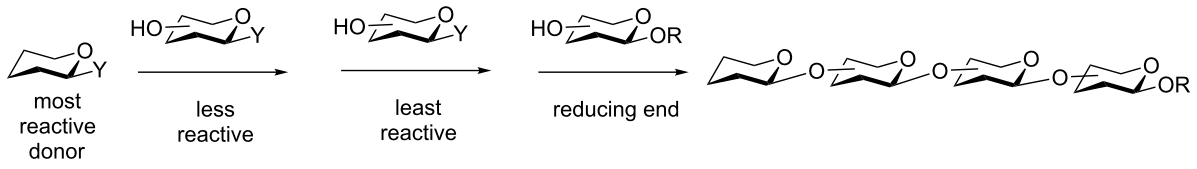
donor and uronic acid thioglycosyl acceptor pairs were screened under preactivation conditions. The anomeric leaving groups of the acceptors had significant impacts on the glycosylation outcomes (Scheme 11a). When donor **54** was utilized to glycosylate iduronic acid **55**, disaccharide **56** was obtained only in 31% yield along with aglycon transfer products, **57** (19%) and **58** (24%). The modest yield of the desired glycoside product resulted from the lower nucleophilicity of 4-OH as compared to the thioethyl moiety, which could compete with the nucleophilic attack by the 4-OH leading to aglycon transfer (Scheme 11b). In contrast, when thiophenyl glycoside **52** was used as the acceptor, no aglycon transfer product was isolated and disaccharide **59** was obtained in 43% yield. The improvement presumably resulted from the lower nucleophilicity of the thiophenyl moiety due to the steric as well as electronegative effects of the phenyl group. The hemiacetal donor **49** glycosylated the thiophenyl glucuronate acceptor **50** in an excellent 91% yield using the preactivation protocol (Scheme 11c). The successful preparation of disaccharides **61** and **59** paved the way for the synthesis of protected heparin pentasaccharide **48** (Scheme 11d).

Preactivation-based chemoselective glycosylation of thioglycosides

Thioglycosides are one of the most commonly utilized building blocks due to their high stabilities under a wide range of synthetic transformations commonly encountered in building block preparation [41]. At the same time, mild promoters are available for thioglycoside activation. The anomeric reactivities of thioglycosides towards glycosylation can be significantly influenced by the protective groups on the glycan ring as well as the size and nucleophilicity of the thioether aglycon [42-44]. Extensive studies on how to fine tune anomeric reactivities culminated in the establishment of the powerful reactivity-based chemoselective glycosylation method [11]. In this strategy, a thioglycosyl donor with high anomeric reactivity is mixed together with a bifunctional thioglycosyl acceptor with lower anomeric reactivity (Scheme 12). Upon the addition of a promoter, the donor is preferentially activated to glycosylate the acceptor. The resulting disaccharide can then be utilized directly as a donor to react with another bifunctional thioglycoside with even lower anomeric reactivity. When building blocks with suitable anomeric reactivities are selected, multiple glycosyla-



Scheme 11: Effects of anomeric leaving groups on glycosylation outcomes. Reagents and conditions: (a) Ph_2SO, Tf_2O, TTBP, CH_2Cl_2, −40 °C; then acceptor, −40 °C to rt, (b) 1-(benzenesulfinyl)piperidine, Tf_2O, CH_2Cl_2, −60 °C, then acceptor, (c) 10% trifluoroacetic acid in Ac_2O, 0 °C to rt, then 6% piperidine in THF.



Scheme 12: Reactivity-based one-pot chemoselective glycosylation.

tion reactions can be carried out in one pot without the need for synthetic manipulations or purification of the advanced oligosaccharide intermediates. This strategy, which has been covered in other reviews [23,42], has been applied to successful synthesis of a range of complex oligosaccharides including human milk oligosaccharides [45], an embryonic stem cell surface

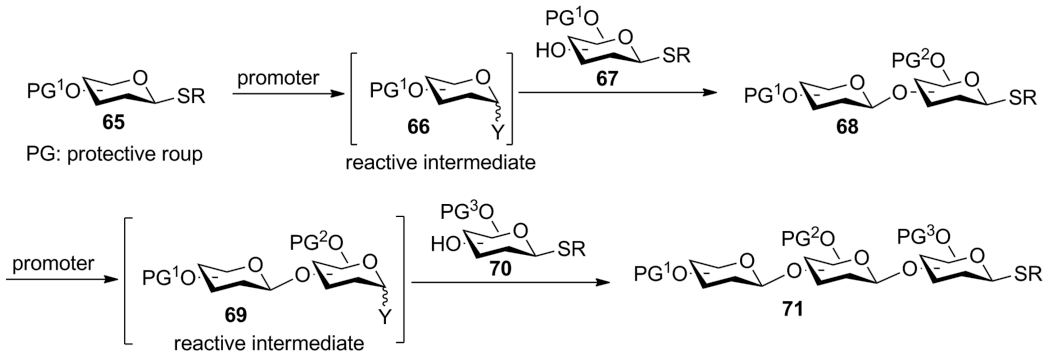
carbohydrate marker Lc4 [46], Globo-H hexasaccharide [47], and heparin-like oligosaccharides [48].

A significant drawback of the reactivity-based chemoselective glycosylation method is the requirement that the glycosyl donor must bear higher anomeric reactivities than the acceptor for preferential donor activation. As a result, extensive protecting group manipulations have to be carried out to prepare building blocks with the required anomeric reactivities. Furthermore, the relative anomeric reactivity values of a building block can vary depending on the structures of acceptors and reaction condition [44], presenting challenges in accurately predicting the reaction outcome.

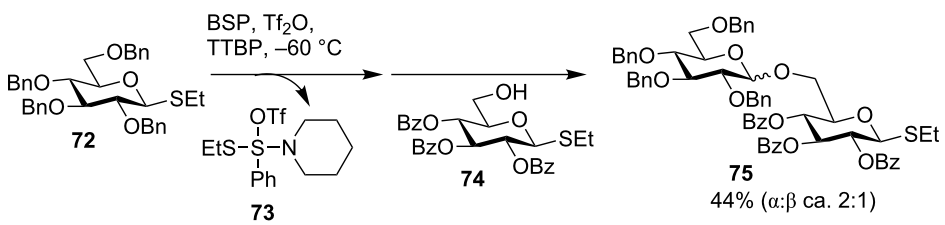
The aforementioned drawbacks of the reactivity-based chemoselective glycosylation can be overcome through preactivation. Under the preactivation protocol, a thioglycosyl donor is activated in the absence of an acceptor to form a reactive intermediate (Scheme 13). Upon complete donor activation, a thioglycosyl acceptor is added, which reacts with the intermediate to form the desired glycoside without the need for additional promoter. The resulting disaccharide bears a thioether aglycon, which can undergo another round of preactivation and glycosylation for rapid chain extension. As donor activation and acceptor glycosylation are carried out in two distinct steps, the preactivation strategy obviates the requirement that the glycosyl donor must have a higher anomeric reactivity than the acceptor for preferential activation, bestowing greater flexibilities in building block design.

A key consideration in designing successful preactivation-based thioglycoside glycosylation is the promoter. It needs to be able to activate a wide range of donors stoichiometrically rather than catalytically to avoid an undesired activation of the acceptor or the product. Many thiophilic activators have been tested, which include *p*-TolSCl/AgOTf [18], *N*-iodosuccinimide (NIS)/TMSOTf [18], dimethyl(methylthio)sulfonium triflate (DMTST) [18], 1-(benzenesulfinyl)piperidine (BSP)/Tf_2O [18,19,49], *S*-(4-methoxyphenyl)benzene-thiosulfinate (MBPT)/Tf_2O [50], Ph_2SO/Tf_2O [36,51], *O*,*O*-dimethylthiophosphonosulfenyl bromide (DMTPSB)/AgOTf [52], and 4-(benzenesulfinyl)morpholine (BSM)/Tf_2O [53].

The combination of BSP/Tf_2O [19,49] has been utilized as the promoter for iterative oligosaccharide synthesis including oligoglucosamine library [20], oligomannan [54] and Lewis^a trisaccharide [55]. During their synthesis, van der Marel and co-workers [19] found that with BSP/Tf_2O promoter, the glycosylation of donor **72** and acceptor **74** gave a moderate yield of 44% (α:β = 2:1) of the desired product **75** (Scheme 14). This was attributed to the formation of (*N*-piperidino)phenyl-(*S*-thioethyl)sulfide triflate (**73**) from the reaction of BSP/Tf_2O



Scheme 13: Preactivation-based iterative glycosylation of thioglycosides.



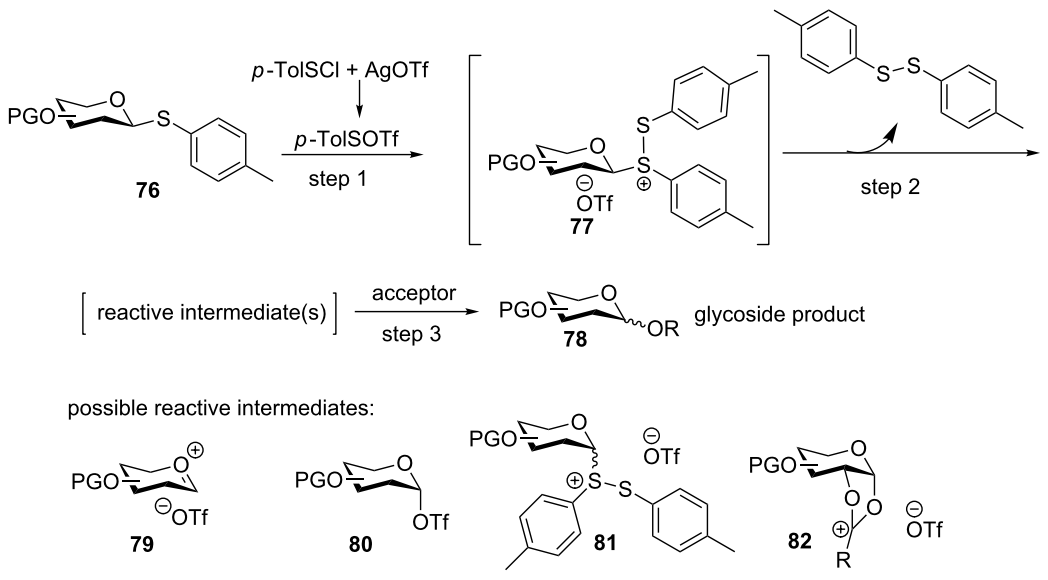
Scheme 14: BSP/Tf_2O promoted synthesis of **75**.

with the thioglycosyl donor. The sulfide triflate **73** could activate the thioglycoside product, which provides a possible explanation for the modest yield. To avoid the side reaction caused by **73**, triethyl phosphite was added as a scavenger to quench **73**, which enhanced the glycosylation yield to 78%.

The need for triethyl phosphite to prevent the undesired acceptor/product activation precludes the possibility of carrying out multiple glycosylation reactions in one pot using BSP/Tf_2O. Other promoter systems such as NIS/TMSOTf, Ph_2SO/Tf_2O and BSM/Tf_2O have similar complications due to the formation of thiophilic or nucleophilic side products following donor activation. Through extensive experimentation, Huang, Ye and co-workers successfully developed an iterative one-pot glycosylation strategy using the *p*-TolSCl/AgOTf promoter system and *p*-tolyl thioglycosides as building blocks [18]. A possible mechanism for this glycosylation has been proposed (Scheme 15). Addition of *p*-TolSCl to the mixture of donor **76** and AgOTf forms *p*-TolSOTf, a powerful electrophile that can electrophilically add to the anomeric sulfur atom of **76** forming disulfonium ion **77** (step 1 in Scheme 15). After ejection of the ditolyl disulfide, **77** can evolve into several reactive species, such as oxocarbenium ion **79**, α-triflate **80**, disulfonium ion **81**, and dioxalenium ion **82**. The nucleophilic attack of the intermediate by a thioglycosyl acceptor would generate the desired glycoside **78**.

Pioneered by Crich and co-workers, low temperature NMR studies have been found to be a powerful approach to analyze intermediates formed during glycosylation reactions [56]. To determine the dominant intermediate in preactivation of thioglycosides, low-temperature an NMR experiments were carried out following donor activation [57]. It was determined that with perbenzoylated donor **83**, the α-glycosyl triflate **84** was formed as the major intermediate [56,58,59]. When the more electron-rich donor **85** was preactivated, the dioxalenium ion **86** via the participation of the 2-benzoyl (Bz) group was found as the dominating species from NMR analysis (Figure 3) [57]. Interestingly, when **87** was preactivated, two major intermediates were produced (α-triflate **88** and dioxalenium ion **89**). The different outcome upon preactivation can be explained in terms of different electron-withdrawing properties of the protective groups present in these three donors. For **83**, the Bz group greatly disfavors the formation of a positively charged dioxalenium ion while the electron-donating benzyl (Bn) group can stabilize the dioxalenium ion. Donor **87** presents an intermediate case. The absence of the disulfonium ion **81** following the donor activation confirms that the disulfide does not significantly impact the structure of the intermediates. The more electron-rich glycosyl donors were found to give higher yields in glycosylation, especially with unreactive and electron-poor secondary acceptors. A representative example is shown in Scheme 16. This was rationalized by higher reactivities of the dioxalenium ion than glycosyl triflate towards nucleophilic attack by the acceptor.

p-TolSCl/AgOTf is a superior promoter system for the preactivation-based thioglycoside glycosylation. Some reactions that failed with the BSP/Tf_2O promoter could be successfully performed with similar substrates using *p*-TolSCl/AgOTf (Scheme 17). This is presumably due to the inertness of the ditolyl disulfide side product from *p*-TolSCl/AgOTf promoted activation, which does not interfere with glycosylation.



Scheme 15: Proposed mechanism for preactivation-based glycosylation strategy.

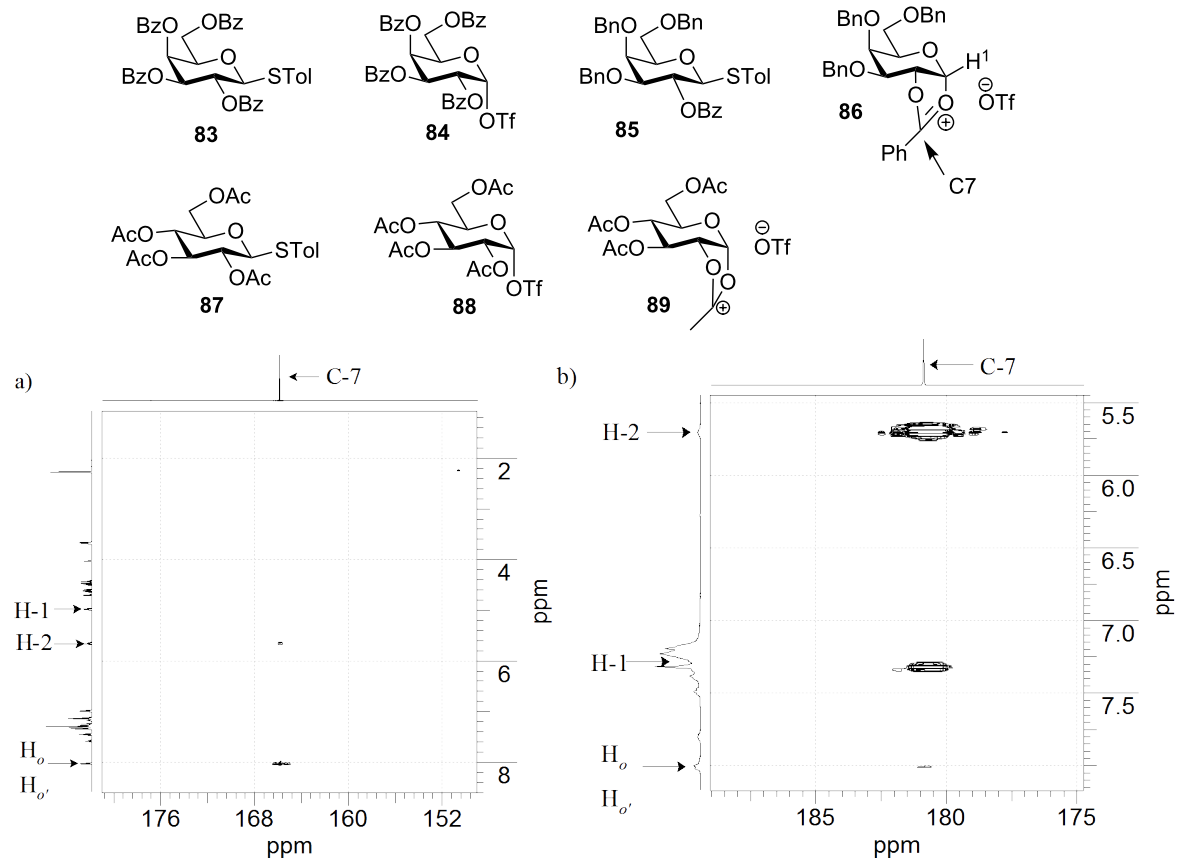
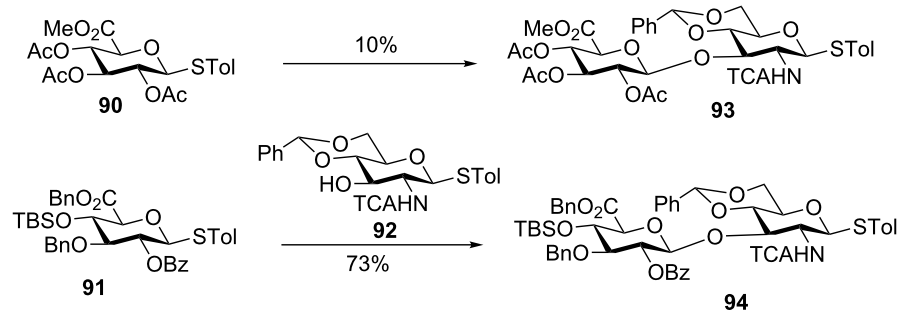
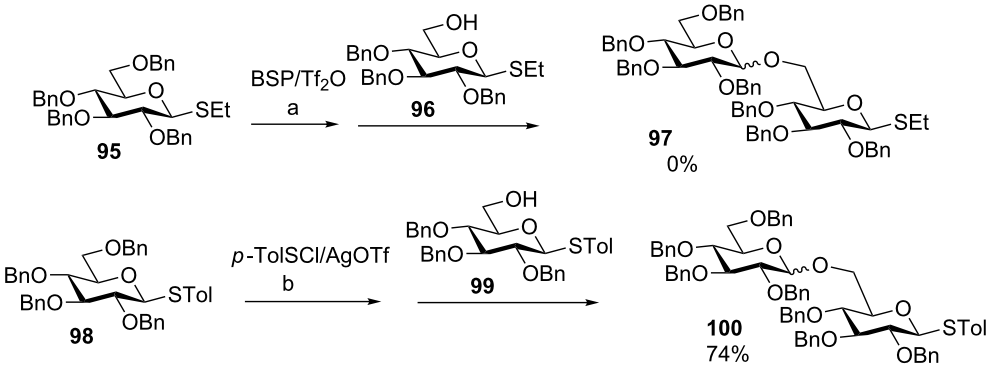


Figure 3: The preactivations of glycosyl donors **83**, **85** and **87** were investigated by low temperature NMR, which gave **84**, **86**, **88**/**89** as dominant intermediates, respectively. gHMBC ($CDCl_3$, 600 MHz) of donor **85** a) before and b) after preactivation at −60 °C. The correlation peak emerged after activation between C-7 and H-1 supports the structure of the dioxalenium ion **86** formed from preactivation.

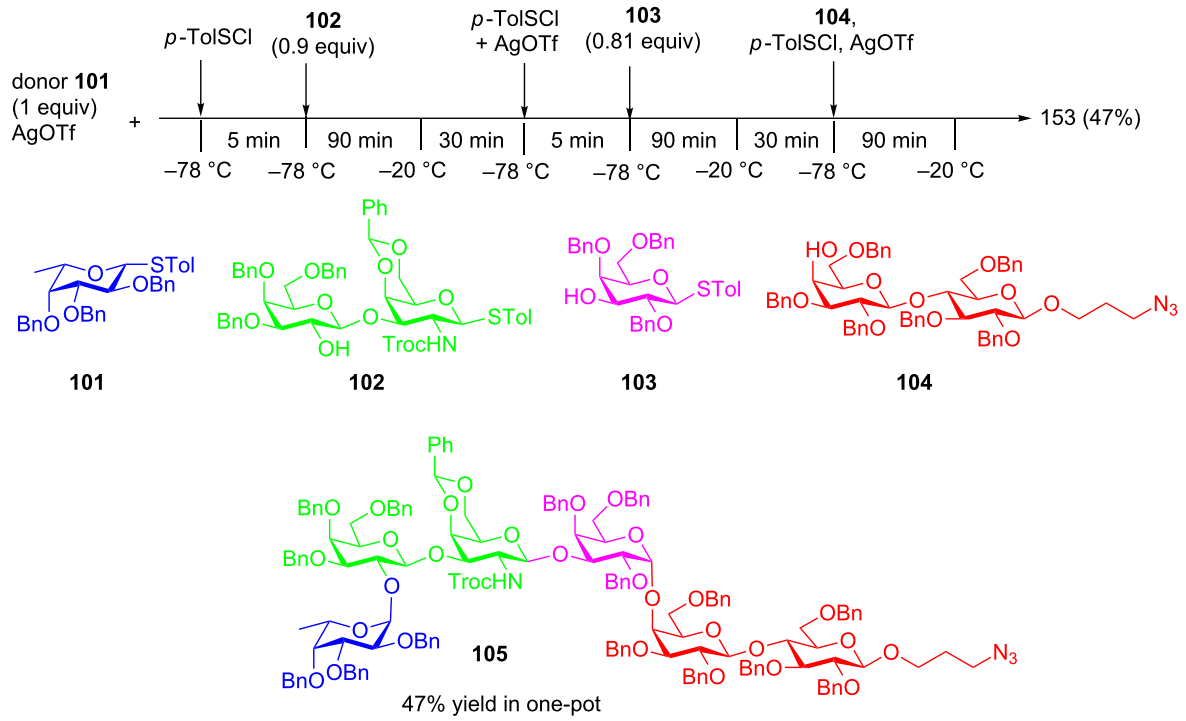


Scheme 16: The more electron-rich glycosyl donor **91** gave a higher glycosylation yield than the glycosyl donor **90** bearing more electron-withdrawing acyl protective groups.

The *p*-TolSCl/AgOTf-promoted preactivation glycosylation has been successfully applied to the total synthesis of complex oligosaccharides including those containing both 1,2-*cis* and 1,2-*trans* linkages, branching sequences and sulfate esters. For example, a four component preactivation-based one-pot synthesis was designed to synthesize Globo-H, an important tumor-associated carbohydrate antigen (Scheme 18) [60]. Globo-H hexasaccharide **105** was prepared within 7 hours in an excellent overall yield of 47% from the sequential one-pot reaction of **101**, **102**, **103** and **104**. Compared to the automated solid-phase synthesis of Globo-H [61], the solution-based preactivation-based synthesis gave a higher overall yield for glyco-assembly (47% vs 30%) without the need for large excess (5–10 equiv) of building blocks.

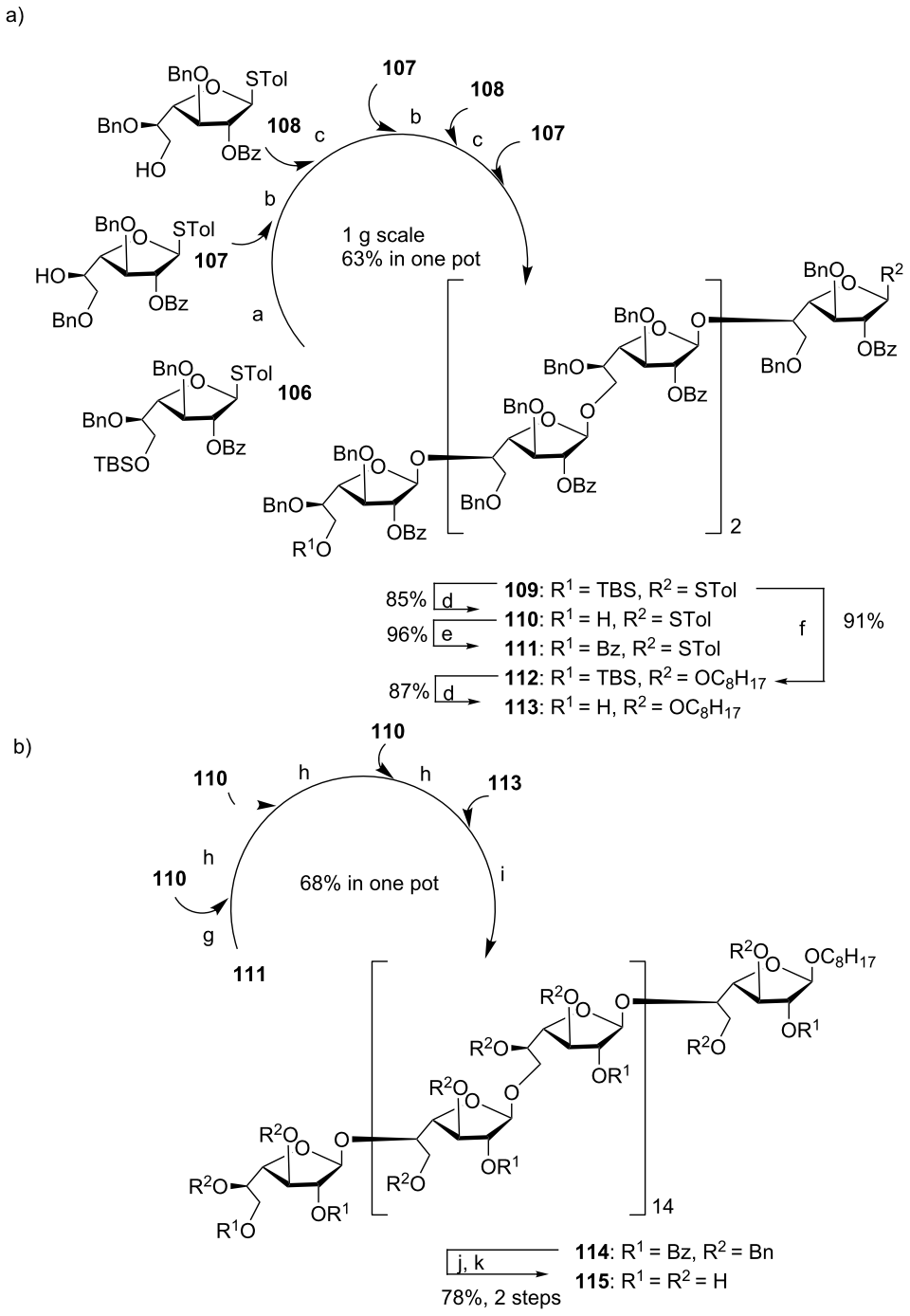


Scheme 17: Comparison of the BSP/Tf_2O and *p*-TolSCl/AgOTf promoter systems in facilitating the preactivation-based thioglycoside glycosylation. Reagents and conditions: (a) BSP, Tf_2O, CH_2Cl_2, TTBP, −60 °C; then **96**, and triethyl phosphite; (b) *p*-TolSCl/AgOTf, −60 °C; then **99**.



Scheme 18: One-pot synthesis of Globo-H hexasaccharide **105** using building blocks **101**, **102**, **103** and **104**.

Recently, using a series of highly efficient preactivation-based glycosylation reactions, Ye and co-workers synthesized a mycobacterial arabinogalactan [62], which is composed of 30 D-galactofuranose residues (Galf_{30}) linked with two arabinan chains each containing 31 D-arabinofuranose residues (Araf_{31}). Both Galf_{30} and Araf_{31} fragments were prepared starting from monosaccharide building blocks. As an example, a six component preactivation-based glycosylation using the *p*-TolSCl/AgOTf promoter system and three monosaccharide building blocks (**106**–**108**) led to the formation of hexasaccharide **109** in an excellent 63% yield in one pot on a gram scale (Scheme 19a). This is the largest number of glycosylation reactions that have been performed in one pot to date. Further iterative five-component one-pot glycosylation (**111** + **110** + **110** + **110** + **113**) successfully produced protected Galf_{30} 30-mer **114** in 68% yield (Scheme 19b). Following similar reaction protocols, Araf_{31} was prepared, which upon glycosylation of a Galf_{30} diol acceptor and deprotection, led to arabinogalactan 92-mer **116** (Figure 4) [62]. This is the largest synthetic glycan that has ever been produced, highlighting the power of the preactivation-based glycosylation strategy.



Scheme 19: Synthesis of (a) oligosaccharides **109–113** towards (b) 30-mer galactan **115.** Reagents and conditions: (a) TTBP, 4 Å MS, CH_2Cl_2, *p*-TolSCl, AgOTf, then **107**, −78 °C to rt; (b) *p*-TolSCl, AgOTf, then **108**, −78 °C to rt; (c) *p*-TolSCl, AgOTf, then **107**, −78 °C to rt; (d) HF-pyridine, THF/H_2O (10:1), 35 °C; (e) Bz_2O, DMAP, pyridine, CH_2Cl_2, reflux; (f) *p*-TolSCl, AgOTf, TTBP, 1-octanol, 4 Å MS, CH_2Cl_2, −78 °C; (g) TTBP, 4 Å MS, CH_2Cl_2, *p*-TolSCl, AgOTf, then **110**, −78 °C to rt; (h) *p*-TolSCl, AgOTf, then **110**, −78 °C to rt; (i) *p*-TolSCl, AgOTf, then **113**, −78 °C to rt; (j) $NaOCH_3$, CH_3OH/CH_2Cl_2 (2:1); (k) Pd/C, H_2, EtOAc/THF/1-PrOH/H_2O (2:1:1:1).

In addition to Globo-H **105** and arabinogalactan **116**, other complex oligosaccharides obtained by the preactivation-based thioglycoside method include branched oligosaccharides from glycolipid family including Lewis^X pentasaccharide **117**, dimeric Lewis^X **118** [63], tristearoyl lipomannan **119** [64], gangliosides GM1 **120** [65] and GM2 **121** (Figure 5) [66], microbial glycans such as the heptasaccharide repeating unit of type V group B Streptococcus capsular polysaccharide **122**

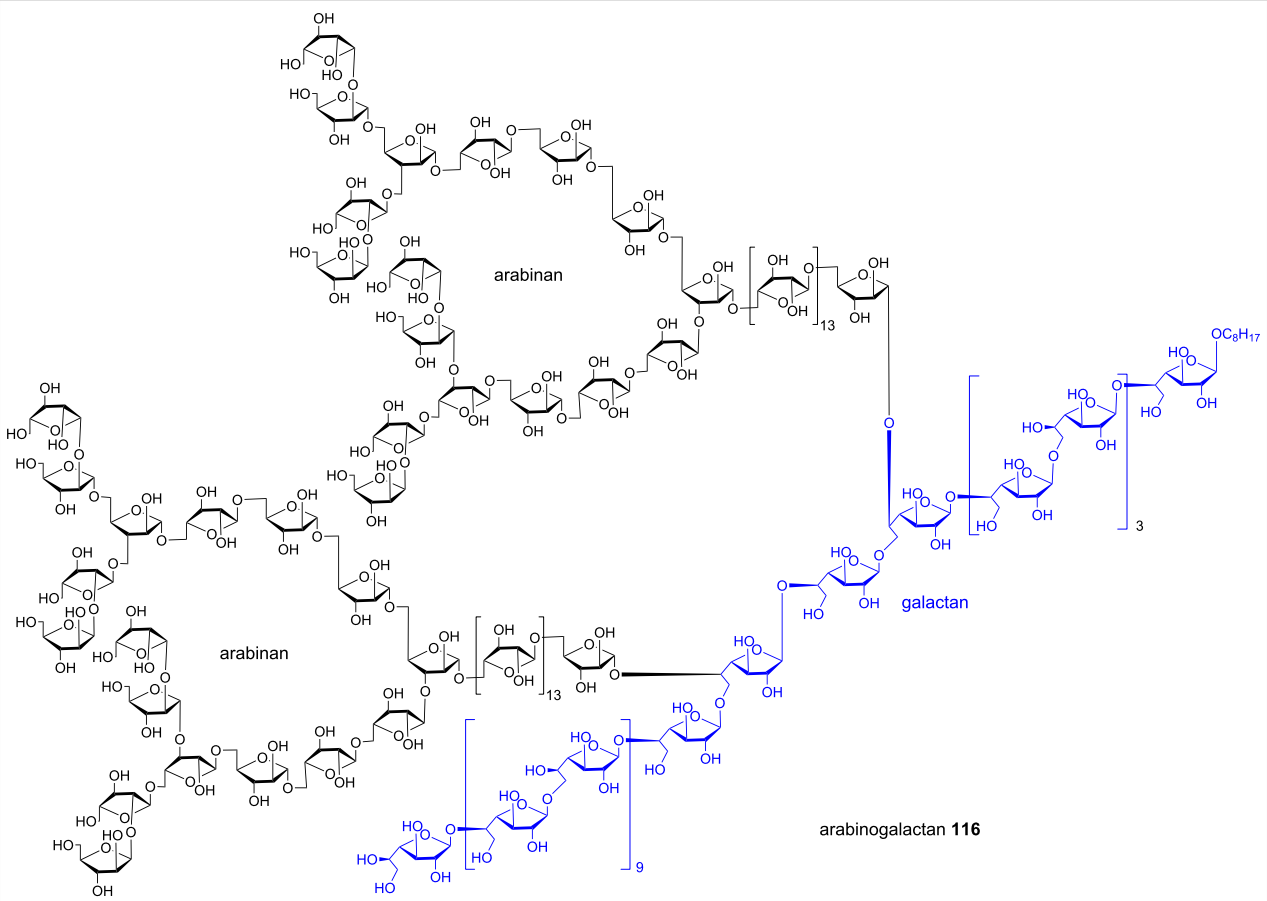


Figure 4: Structure of mycobacterial arabinogalactan **116**.

[67], β-glucan oligosaccharides **123** from fungal cells [68,69], oligomannan containing multiple challenging β-mannosyl linkages **124** [54] (Figure 6), chitotetraose [70], mammalian glycans including complex type bisected N-glycan dodecasaccharide **125** [71], glycosaminoglycans including hyaluronic acid oligosaccharides **126** [72,73] (Figure 7), and heparan sulfate oligosaccharides including those bearing sulfate esters [74,75] and other sialylated glycans [76,77].

As the preactivation-based glycosylation does not require the donor to have higher anomeric reactivity than the acceptor, this approach is particularly suitable for the synthesis of libraries of oligosaccharides by divergently combining building blocks. An example of this is the preparation of a library of heparan sulfate oligosaccharides (Figure 8) [74]. Alternating use of disaccharide building blocks **127** and **128** in preactivation-based one-pot glycosylation led to a panel of 7 heparan sulfate hexasaccharides **129**–**135** following the standard glycosylation protocol. The yields for one-pot glycosylation of all these hexasaccharides range from 50% to 70% highlighting the robustness of the protocol.

Besides the more “classical” chemical activation of thioglycosides, Nokami, Yoshida and co-workers developed an alternative method taking advantage of electrochemistry for donor activation [59]. They have demonstrated that thioglycosides can be electrochemically oxidized in the presence of tetrabutylammonium triflate to yield a glycosyl triflate, which can be subsequently glycosylated. This approach has been adapted to an automated solution-phase synthesis of poly-β-D-(1-6)-*N*-acetylglucosamine [78]. The aryl group in arylthioglycosides was first optimized for both the donor and the acceptor, where the electron-withdrawing fluorine on the phenyl ring gave the best result. The thioglycoside donor **136** was preactivated through anodic oxidation, followed by the addition of the acceptor **137** to afford disaccharide **138** (Scheme 20). Repeating this process, a series of oligo-glucosamine ranging from tri- to hexa-saccharides **139**–**142** was successfully prepared.

2-Deoxy and 2,6-dideoxyglycosides are present in many natural products. Based on the preactivation protocol, the Wang group reported a stereoselective glycosyl chloride-mediated synthesis of 2-deoxyglucosides [79]. They found that the addition of AgOTf and *p*-TolSCl to donor **143** afforded the stable glycosyl

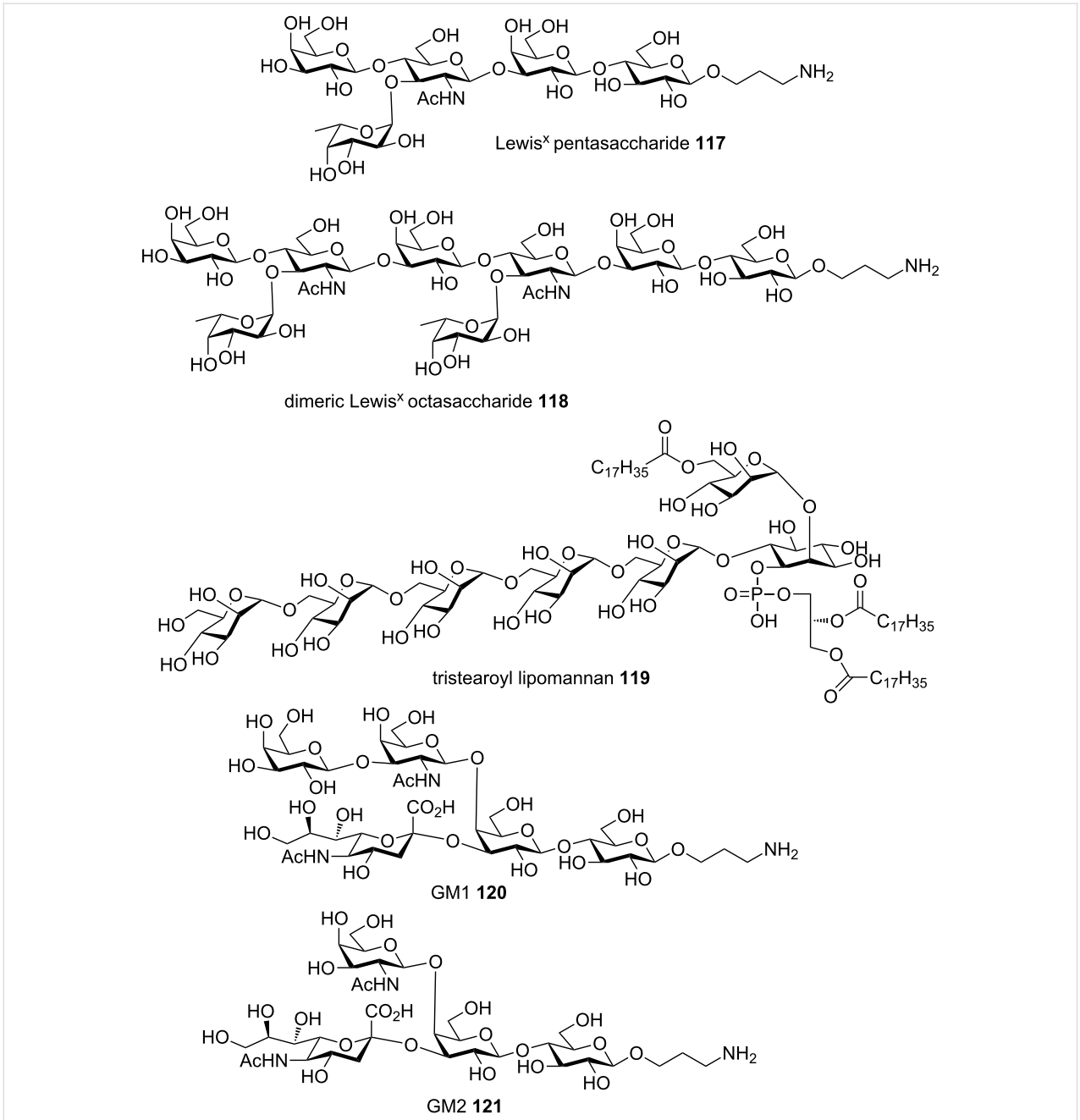
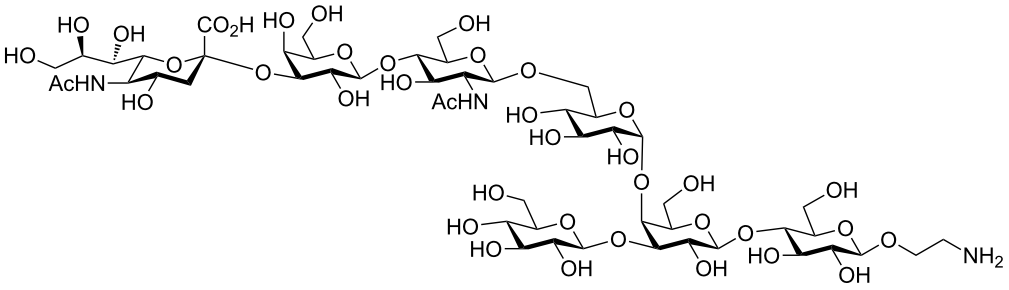


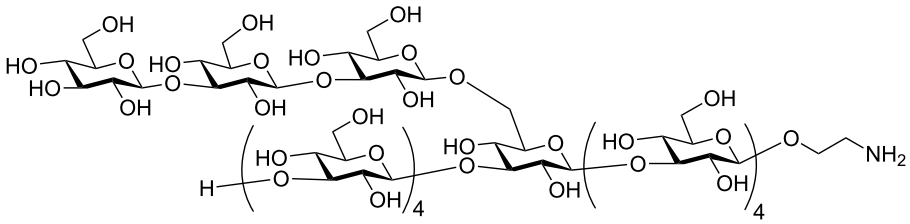
Figure 5: Representative complex glycans from glycolipid family synthesized by the preactivation-based thioglycoside method.

chloride **144** as detected by NMR (Scheme 21a). The formation of the glycosyl chloride was possibly due to the presence of Lewis basic molecule sieves (MS 4 Å) in the reaction system lowering the reactivity of AgOTf [18]. As a result, *p*-TolSCl could directly activate the glycosyl donor forming glycosyl chloride due to the higher anomeric reactivities of deoxy glycosides compared to the corresponding pyranosides. Upon the addition of the acceptor, the glycosyl chloride could be activated by AgOTf producing the glycosylation product with good α selectivity. To test the applicability to iterative synthesis, donor **143** was preactivated with *p*-TolSCl and AgOTf at −78 °C followed by the addition of acceptor **146** to afford disaccharide **147** in 70% yield with complete α selectivity (Scheme 21b). This high α selectivity remained when disaccharide **147** was reacted with acceptor **148** to give trisaccharide **149** using the same promoter system.

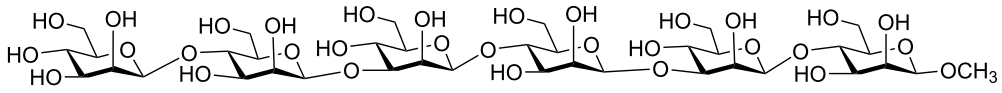
The preactivation-based one-pot approach can greatly accelerate oligosaccharide assembly. To facilitate isolation of the desired product from the reaction mixture, the Huang group re-



heptasaccharide repeating unit of type V group B Streptococcus capsular polysaccharide **122**

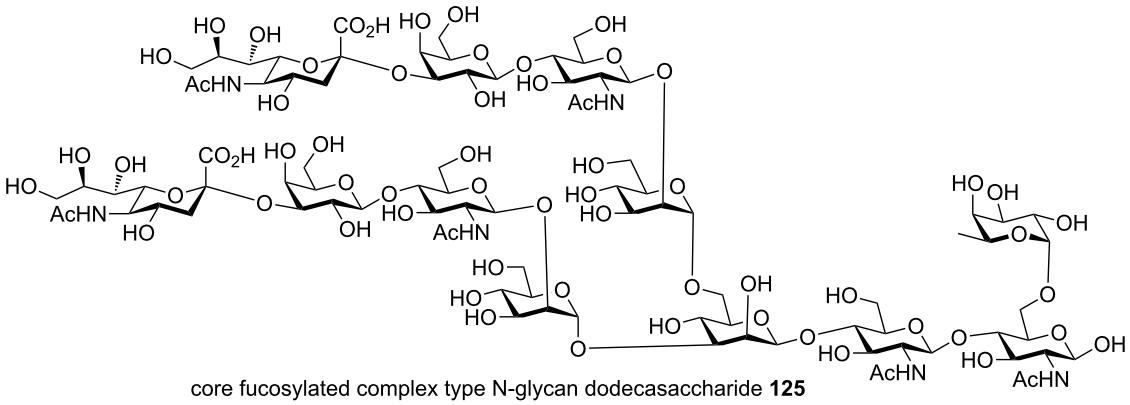


branched β-glucan oligosaccharide **123**

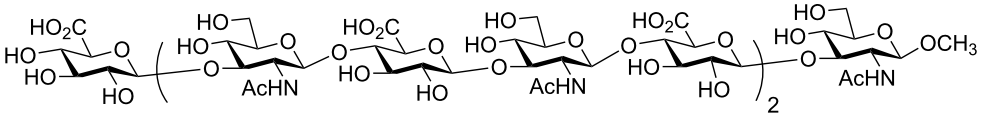


oligomannan **124** containing multiple challenging β-mannosyl linkages

Figure 6: Representative microbial and mammalian oligosaccharides synthesized by the preactivation-based thioglycoside method.



core fucosylated complex type N-glycan dodecasaccharide **125**



hyaluronan decasaccharide **126**

Figure 7: Some representative mammalian oligosaccharides synthesized by the preactivation-based thioglycoside method.

ported a fluorine-assisted one-pot method, where no silica gel column chromatography was required [80]. To demonstrate the applicability of this method, a linear tetrasaccharide was synthesized bearing a ketone tag at the reducing end using building blocks **83**, **150** and **151** following the preactivation-based one-pot protocol (Scheme 22). After completion of the synthesis, a

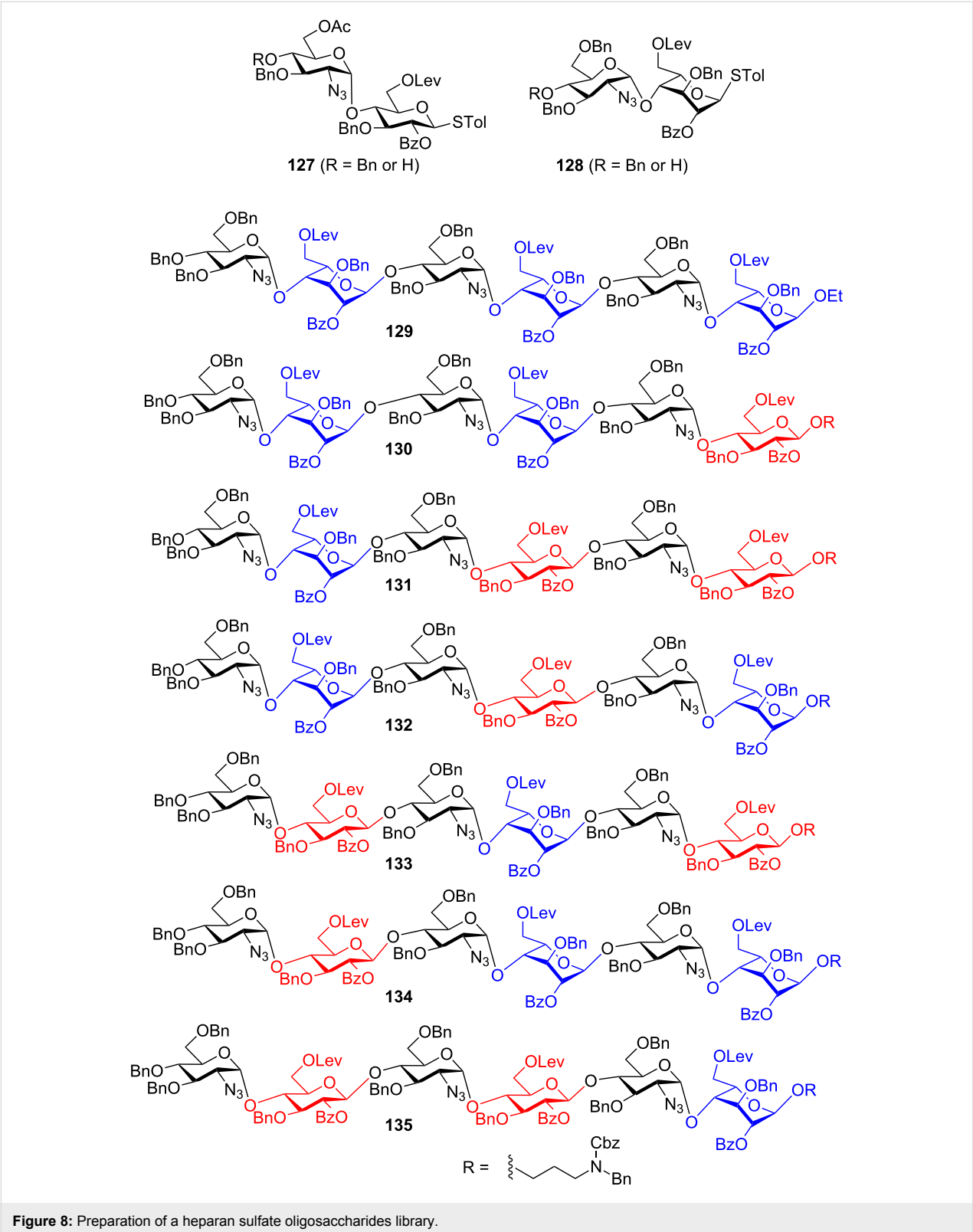
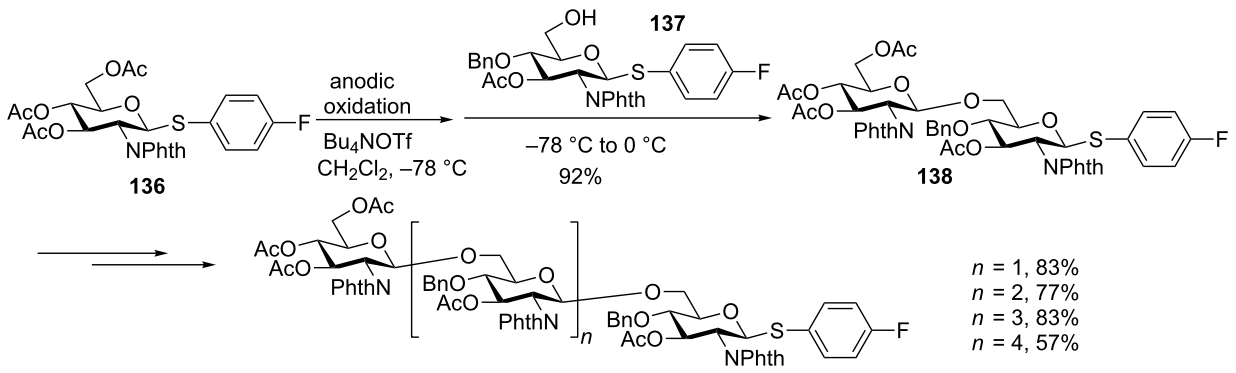
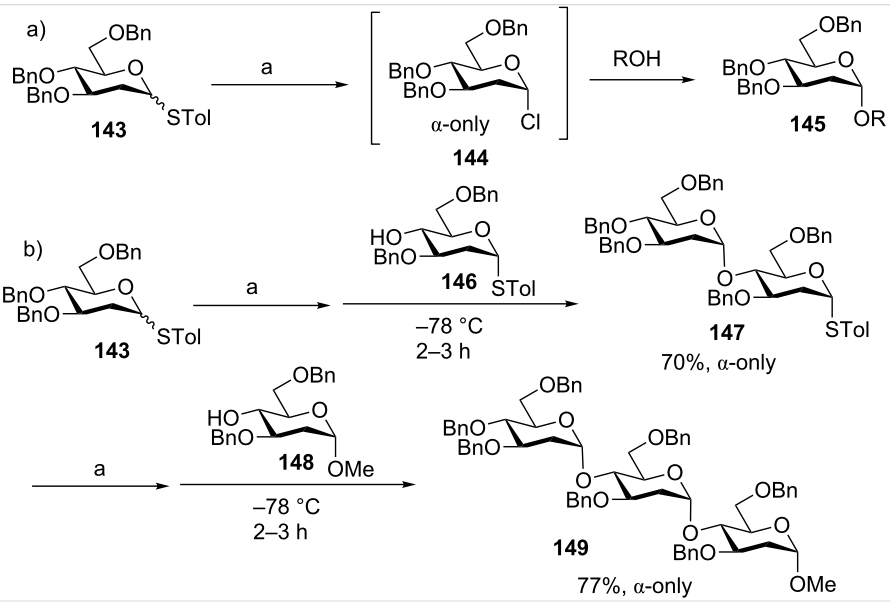


Figure 8: Preparation of a heparan sulfate oligosaccharides library.

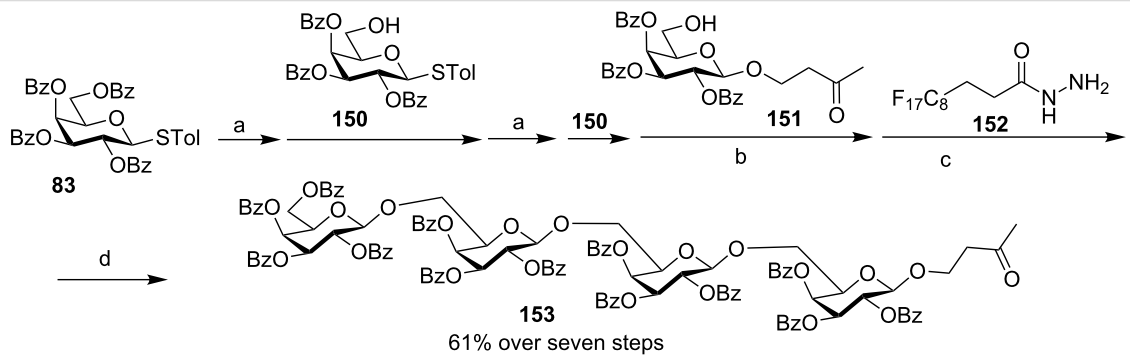
fluorinated hydrazide **152** was added to the reaction mixture to selectively "catch" the desired tetrasaccharide **153**, which was rapidly separated from non-fluorinated impurities by fluorous solid-phase extraction (F-SPE). Subsequent release of the compound from the fluorous tag and F-SPE yielded pure **153** in 61% overall yield from donor **83**.



Scheme 20: Synthesis of oligo-glucosamines through electrochemical promoted preactivation-based thioglycoside glycosylation.

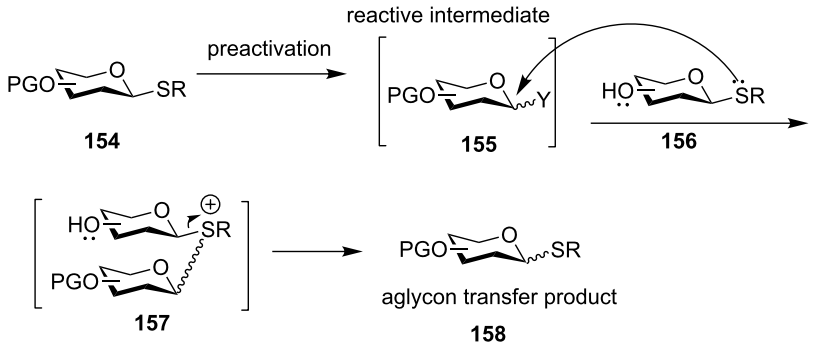


Scheme 21: Synthesis of 2-deoxyglucosides through preactivation. Reagents and conditions: a) AgOTf, *p*-TolSCl, −78 °C.



Scheme 22: Synthesis of tetrasaccharide **153**. Reagents and conditions: (a) AgOTf, *p*-TolSCl, CH_2Cl_2, −78 °C; then **150**; (b) AgOTf, *p*-TolSCl, CH_2Cl_2, **150**, −78 °C to rt; (c) CH_2Cl_2/MeOH, then F-SPE; (d) acetone/trifluoroacetic acid, then F-SPE.

One potential side reaction in using a thioglycosyl acceptor is the transfer of the thioaglycon of the acceptor to the activated donor presumably due to the high nucleophilicity of the aglycon compared with the hydroxy group of the acceptor (Scheme 23). Occasionally, the donor could be found regenerated upon addition of the acceptor following preactivation. This aglycon



Scheme 23: Aglycon transfer from a thioglycosyl acceptor to an activated donor can occur during preactivation-based glycosylation reaction. This side reaction can be suppressed by tuning the reactivity of acceptor aglycon or manipulating the reaction temperature.

transfer phenomenon is not restricted to preactivation or thioglycosyl donors, as aglycon transfer products have been reported in premixed glycosylations with either glycosyl bromide or glycosyl trichloroacetimidate (Scheme 11b) [81-86]. The amounts of aglycon transfer products can be reduced by decreasing the nucleophilicity of the acceptor aglycon through steric effects [87] or tuning protective groups of acceptors [84,86], in some cases by lowering the reaction temperature [85].

Conclusion

While conceptually simple, the temporal separation of donor activation and acceptor glycosylation in the preactivation protocol can enable chemoselective activation of the glycosyl donor without undesired acceptor activation. As a result, even an acceptor having higher anomeric reactivities than the glycosyl donor can be successfully glycosylated [18]. This protocol is found to be applicable to a wide range of glycosyl-donor types including thioglycosides, glycosyl sulfoxides, glycosyl hemiacetals, selenoglycosides, and 2-pyridyl glycosides. The newly formed oligosaccharide intermediate could be directly subjected to another round of preactivation and acceptor glycosylation without the need for additional synthetic operations to modify either protective groups or aglycon leaving groups. This can enable rapid glycan chain extension and improve overall synthetic efficiencies for glycan assembly.

Compared to the more traditional premixed method where both the glycosyl donor and the acceptor are present when the promoter is added, preactivation can generate reactive intermediates as the resting state allowing spectroscopic analysis such as low temperature NMR studies to better characterize the intermediate. This can help gaining a deeper insight into the reaction mechanism, which is critical for efforts to enhance the glycosylation yield.

The preactivation strategy is a powerful method for glycoassembly, which is evident from the successful synthesis of many complex oligosaccharides and glycoconjugates. However, glycosylation reactions are intrinsically sensitive to factors including protective groups on the glycan ring, reaction solvent, and additives present. As a result, further experimentation and analysis are needed to enable robust syntheses and achieve automation with comparable efficiencies of automated peptide and nucleic acid synthesis. With continuous development, the preactivation strategy will achieve wider applications in complex carbohydrate synthesis.

Acknowledgements

We are grateful for the financial support from National Science Foundation (CHE 1507226) and the National Institute of General Medical Sciences, NIH (R01GM072667).

References

1. Varki, A. *Glycobiology* **1993,** *3,* 97–130. doi:10.1093/glycob/3.2.97
2. Dwek, R. A. *Chem. Rev.* **1996,** *96,* 683–720. doi:10.1021/cr940283b
3. Bertozzi, C. R.; Kiessling, L. L. *Science* **2001,** *291,* 2357–2364. doi:10.1126/science.1059820
4. Zhu, X.; Schmidt, R. R. *Angew. Chem., Int. Ed.* **2009,** *48,* 1900–1934. doi:10.1002/anie.200802036
5. Huang, X.; Wang, Z. Strategies in Oligosaccharide Synthesis. In *Comprehensive Glycoscience from Chemistry to Systems Biology;* Kamerling, J. P., Ed.; Elsevier, 2007; Vol. 1, pp 379–413. doi:10.1016/B978-044451967-2/00011-8
6. Johnson, M.; Arles, C.; Boons, G.-J. *Tetrahedron Lett.* **1998,** *39,* 9801–9804. doi:10.1016/S0040-4039(98)02177-7
7. Roy, R.; Andersson, F. O.; Letellier, M. *Tetrahedron Lett.* **1992,** *33,* 6053–6056. doi:10.1016/S0040-4039(00)60004-7
8. Kim, K. S.; Kim, J. H.; Lee, Y. J.; Lee, Y. J.; Park, J. *J. Am. Chem. Soc.* **2001,** *123,* 8477–8481. doi:10.1021/ja015842s
9. Demchenko, A. V.; Pornsuriyasak, P.; De Meo, C.; Malysheva, N. N. *Angew. Chem., Int. Ed.* **2004,** *43,* 3069–3072. doi:10.1002/anie.200454047

10. Kanie, O.; Ito, Y.; Ogawa, T. *J. Am. Chem. Soc.* **1994,** *116,* 12073–12074. doi:10.1021/ja00105a066
11. Zhang, Z.; Ollmann, I. R.; Ye, X.-S.; Wischnat, R.; Baasov, T.; Wong, C.-H. *J. Am. Chem. Soc.* **1999,** *121,* 734–753. doi:10.1021/ja982232s
12. Douglas, N. L.; Ley, S. V.; Lücking, U.; Warriner, S. L. *J. Chem. Soc., Perkin Trans. 1* **1998,** 51–66. doi:10.1039/a705275h
13. Mootoo, D. R.; Konradsson, P.; Udodong, U.; Fraser-Reid, B. *J. Am. Chem. Soc.* **1988,** *110,* 5583–5584. doi:10.1021/ja00224a060
14. Huang, L.; Wang, Z.; Huang, X. *Chem. Commun.* **2004,** 1960–1961. doi:10.1039/b405886k
15. Miura, T.; Hirose, Y.; Ohmae, M.; Inazu, T. *Org. Lett.* **2001,** *3,* 3947–3950. doi:10.1021/ol016838o
16. Roychoudhury, R.; Pohl, N. L. B. Light Fluorous-Tag-Assisted Synthesis of Oligosaccharides. In *Modern Synthetic Methods in Carbohydrate Chemistry: From Monosaccharides to Complex Glycoconjugates;* Werz, D. B.; Vidal, S., Eds.; Wiley-VCH: Weinheim, 2013; pp 221–239. doi:10.1002/9783527658947.ch8
17. Plante, O. J.; Palmacci, E. R.; Seeberger, P. H. *Science* **2001,** *291,* 1523–1527. doi:10.1126/science.1057324
18. Huang, X.; Huang, L.; Wang, H.; Ye, X.-S. *Angew. Chem., Int. Ed.* **2004,** *43,* 5221–5224. doi:10.1002/anie.200460176
19. Codée, J. D. C.; van den Bos, L. J.; Litjens, R. E. J. N.; Overkleeft, H. S.; van Boeckel, C. A. A.; van Boom, J. H.; van der Marel, G. A. *Tetrahedron* **2004,** *60,* 1057–1064. doi:10.1016/j.tet.2003.11.084
20. Yamago, S.; Yamada, T.; Maruyama, T.; Yoshida, J.-i. *Angew. Chem., Int. Ed.* **2004,** *43,* 2145–2148. doi:10.1002/anie.200353552
21. Nguyen, H. M.; Poole, J. L.; Gin, D. Y. *Angew. Chem., Int. Ed.* **2001,** *40,* 414–417. doi:10.1002/1521-3773(20010119)40:2<414::AID-ANIE414>3.0.CO;2-6
22. Codée, J. D. C.; Litjens, R. E. J. N.; van den Bos, L. J.; Overkleeft, H. S.; van der Marel, G. A. *Chem. Soc. Rev.* **2005,** *34,* 769–782. doi:10.1039/b417138c
23. Wang, Y.; Ye, X.-S.; Zhang, L.-H. *Org. Biomol. Chem.* **2007,** *5,* 2189–2200. doi:10.1039/b704586g
24. Bouhall, S. K.; Sucheck, S. J. *J. Carbohydr. Chem.* **2014,** *33,* 347–367. doi:10.1080/07328303.2014.931964
25. Kahne, D.; Walker, S.; Cheng, Y.; Van Engen, D. *J. Am. Chem. Soc.* **1989,** *111,* 6881–6882. doi:10.1021/ja00199a081
26. Gargiulo, D.; Blizzard, T. A.; Nakanishi, K. *Tetrahedron* **1989,** *45,* 5423–5432. doi:10.1016/S0040-4020(01)89488-6
27. Yamago, S.; Yamada, T.; Hara, O.; Ito, H.; Mino, Y.; Yoshida, J.-i. *Org. Lett.* **2001,** *3,* 3867–3870. doi:10.1021/ol016713j
28. Yamago, S.; Yamada, T.; Ito, H.; Hara, O.; Mino, Y.; Yoshida, J.-i. *Chem. – Eur. J.* **2005,** *11,* 6159–6174. doi:10.1002/chem.200500126
29. Hanessian, S.; Lou, B. *Chem. Rev.* **2000,** *100,* 4443–4464. doi:10.1021/cr9903454
30. Xiong, D.-C.; Yang, A.-Q.; Yu, Y.; Ye, X.-S. *Tetrahedron Lett.* **2015,** *56,* 211–214. doi:10.1016/j.tetlet.2014.11.066
31. Hanessian, S. *Preparative Carbohydrate Chemistry;* Marcel Dekker Inc.: New York, 1997.
32. Toshima, K.; Tatsuta, K. *Chem. Rev.* **1993,** *93,* 1503–1531. doi:10.1021/cr00020a006
33. Sinaÿ, P. *Pure Appl. Chem.* **1991,** *63,* 519–528. doi:10.1351/pac199163040519
34. Danishefsky, S. J.; Bilodeau, M. T. *Angew. Chem., Int. Ed. Engl.* **1996,** *35,* 1380–1419. doi:10.1002/anie.199613801
35. Schmidt, R. R.; Kinzy, W. *Adv. Carbohydr. Chem. Biochem.* **1994,** *50,* 21–123. doi:10.1016/S0065-2318(08)60150-X
36. Garcia, B. A.; Poole, J. L.; Gin, D. Y. *J. Am. Chem. Soc.* **1997,** *119,* 7597–7598. doi:10.1021/ja971067y
37. Garcia, B. A.; Gin, D. Y. *J. Am. Chem. Soc.* **2000,** *122,* 4269–4279. doi:10.1021/ja993595a
38. Codée, J. D. C.; van den Bos, L. J.; Litjens, R. E. J. N.; Overkleeft, H. S.; van Boom, J. H.; van der Marel, G. A. *Org. Lett.* **2003,** *5,* 1947–1950. doi:10.1021/ol034528v
39. Dinkelaar, J.; Codée, J. D. C.; van den Bos, L. J.; Overkleeft, H. S.; van der Marel, G. A. *J. Org. Chem.* **2007,** *72,* 5737–5742. doi:10.1021/jo070704s
40. Codée, J. D. C.; Stubba, B.; Schiattarella, M.; Overkleeft, H. S.; van Boeckel, C. A. A.; van Boom, J. H.; van der Marel, G. A. *J. Am. Chem. Soc.* **2005,** *127,* 3767–3773. doi:10.1021/ja045613g
41. Garegg, P. J. *Adv. Carbohydr. Chem. Biochem.* **1997,** *52,* 179–205. doi:10.1016/S0065-2318(08)60091-8
42. Koeller, K. M.; Wong, C.-H. *Chem. Rev.* **2000,** *100,* 4465–4493. doi:10.1021/cr990297n and references cited therein.
43. Geurtsen, R.; Holmes, D. S.; Boons, G.-J. *J. Org. Chem.* **1997,** *62,* 8145–8154. doi:10.1021/jo971233k
44. Li, X.; Huang, L.; Hu, X.; Huang, X. *Org. Biomol. Chem.* **2009,** *7,* 117–127. doi:10.1039/B813048E
45. Jennum, C. A.; Fenger, T. H.; Bruun, L. M.; Madsen, R. *Eur. J. Org. Chem.* **2014,** 3232–3241. doi:10.1002/ejoc.201400164
46. Hsu, Y.; Lu, X.-A.; Zulueta, M. M. L.; Tsai, C.-M.; Lin, K.-I.; Hung, S.-C.; Wong, C.-H. *J. Am. Chem. Soc.* **2012,** *134,* 4549–4552. doi:10.1021/ja300284x
47. Burkhart, F.; Zhang, Z.; Wacowich-Sgarbi, S.; Wong, C.-H. *Angew. Chem., Int. Ed.* **2001,** *40,* 1274–1277. doi:10.1002/1521-3773(20010401)40:7<1274::AID-ANIE1274>3.0.CO;2-W
48. Polat, T.; Wong, C.-H. *J. Am. Chem. Soc.* **2007,** *129,* 12795–12800. doi:10.1021/ja073098r
49. Crich, D.; Smith, M. *J. Am. Chem. Soc.* **2001,** *123,* 9015–9020. doi:10.1021/ja0111481
50. Crich, D.; Smith, M. *Org. Lett.* **2000,** *2,* 4067–4069. doi:10.1021/ol006715o
51. Codée, J. D. C.; Litjens, R. E. J. N.; den Heeten, R.; Overkleeft, H. S.; van Boom, J. H.; van der Marel, G. A. *Org. Lett.* **2003,** *5,* 1519–1522. doi:10.1021/ol034312t
52. Peng, P.; Ye, X.-S. *Org. Biomol. Chem.* **2011,** *9,* 616–622. doi:10.1039/C0OB00380H
53. Wang, C.; Wang, H.; Huang, X.; Zhang, L.-H.; Ye, X.-S. *Synlett* **2006,** 2846–2850. doi:10.1055/s-2006-950247
54. Crich, D.; Li, W.; Li, H. *J. Am. Chem. Soc.* **2004,** *126,* 15081–15086. doi:10.1021/ja0471931
55. Son, S.-H.; Tano, C.; Furuike, T.; Sakairi, N. *Carbohydr. Res.* **2009,** *344,* 285–290. doi:10.1016/j.carres.2008.11.008
56. Crich, D.; Sun, S. *J. Am. Chem. Soc.* **1997,** *119,* 11217–11223. doi:10.1021/ja971239r
57. Zeng, Y.; Wang, Z.; Whitfield, D.; Huang, X. *J. Org. Chem.* **2008,** *73,* 7952–7962. doi:10.1021/jo801462r
58. Suga, S.; Suzuki, S.; Yamamoto, A.; Yoshida, J.-i. *J. Am. Chem. Soc.* **2000,** *122,* 10244–10245. doi:10.1021/ja002123p
59. Nokami, T.; Shibuya, A.; Tsuyama, H.; Suga, S.; Bowers, A. A.; Crich, D.; Yoshida, J.-i. *J. Am. Chem. Soc.* **2007,** *129,* 10922–10928. doi:10.1021/ja072440x

60. Wang, Z.; Zhou, L.; El-Boubbou, K.; Ye, X.-S.; Huang, X. *J. Org. Chem.* **2007,** *72,* 6409–6420. doi:10.1021/jo070585g
61. Werz, D. B.; Castagner, B.; Seeberger, P. H. *J. Am. Chem. Soc.* **2007,** *129,* 2770–2771. doi:10.1021/ja069218x
62. Wu, Y.; Xiong, D.-C.; Chen, S.-C.; Wang, Y.-S.; Ye, X.-S. *Nat. Commun.* **2017,** *8,* No. 14851. doi:10.1038/ncomms14851
63. Miermont, A.; Zeng, Y.; Jing, Y.; Ye, X.-s.; Huang, X. *J. Org. Chem.* **2007,** *72,* 8958–8961. doi:10.1021/jo701694k
64. Gao, J.; Guo, Z. *J. Org. Chem.* **2013,** *78,* 12717–12725. doi:10.1021/jo4021979
65. Sun, B.; Yang, B.; Huang, X. *Sci. China: Chem.* **2012,** *55,* 31–35. doi:10.1007/s11426-011-4449-x
66. Yin, Z.; Dulaney, S.; McKay, C. S.; Baniel, C.; Kaczanowska, K.; Ramadan, S.; Finn, M. G.; Huang, X. *ChemBioChem* **2016,** *17,* 174–180. doi:10.1002/cbic.201500499
67. Gao, J.; Guo, Z. *Org. Lett.* **2016,** *18,* 5552–5555. doi:10.1021/acs.orglett.6b02796
68. Liao, G.; Zhou, Z.; Burgula, S.; Liao, J.; Yuan, C.; Wu, Q.; Guo, Z. *Bioconjugate Chem.* **2015,** *26,* 466–476. doi:10.1021/bc500575a
69. Liao, G.; Burgula, S.; Zhou, Z.; Guo, Z. *Eur. J. Org. Chem.* **2015,** 2942–2951. doi:10.1002/ejoc.201500229
70. Huang, L.; Wang, Z.; Li, X.; Ye, X.-s.; Huang, X. *Carbohydr. Res.* **2006,** *341,* 1669–1679. doi:10.1016/j.carres.2006.01.007
71. Sun, B.; Srinivasan, B.; Huang, X. *Chem. – Eur. J.* **2008,** *14,* 7072–7081. doi:10.1002/chem.200800757
72. Huang, L.; Huang, X. *Chem. – Eur. J.* **2007,** *13,* 529–540. doi:10.1002/chem.200601090
73. Lu, X.; Kamat, M. N.; Huang, L.; Huang, X. *J. Org. Chem.* **2009,** *74,* 7608–7617. doi:10.1021/jo9016925
74. Wang, Z.; Xu, Y.; Yang, B.; Tiruchinapally, G.; Sun, B.; Liu, R.; Dulaney, S.; Liu, J.; Huang, X. *Chem. – Eur. J.* **2010,** *16,* 8365–8375. doi:10.1002/chem.201000987
75. Tiruchinapally, G.; Yin, Z.; El-Dakdouki, M.; Wang, Z.; Huang, X. *Chem. – Eur. J.* **2011,** *17,* 10106–10112. doi:10.1002/chem.201101108
76. Sun, B.; Jiang, H. *Tetrahedron Lett.* **2011,** *52,* 6035–6038. doi:10.1016/j.tetlet.2011.09.022
77. Wang, C.-C.; Lee, J.-C.; Luo, S.-Y.; Kulkarni, S. S.; Huang, Y.-W.; Lee, C.-C.; Chang, K.-L.; Hung, S.-C. *Nature* **2007,** *446,* 896–899. doi:10.1038/nature05730
78. Nokami, T.; Hayashi, R.; Saigusa, Y.; Shimizu, A.; Liu, C.-Y.; Mong, K.-K. T.; Yoshida, J.-i. *Org. Lett.* **2013,** *15,* 4520–4523. doi:10.1021/ol402034g
79. Verma, V. P.; Wang, C.-C. *Chem. – Eur. J.* **2013,** *19,* 846–851. doi:10.1002/chem.201203418
80. Yang, B.; Jing, Y.; Huang, X. *Eur. J. Org. Chem.* **2010,** 1290–1298. doi:10.1002/ejoc.200901155
81. Auzanneau, F.-I.; Bundle, D. R. *Carbohydr. Res.* **1991,** *212,* 13–24. doi:10.1016/0008-6215(91)84041-C
82. Leigh, D. A.; Smart, J. P.; Truscello, A. M. *Carbohydr. Res.* **1995,** *276,* 417–424. doi:10.1016/0008-6215(95)00182-S
83. Belot, F.; Jacquinet, J.-C. *Carbohydr. Res.* **1996,** *290,* 79–86. doi:10.1016/0008-6215(96)00116-4
84. Zhu, T.; Boons, G.-J. *Carbohydr. Res.* **2000,** *329,* 709–715. doi:10.1016/S0008-6215(00)00252-4
85. Yu, H.; Yu, B.; Wu, X.; Hui, Y.; Han, X. *J. Chem. Soc., Perkin Trans. 1* **2000,** 1445–1453. doi:10.1039/a909218h
86. Li, Z.; Gildersleeve, J. C. *Tetrahedron Lett.* **2007,** *48,* 559–562. doi:10.1016/j.tetlet.2006.11.126
87. Li, Z.; Gildersleeve, J. C. *J. Am. Chem. Soc.* **2006,** *128,* 11612–11619. doi:10.1021/ja063247q

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

What contributes to an effective mannose recognition domain?

Christoph P. Sager, Deniz Eriş, Martin Smieško, Rachel Hevey and Beat Ernst*§

Review

Open Access

Address:
Department of Pharmaceutical Sciences, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

Email:
Beat Ernst* - beat.ernst@unibas.ch

* Corresponding author
§ phone: +41 61 267 15 51

Keywords:
carbohydrate–lectin interactions; desolvation penalty; dielectric constant; multivalency; pre-organization

Beilstein J. Org. Chem. **2017,** *13,* 2584–2595.
doi:10.3762/bjoc.13.255

Received: 03 August 2017
Accepted: 15 November 2017
Published: 04 December 2017

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

In general, carbohydrate–lectin interactions are characterized by high specificity but also low affinity. The main reason for the low affinities are desolvation costs, due to the numerous hydroxy groups present on the ligand, together with the typically polar surface of the binding sites. Nonetheless, nature has evolved strategies to overcome this hurdle, most prominently in relation to carbohydrate–lectin interactions of the innate immune system but also in bacterial adhesion, a process key for the bacterium's survival. In an effort to better understand the particular characteristics, which contribute to a successful carbohydrate recognition domain, the mannose-binding sites of six C-type lectins and of three bacterial adhesins were analyzed. One important finding is that the high enthalpic penalties caused by desolvation can only be compensated for by the number and quality of hydrogen bonds formed by each of the polar hydroxy groups engaged in the binding process. In addition, since mammalian mannose-binding sites are in general flat and solvent exposed, the half-lives of carbohydrate–lectin complexes are rather short since water molecules can easily access and displace the ligand from the binding site. In contrast, the bacterial lectin FimH benefits from a deep mannose-binding site, leading to a substantial improvement in the off-rate. Together with both a catch-bond mechanism (i.e., improvement of affinity under shear stress) and multivalency, two methods commonly utilized by pathogens, the affinity of the carbohydrate–FimH interaction can be further improved. Including those just described, the various approaches explored by nature to optimize selectivity and affinity of carbohydrate–lectin interactions offer interesting therapeutic perspectives for the development of carbohydrate-based drugs.

Review

Recognition of carbohydrate ligands

For the recognition of carbohydrate ligands, nature has explored binding sites of different shapes and properties. The large family of C-type lectins (CLECs) exhibits carbohydrate-recognition domains (CRDs) which incorporate a calcium ion [1-4]. CLECs are involved in a wide range of biological processes, such as pathogen recognition and intercellular adhesion [5-7]. A

large number of CLEC structures, including animal, plant and bacterial lectins, are available in the Protein Data Bank [8]. A second large family of lectins, the bacterial adhesins, play an important role in the initial interaction of the bacterium with host tissue [9,10]. This primary contact is a prerequisite for the infection of host cells and subsequent biofilm formation, and grants the bacteria a significant advantage by resisting clearance and killing by immune factors, bacteriolytic enzymes, or antibiotics.

In this review, with focus on lectins relevant for drug discovery and development, the mannose-binding sites of six CLECs and three bacterial lectins are analyzed and compared with one another to answer the question: What makes for a successful mannose recognition domain? In general, lectins are characterized by high ligand specificity, whereas the affinity for their carbohydrate ligands is comparatively low. A prominent example is sialyl Lewis^x (sLe^x), a tetrasaccharide typically O-linked to cell surfaces and known to play a vital role in cell-to-cell recognition processes [11]. Although highly specific, its interaction with E-selectin exhibits a dissociation constant (K_D) of only 800 μM [12]. To address this obstacle of low affinity, nature applies the principal of multivalency by providing several binding sites to the carbohydrate ligand and/or a multivalent display of the ligand [13-15]. This accumulation of individual binding events increases the overall binding strength either by avidity or local concentration effects [16,17]. However, other approaches, such as the reduction of desolvation costs or ligand and binding site pre-organization, are more difficult to assess and accordingly have been highlighted in this review.

Mannose-binding CLECs are involved in various pathways of the human innate immune response, including the blood dendritic cell antigen 2 (BDCA-2, also known as CD303) [18], langerin (CD207) [19,20], pulmonary surfactant-associated protein D (SP-D) [21], dendritic cell-specific ICAM-3-grabbing non-integrins 1 and 2 (DC-SIGN, also known as CD209; and DC-SIGNR, also known as CD299) [22,23], and mannose-binding protein (MBP) [24]. These CLECs exert their function through different mechanisms, for instance by pathogen internalization as in the case of BDCA-2 and langerin, by pathogen opsonization as mediated by SP-D and MBP, or by T-cell interactions as mediated by DC-SIGN and DC-SIGNR [25,26].

In contrast, pathogens have developed numerous adhesins that mediate their interaction with glycosides on mammalian cell surfaces. After this initial contact, they can infect host cells and form biofilms, both of which are key factors for their survival [9,27,28]. Examples of such opportunistic bacterial species binding to mannosides on host cells include *Pseudomonas aeruginosa* with its membrane lectin LecB [29,30] and *Burkholderia cenocepacia* with its characteristic *B. cenocepacia* lectin A (BC2L-A) [31,32], both playing an important role in the social life of bacterial cells. A further example is the bacterial adhesin FimH, which plays a crucial role in urinary tract infections (UTIs). FimH enables uropathogenic *Escherichia coli* (UPEC) to adhere to urothelial host cells [33,34], which represents the first and most critical step in UTI, triggering a cascade of pathogenic processes ultimately leading to infection. The ligand on urothelial cells binding to the N-terminal lectin domain of FimH is the highly mannosylated glycoprotein uroplakin 1a (UPIa) [35,36]. The binding pocket of FimH accommodates a single α-D-mannose (**1**) with an extended hydrogen-bond network [37,38]. Accordingly, any modifications on the hydroxy groups of the mannose virtually abolish binding affinity [37-39].

Crystal structures of mannose–lectin complexes

The X-ray structures of six mannose-binding receptors in complex with either α-D-mannose (**1**) or methyl α-D-mannopyranoside (**2**) were analyzed (Figure 1 and Table 1, **A**–**C** and **G**–**I**). Since for DC-SIGNR (Figure 1, **D**) and DC-SIGN (Figure 1, **E**) neither complexes with **1** nor **2** were available, we instead modeled the monosaccharide–receptor interactions based on the available oligomannose crystal structures (PDB codes: 1K9J and 1SL4). In addition, because none of the available crystal structures of human MBP met our threshold of a resolution below 2 Å, we used a structure based on a homologous MBP lectin domain from *Rattus norvegicus* and accordingly compared the measured binding affinity of rat MBP (Figure 1, **F**). Finally, a special case is the bacterial adhesin FimH, which can adopt three different affinity states (see below). For our discussion we focus specifically on the high-affinity state of FimH present in the isolated lectin domain of FimH, called $FimH_{LD}$ (Figure 1, **I**).

Although the receptors **A**–**F** play important roles in human immune responses, they exhibit affinities only in the millimolar range (9.4–1.3 mM) for α-D-mannose (**1**) and methyl α-D-mannopyranoside (**2**) [40-44]. In contrast, the receptors **G** and **H** of bacterial origin show affinities in the micromolar range (71 and 2.8 μM, respectively) for methyl α-D-mannose (**2**) [31,45]. Despite the 71 μM affinity, LecB (**G**) preferably binds L-fucose (3 μM) and methyl α-L-fucoside (0.4 μM) [45]. The enhanced affinity for fucosides originates from the C5-methyl group, absent in both **1** and **2**, which can form a hydrophobic contact with Thr45 [45].

The analyzed CLECs **A**–**F** share a common binding motif, with a calcium ion coordinating to O–C3 and O–C4 of the mannose

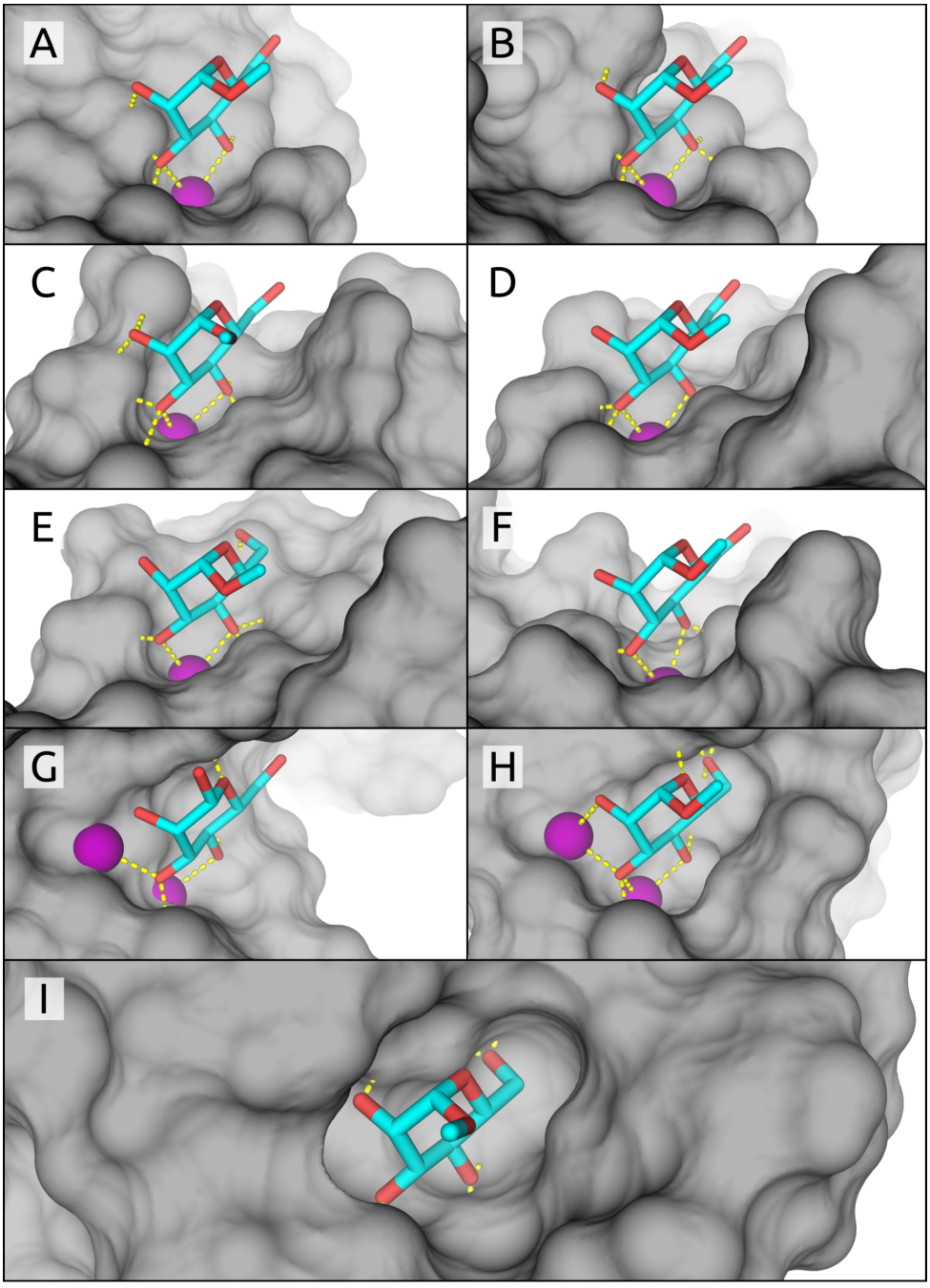


Figure 1: CRDs of the analyzed crystal structures, with mannose pyranosyl units similarly aligned in each structure. Mammalian lectins: (**A**) BDCA-2, (**B**) langerin, (**C**) SP-D, (**D**) DC-SIGNR, (**E**) DC-SIGN, (**F**) rat MBP; bacterial lectins: (**G**) LecB, (**H**) BC2L-A, and (**I**) FimH lectin domain.

ligand [5,7]. In instances where the binding site hosts a second calcium ion (**G** and **H**), advantageous interactions between O–C2 and O–C3 can also occur. Additional contributions from H-bonds formed in the buried binding pockets further improve affinity. In contrast, the calcium-free, buried binding site of the bacterial lectin FimH (**I**) forms a complex network of eight hydrogen bonds with ligand **2**, one of them mediated by a conserved water [37].

Various approaches to realize binding affinity

The immense variability of binding affinities among mannose-binding receptors is remarkable, albeit not surprising. While CRDs involved in the human immune system (Table 1, **A**–**F**) recognize a broader spectrum of binding partners (i.e., various pathogenic oligosaccharides), bacterial CRDs **G**–**I** strive for tight binding to host glycans to improve their chances of survival. To achieve these enhanced affinities, pathogens

Table 1: Crystal structures of mannose-binding lectins, and their affinity for α-D-mannose (**1**) or methyl α-D-mannopyranoside (**2**).

lectin	target	affinity [μM]	ligand efficiency	PDB code	resolution [Å]	reference
A	BDCA-2	9.4×10^3 [a]	0.22	4ZES	1.65	[40]
B	langerin	4.4×10^3 [a]	0.25	4N37	2.00	[41]
C	SP-D	3.8×10^3 [b]	0.28	3G81	1.80	[42]
D	DC-SIGNR	2.5×10^3 [b]	0.30	1K9J[c]	1.90	[43,46]
E	DC-SIGN	2.3×10^3 [b]	0.31	1SL4[c]	1.55	[22,43]
F	rat MBP	1.3×10^3 [b]	0.34	1KWU	1.95	[44,47]
G	LecB	71[a]	0.45	1OUR	1.42	[29,45]
H	BC2L-A	2.8[a]	0.60	2VNV	1.70	[31]
I	$FimH_{LD}$	1.2[a]	0.61	5JCQ	1.60	[48]

[a]Affinity of methyl α-D-mannoside (**2**); [b]affinity of α-D-mannose (**1**); [c]modified oligopyranomannose crystal structure.

apply a variety of strategies such as binding sites with minimal solvent-exposed surface areas, increased number of ligand interactions, "shared" desolvation costs, and multivalency.

Degree of solvent exposure in the binding site (Figure 2). Because of the electrostatic character of H-bonds, the dielectric constant ε becomes especially important in carbohydrate–lectin interactions. In buried cavities of the binding site, ε is lower

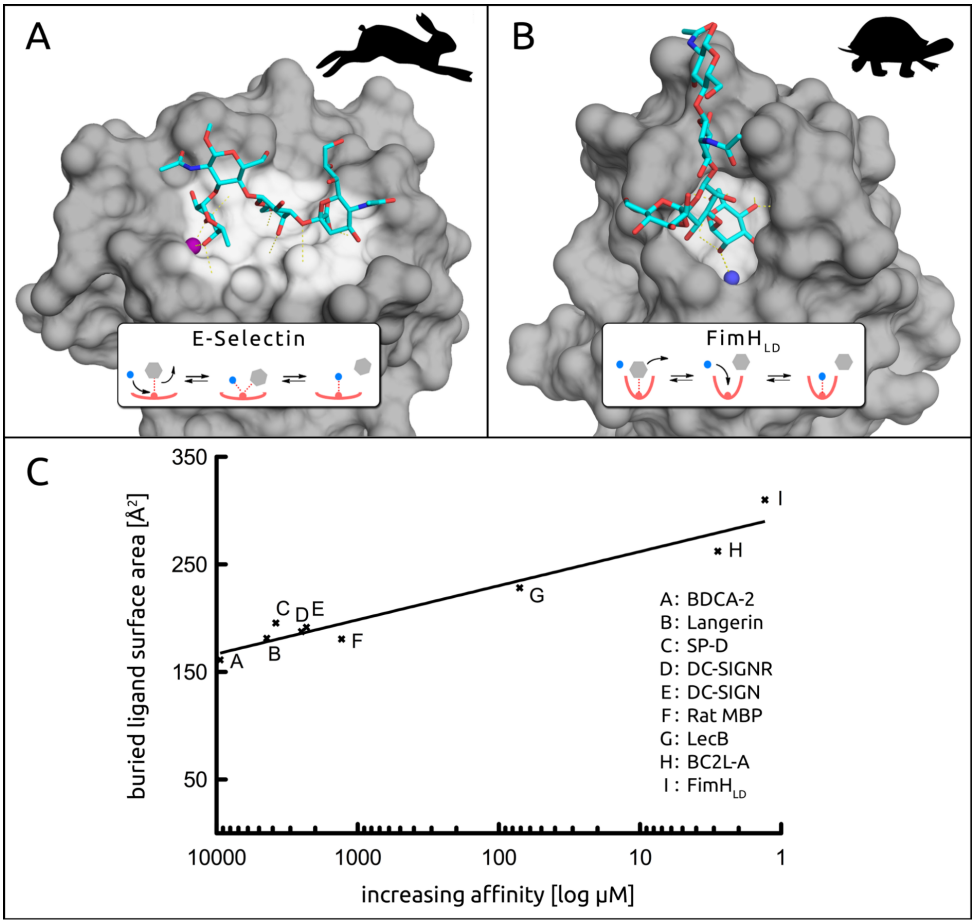


Figure 2: A) The solvent exposed binding site of E-selectin interacting with sLe^x (PDB: 1G1T) [53]. B) The buried binding site of $FimH_{LD}$ in complex with a high-mannose epitope and a conserved water (blue sphere) (PDB: 2VCO) [36]. C) The buried ligand surface area of analyzed crystal structures correlates to affinity [μM] on a logarithmic scale (for references see Table 1). The rabbit and turtle metaphor was adapted from Schmidtke et al. [50].

(ε ≈ 5–10) compared to protein surfaces (ε ≈ 20) or bulk water (ε ≈ 80), making an H-bond thermodynamically up to 10-fold more valuable in buried cavities [49]. This at least partially explains the generally weak interactions of carbohydrates that bind on the solvent exposed surface of proteins, as compared to those of the majority of marketed drugs that most frequently bind to protein cavities. Additionally, buried and less solvent exposed ligands show slower exchange rates, characterized by a high-energy transition state. This can be explained by the stepwise dissociation and subsequent rehydration that are required for ligand displacement (inset, Figure 2B), due to the inherently shielded nature of the buried binding site. In contrast, solvent exposed H-bonds can be more easily substituted by surrounding water molecules in a concerted, bimolecular process (inset, Figure 2A), resulting in faster off-rates and therefore poor pharmacodynamics [50,51]. Similarly, water molecules in buried binding sites show residence times in the micro- to nanosecond range as opposed to surface water molecules which exhibit short residence times in the low picosecond range [52].

Whereas E-selectin in complex with sLe^x is an excellent example of a solvent exposed interaction [12,53], the interaction of $FimH_{LD}$ with mannosides well illustrates the counter situation for a deep CRD [36] (Figure 2A and B, respectively). This difference in solvent exposure leads to considerably different residence times for their physiological ligands. Whereas sLe^x has a residence time of less than a second, the natural substrate of $FimH_{LD}$ (**I**) displays a residence time of more than a minute, and for some synthetic $FimH_{LD}$ antagonists even longer [48,54].

Among the analyzed CLECs **A**–**F** and bacterial lectins **G**–**I**, affinity increases with a decrease in solvent exposure of the binding site (Figure 2C). The buried ligand surface area, an alternative way of expressing solvent exposure of the binding site, is between 160–180 $Å^2$ for **A**–**F**, 228 $Å^2$ for **G**, 262 $Å^2$ for **H**, and 310 $Å^2$ for **I**. The decreased dielectric constant ε in the deep cavities of **H** and **I**, as well as the resulting occlusion of the ligand from surrounding water molecules, leads to a more stable hydrogen-bond network and thus to higher affinities. Furthermore, the binding site of **F** features the aromatic His189, that can engage in CH–π interactions, associated with contributions to the binding affinity in the range of 0–6.3 kJ/mol [55,56].

Analysis of the dynamics of mannose–lectin interactions (Figure 3). In a next step, the stability of H-bond and metal interactions, as well as the influence of highly mobile vs conserved waters were analyzed. For the assessment of the dynamic behavior of the ligand complexes of the seven calcium-dependent lectins, 20 ns molecular dynamics (MD) simulations were performed [57]. The most prominent interactions of O–C3 and O–C4 of the mannose moiety with the calcium ion of CLECs **A**–**F** were stable throughout the entire simulation [5,7]. With the bacterial lectins LecB (**G**) and BC2L-A (**H**) each featuring two calcium ions the carbohydrate ligand forms up to four interactions: O–C2 and O–C4 provide one each, while O–C3 engages with both calcium ions.

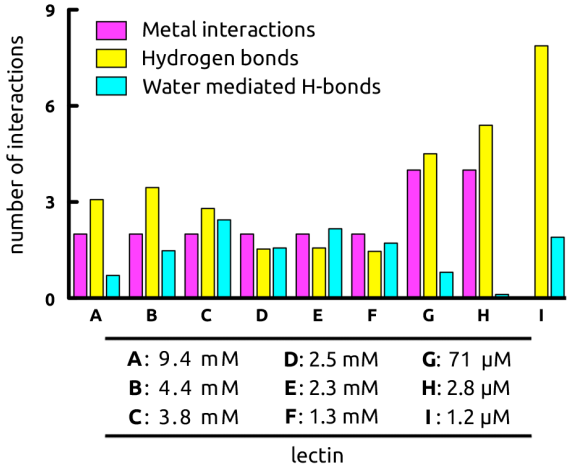


Figure 3: Dynamic mannose–receptor interactions (20 ns MD simulations), grouped according to the nature of the interaction. Metal interactions with Ca^{2+} are colored in purple, hydrogen bonds highlighted in yellow, and water-mediated interactions in cyan. (**A**) BDCA-2, (**B**) langerin, (**C**) SP-D, (**D**) DC-SIGNR, (**E**) DC-SIGN, (**F**) rat MBP, (**G**) LecB, (**H**) BC2L-A, (**I**) $FimH_{LD}$.

During MD simulations, the number of ligand–protein hydrogen-bond interactions for lectins **A**–**F** varied from 1.5 to 3.5, and subsequently increased to 4.5 and 5.4 for LecB (**G**) and BC2L-A (**H**), respectively. Lastly, FimH (**I**) forms on average 7.9 hydrogen bonds with methyl α-D-mannopyranoside (**2**). For H-bonds that were only partially present during the MD simulation, non-integer numbers of hydrogen bonds arise.

The number of water-bridged H bonds between ligand and lectin varied greatly (Figure 3), from 0.1 to 2.4 for the buried binding site of BC2L-A (**H**) versus the solvent exposed binding site of SP-D (**C**), respectively. Interestingly, although the structurally similar bacterial CRDs of **G** and **H** differ by only one amino acid in the β8-β9-loop, a large impact on the number of water-mediated H-bonds was observed. Thus, Thr98 in the bacterial lectin **G** allows for the entry of a water molecule close to the first calcium ion, a process which is hindered by His112 in **H**, leading to a 25-fold difference in affinity. However, in the case of highly mobile water molecules, water-mediated H-bonds as observed in MD simulations destabilize the carbohydrate–lectin interaction, whereas a pre-constrained water molecule does not lead to an additional entropy penalty upon

H-bonding to the ligand. As a result, the interaction benefits from an enthalpic gain without suffering from an entropic penalty [58]. Examples of such highly conserved water molecules are found in both, L-arabinose binding protein (ABP) [59] and FimH (**I**), where in the latter the water mediated H-bond originates solely from one stable water interacting with O–C2 (Figure 2B).

The cost of desolvating hydroxy groups (Figure 4). In general, when the low affinity issue regarding carbohydrate–lectin interactions is discussed, the costs of desolvation are often neglected. Because of the large number of hydroxy groups present in carbohydrate ligands, and the polar amino acid side chains of the lectin binding sites, desolvation generates an essential enthalpic penalty which can hardly be compensated for by the newly formed electrostatic interactions [60]. Cabani et al. calculated that the desolvation of an isolated hydroxy group causes an enthalpic penalty of ΔH = 35 kJ/mol, which is slightly reduced by a beneficial entropic term of ΔS = 10 kJ/mol due to the release of solvating water molecules into bulk [61]. As a result, the desolvation penalty for one hydroxy group amounts to ΔG = 25 kJ/mol (Figure 4A) and cannot be compensated for by a single hydroxy H-bond, which has been associated with a maximal energy gain of approximately ΔG = 18 kJ/mol [62,63]. However, for vicinal hydroxy groups as are present in carbohydrate ligands, the overall desolvation penalty is slightly reduced resulting in an overall desolvation cost of ΔG = 34 kJ/mol for both hydroxy groups (Figure 4B). Since carbohydrates in general exhibit a number of adjacent hydroxy groups, their desolvation penalties are difficult to assess but it is most likely that each additional hydroxy group would not contribute the maximum penalty associated with an isolated one.

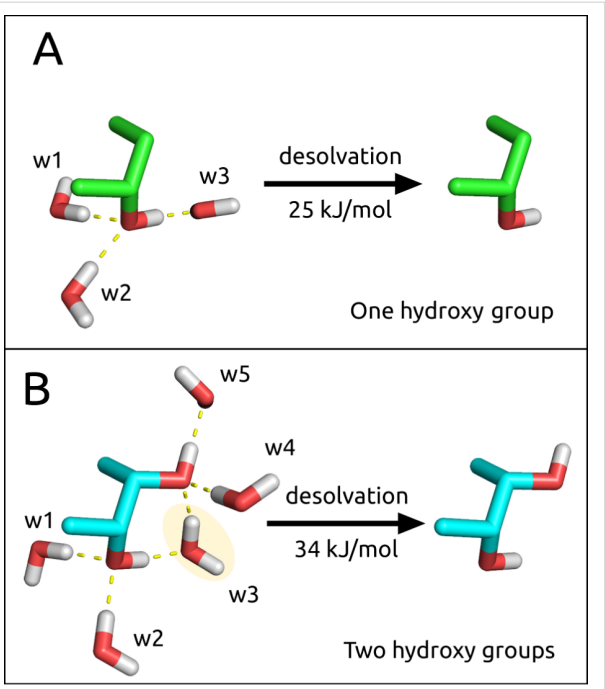


Figure 4: Desolvation of hydroxy groups. A) The desolvation cost of a single hydroxy group associated with three water molecules is ΔG = 25 kJ/mol, as calculated by Cabani et al. [61]. B) the desolvation cost of two adjacent hydroxy groups associated with five water molecules is ΔG = 34 kJ/mol, as opposed to 50 kJ/mol (equal to twice 25 kJ/mol). This can be explained by water w3, which is shared between the two hydroxy groups.

The cost of desolvating calcium ions (Figure 5). Opportunistic bacteria such as *Pseudomonas aeruginosa* or *Burkholderia cenocepacia* have incorporated a second calcium ion into their binding site, coordinating three water molecules which are released into bulk water upon mannose binding and thereby contribute to a favorable entropic effect. The cost to

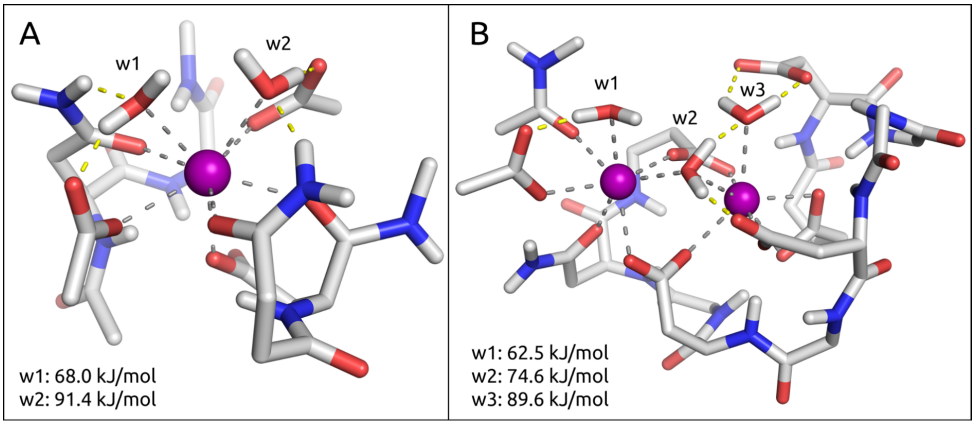


Figure 5: Model view of the binding site interactions of DC-SIGN (**E**) and BC2L-A (**H**) with water. A) The solvent exposed binding site of **E** illustrates two water molecules (w1 and w2) interacting with a single calcium ion, and the cost of desolvation for these waters. B) The buried binding site of **H** exhibits three water molecules (w1–3) bound to two calcium ions, together with their respective desolvation costs.

remove one water molecule from a calcium–malonate model system was calculated quantum mechanically (QM) to be 56.9 kJ/mol by Charifson et al. [64]. This is in agreement with electrospray mass spectrometry experiments from Blades et al., who reported water–calcium interaction energies in the range of 62.8 kJ/mol [65]. In-house QM calculations, based on binding site models of DC-SIGN (**E**) and BC2L-A (**H**) (Figure 5), suggest that the average desolvation cost of a single water molecule coordinated to the calcium ion (calculated as a simple difference of the electronic energies of three molecular species: $E_{desolv} = E_{receptor\cdots water} - E_{receptor} - E_{water}$) is approximately 77 kJ/mol [66]. Interestingly, the calculated desolvation penalty per calcium ion is more favorable for the binding site of **H** (113 kJ/mol per Ca^{2+}), as compared to the one for **E** (159 kJ/mol per Ca^{2+}). Similar to the observations made for vicinal hydroxy groups, the rather high desolvation penalty of two calcium ions in the cases of LecB (**G**) and BC2L-A (**H**) (Figure 5B), is in fact reduced when compared to the sum of desolvating two individual calcium ions, again a result of them sharing a common water molecule.

However, the absolute values of the calculated desolvation energies strongly depend on the local environment of each water molecule. For example, w3 in the binding site of BC2L-A (**H**) exhibits a desolvation energy of 89.6 kJ/mol due to the additional interactions to a glutamate and w2. On the other hand, w1 in the exact same binding site is the least costly among the three waters, as it forms fewer interactions and its loss can also be partially compensated by w2 (Figure 5B).

Profiling the pharmacodynamic difference in binding sites. A comparison of the thermodynamic fingerprints of sLe^x interacting with the solvent exposed CRD of E-selectin versus *n*-heptyl α-D-mannoside bound to the buried binding pocket of $FimH_{LD}$ (**I**) represent two different binding scenarios (Figure 2A and B). With the entropically driven sLe^x interaction, surface waters are displaced to the bulk and penalized by a positive enthalpy term resulting from a desolvation penalty that is not compensated by the newly formed electrostatic interactions (Figure 6) [67]. According to Dunitz [68], the entropy that can be gained by such waters ranges from 0 kJ/mol for highly mobile waters to 8 kJ/mol for ordered and firmly bound waters. In contrast, the thermodynamic fingerprint of FimH ligands is enthalpically driven because an optimized, stable H-bond network is formed, and as a result, overcompensates the desolvation penalty [69,70].

Pre-organization vs flexibility. Carbohydrate–lectin interactions benefit from the low conformational flexibility of pyranoses. This could be impressively demonstrated in a case study comparing a septanose with a *manno*-configured pyranose de-

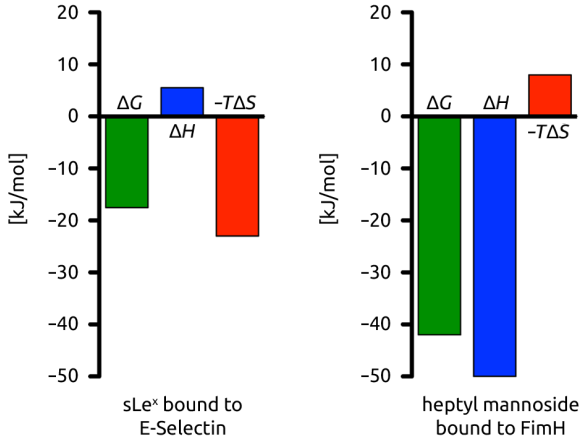


Figure 6: Thermodynamic fingerprints of sLe^x bound to E-selectin and *n*-heptyl α-D-mannoside bound to $FimH_{LD}$ (**I**). The sLe^x–E-selectin interaction is entropically driven, whereas the *n*-heptyl α-D-mannoside–$FimH_{LD}$ (**I**) is enthalpically driven.

rivative [71]. Although in both cases an identical H-bond network with the conformationally rigid $FimH_{LD}$ (**I**) was established, the higher flexibility of the seven-membered ring septanose led to a 10-fold loss in affinity. In fact, the number of possible solution conformations was considerably higher for the septanose ligand as compared to the six-membered ring counterpart, effectively increasing the entropic cost of binding to $FimH_{LD}$ (**I**), while the enthalpic fingerprint observed for both ligands was identical.

However, depending on its needs, UPEC can vary the conformational state of FimH. In the unbound state, FimH exhibits the low-affinity conformation (Figure 7A), which upon binding to mannose, switches to the medium-affinity conformation (Figure 7B). In this state, weak interactions are beneficial because the bacterium can still easily dissociate (slip-bond behavior) and explore its surroundings for optimal nutrient supply. During voiding of the bladder, shear force acts on the FimH protein and pulls the lectin domain ($FimH_{LD}$) away from the pilin domain ($FimH_{PL}$), inducing the high-affinity conformation (Figure 7C), which exhibits an approximate 100-fold higher affinity. Generally, this type of shear force-dependent adhesive bond is known as a catch-bond and in the case of UPECs enables them to evade clearance during micturition. When shear force ceases, FimH reverts back to the equilibrium between low-affinity and medium-affinity conformations [72].

In general, flexible receptors are associated with higher entropic costs resulting from induced-fit binding, which also correlates to facilitated ligand dissociation: due to increased water exposure, the residence time of flexible ligand–lectin com-

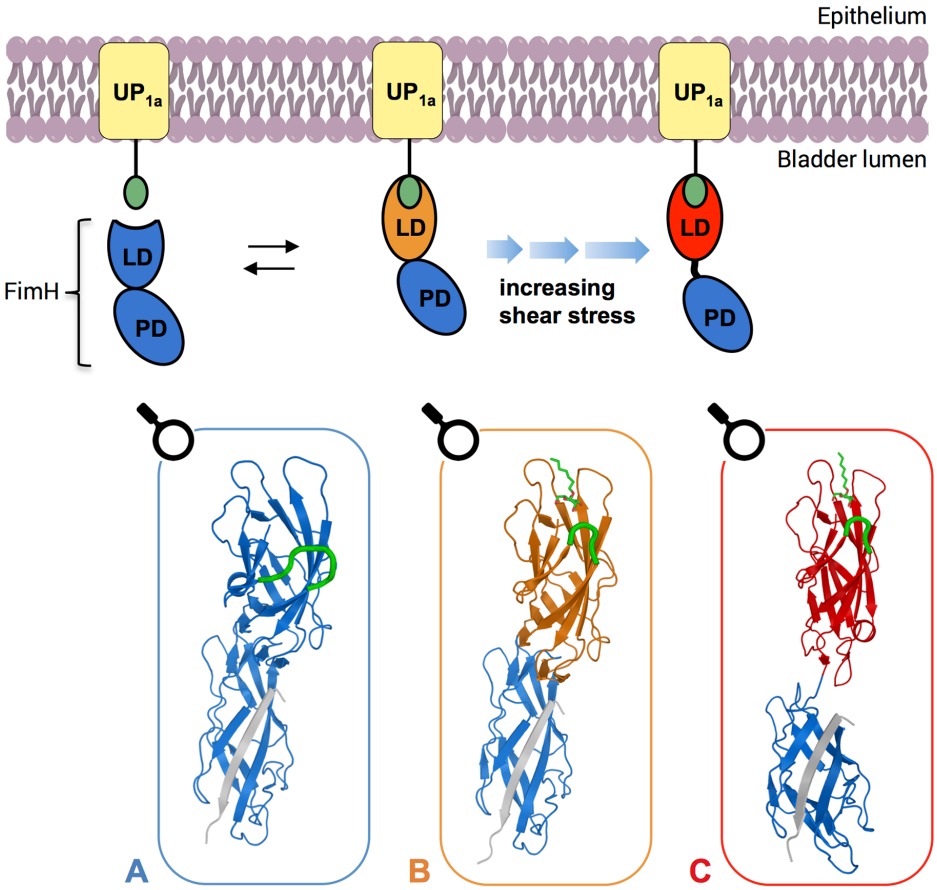


Figure 7: Schematic overview of the conformational changes of FimH (**I**). FimH crystal structures, which correspond to each individual state are shown in boxes. In the absence of urine flow FimH is in the low-affinity conformation, characterized by an open binding pocket and intertwined domains (A, PDB: 4XOD). Upon ligand binding, FimH adopts the medium-affinity conformation (B, PDB: 4XOE), in which a loop (highlighted in green in the crystal structures) closes in the ligand, forming a deep and well-defined binding pocket. As urine begins to flow, the two domains are pulled apart, inducing the high-affinity conformation (C, PDB: 4XOB). LD, lectin domain; PD, pilin domain [72].

plexes is shortened [51]. A comparison of the apo crystal structures of BDCA-2 (**A**) and LecB (**G**) (PDB codes: 3WBP [73] and 1OUX, respectively) to their holo forms excellently demonstrates the entropic costs generated by receptor flexibility. Whereas the binding site of the bacterial lectin **G** does not undergo conformational changes upon ligand binding (RMSD: 0.3 Å; Figure 8A), a conformational change involving a binding site loop allows for the formation of a homodimer of **A** (Figure 8B) [40]. It is believed that this dimer enables transport of the lectin from the Golgi apparatus to cell membranes [73]. Due to a dislocated glutamate in the side chain of the homodimer (Figure 8B), the affinity for calcium binding and therefore also carbohydrate binding is extensively reduced. This remarkable form of inactivation is only possible due to loop flexibility. However, it is also the origin of the low affinity (9.4 mM) towards methyl mannoside (**2**) due to entropic costs associated with the formation of the binding site.

Multivalency. Dam and Brewer reviewed the role of density and number of glycan epitopes involved in multivalent carbohydrate interactions for legume lectins as well as for lectins of the innate immune system [74]. As an example, HIV-1 establishes multivalent contacts to DC-SIGN (**E**)-decorated dendritic cells in order to bypass host immune attack. Thus, DC-SIGN plays a key role in the dissemination of HIV-1 by capturing of HIV-1 at entry sites of infection and subsequent transport of the virus to $CD4^+$ T cells in lymphoid tissues. The weak monovalent binding affinity of DC-SIGN (**E**) is compensated for by a multivalent display of oligomannosides on viral envelop glycoprotein gp120, facilitating stronger adhesion between dendritic cells and HIV-1 [43,75,76]. This multivalent binding interaction results in an enhancement in binding by several orders of magnitude, from a K_D of 26 μM for monovalent $Man_9GlcNAc_2$, as compared to 1.7 nM for glycosylated gp120 (25 glycosylation sites) [43,77]. In the case of UPEC, each bacterium

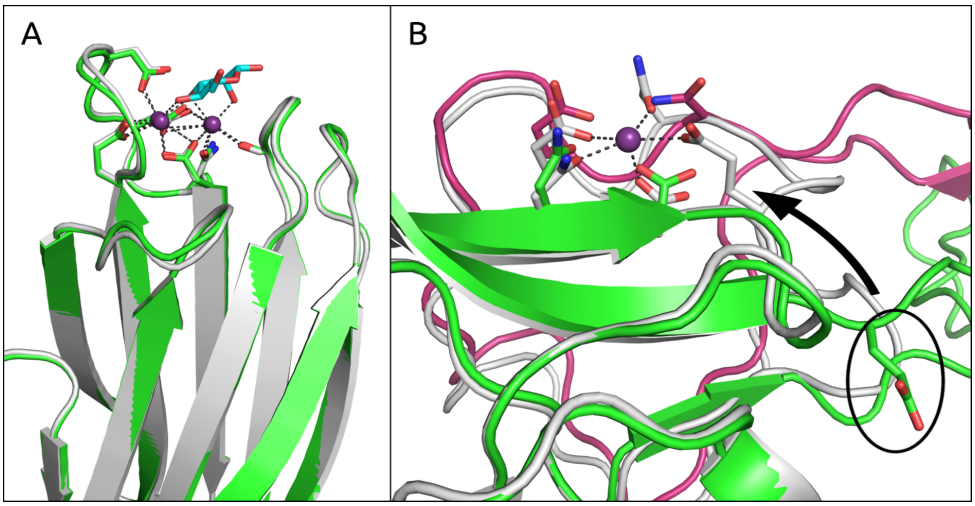


Figure 8: Comparison of the holo (white) and apo (green, magenta) binding sites of LecB (**G**) and BDCA-2 (**A**), respectively. A) The superimposition of both binding sites of **G** reveals nearly identical structures (RMSD of 0.3 Å). B) The binding site of **A** exhibits a flexible loop which enables homodimerization (chain A in green and chain B in magenta), in which a glutamate residue that is essential for Ca^{2+} binding ends up projecting away from the binding site (illustrated for chain A). The situation is mirrored for chain B. Interestingly the loop from chain B closely mimics the binding site of the holo structure (white).

contains three to five hundred fimbriae to potentiate multivalency, as each $FimH_{LD}$ (**I**) at the fimbrial tip can interact with mammalian UPIa [78].

Multivalent glycosides have also been investigated in the context of a novel therapeutic approach against viral and bacterial infections [79]. However, carbohydrate valency, spacing, and branching all need to be thoughtfully considered with this class of therapeutics [15,80].

Conclusion

Mannose-recognizing lectins fulfill a myriad of purposes and depending on the particular biological role either high selectivity and/or high affinity can be required.

On the one hand, lectins of the human immune system tend to exhibit lower affinities due to a higher degree of solvent exposure of their CRDs: fewer H-bond interactions can barely compensate for the high desolvation penalties and constrainment of flexible loop motifs which together contribute to a significant energy penalty upon binding. Nonetheless, these qualities enable ligand promiscuity and can facilitate other features such as the inactivation via homodimerization as exemplified in BDCA-2.

In contrast, bacterial lectins are under constant pressure for survival, hence multiple strategies to 'get it right the first time' are employed. For example, the desolvation of a binding site containing two calcium ions costs 113 kJ/mol/Ca^{2+} and therefore is less costly per calcium ion than a binding site containing only a single ion (159 kJ/mol/Ca^{2+}; Figure 5). However, in the binding site containing two calcium ions, the ions are able to establish four interactions with the carbohydrate ligand, whereas in the latter example the number of interactions is reduced to two. This leads to an overall enthalpic benefit by forming additional interactions at a reduced cost. In addition, the entropy gained by releasing three water molecules into bulk, as compared to only two, should also be taken into account.

The formation of multiple H-bonds in rigid, buried binding sites is an alternative way to gain enthalpy, and thereby increase affinity. UPEC perfects this approach with the calcium-devoid binding site of $FimH_{LD}$ (**I**). A possible explanation for the lack of a calcium ion in the FimH binding site may relate to the slight acidity of urine (pH 5.5–7.0), with a calcium clearance of 20–300 mg/day [81]. Calcium-dependent lectins require a non-acidic environment, such as found in blood, since at lower pH the glutamate and/or aspartate side chains essential for calcium binding can become partially protonated. Instead, $FimH_{LD}$ forms an extensive hydrogen-bond network in a buried, rigid binding site, which lowers the dielectric constant resulting in better shielded, stronger hydrogen bonds, and also reduces the entropic penalty of binding [82]. In addition, the recently described catch-bond behavior of $FimH_{LD}$ is responsible for a 100-fold increased affinity under selective pressure [70,83]. Together with the multivalency of the interaction this results in the high affinity of **2** to $FimH_{LD}$ (**I**).

The examples apparent in nature of effective mannose recognition domains rely on a combination of partially opposed effects.

They nevertheless offer interesting perspectives for the development of carbohydrate-based drugs. One such example of a therapeutic application can be found in a recent novel approach to treating anti-myelin-associated glycoprotein (anti-MAG) neuropathy, a rare, disabling autoimmune disorder. The use of a multivalent glycopolymer mimicking the natural HNK-1 epitope proved to be a valid approach to selectively sequester the autoantibodies associated with anti-MAG neuropathy onset. By applying a multivalent strategy, the inhibitory potential of the monomeric carbohydrate epitope (K_D 124–793 μM from individual patient sera) could be improved by up to a factor of 230,000 in the multivalent display (K_D 3.6–5.4 nM/epitope) [84].

Acknowledgements

This work was supported by the Swiss National Science Foundation (grant number 146202).

ORCID® iDs

Christoph P. Sager - https://orcid.org/0000-0002-1481-9514
Martin Smieško - https://orcid.org/0000-0003-2758-2680
Rachel Hevey - https://orcid.org/0000-0002-2649-3427
Beat Ernst - https://orcid.org/0000-0001-5787-2297

References

1. Weis, W. I.; Taylor, M. E.; Drickamer, K. *Immunol. Rev.* **1998,** *163,* 19–34. doi:10.1111/j.1600-065X.1998.tb01185.x
2. Ernst, B.; Magnani, J. L. *Nat. Rev. Drug Discovery* **2009,** *8,* 661–677. doi:10.1038/nrd2852
3. Takeuchi, O.; Akira, S. *Cell* **2010,** *140,* 805–820. doi:10.1016/j.cell.2010.01.022
4. Drickamer, K. *J. Biol. Chem.* **1988,** *263,* 9557–9560.
5. Zelensky, A. N.; Gready, J. E. *FEBS J.* **2005,** *272,* 6179–6217. doi:10.1111/j.1742-4658.2005.05031.x
6. Drickamer, K. Multiple Subfamilies of Carbohydrate Recognition Domains in Animal Lectins. In *Ciba Foundation Symposium 145 – Carbohydrate Recognition in Cellular Function;* Bock, G.; Harnett, S., Eds.; John Wiley and Sons: Chichester, UK, 1989; pp 45–61. doi:10.1002/9780470513828.ch4
7. Drickamer, K.; Taylor, M. E. *Curr. Opin. Struct. Biol.* **2015,** *34,* 26–34. doi:10.1016/j.sbi.2015.06.003
8. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucleic Acids Res.* **2000,** *28,* 235–242. doi:10.1093/nar/28.1.235
9. Sharon, N.; Lis, H. *Science* **1989,** *246,* 227–234. doi:10.1126/science.2552581
10. Sharon, N. *Biochim. Biophys. Acta, Gen. Subj.* **2006,** *1760,* 527–537. doi:10.1016/j.bbagen.2005.12.008
11. Ley, K. *Trends Mol. Med.* **2003,** *9,* 263–268. doi:10.1016/S1471-4914(03)00071-6
12. Preston, R. C.; Jakob, R. P.; Binder, F. P. C.; Sager, C. P.; Ernst, B.; Maier, T. *J. Mol. Cell Biol. (Oxford, U. K.)* **2016,** *8,* 62–72. doi:10.1093/jmcb/mjv046
13. Pieters, R. J. *Org. Biomol. Chem.* **2009,** *7,* 2013–2025. doi:10.1039/b901828j
14. Ting, S. R. S.; Chen, G.; Stenzel, M. H. *Polym. Chem.* **2010,** *1,* 1392–1412. doi:10.1039/c0py00141d
15. Müller, C.; Despras, G.; Lindhorst, T. K. *Chem. Soc. Rev.* **2016,** *45,* 3275–3302. doi:10.1039/C6CS00165C
16. Mammen, M.; Choi, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998,** *37,* 2754–2794. doi:10.1002/(SICI)1521-3773(19981102)37:20<2754::AID-ANIE2754>3.0.CO;2-3
17. Reynolds, M.; Pérez, S. *C. R. Chim.* **2011,** *14,* 74–95. doi:10.1016/j.crci.2010.05.020
18. Dzionek, A.; Sohma, Y.; Nagafune, J.; Cella, M.; Colonna, M.; Facchetti, F.; Günther, G.; Johnston, I.; Lanzavecchia, A.; Nagasaka, T.; Okada, T.; Vermi, W.; Winkels, G.; Yamamoto, T.; Zysk, M.; Yamaguchi, Y.; Schmitz, J. *J. Exp. Med.* **2001,** *194,* 1823–1834. doi:10.1084/jem.194.12.1823
19. van der Vlist, M.; Geijtenbeek, T. B. H. *Immunol. Cell Biol.* **2010,** *88,* 410–415. doi:10.1038/icb.2010.32
20. Valladeau, J.; Ravel, O.; Dezutter-Dambuyant, C.; Moore, K.; Kleijmeer, M.; Liu, Y.; Duvert-Frances, V.; Vincent, C.; Schmitt, D.; Davoust, J.; Caux, C.; Lebecque, S.; Saeland, S. *Immunity* **2000,** *12,* 71–81. doi:10.1016/S1074-7613(00)80160-0
21. Wright, J. R. *Nat. Rev. Immunol.* **2005,** *5,* 58–68. doi:10.1038/nri1528
22. Guo, Y.; Feinberg, H.; Conroy, E.; Mitchell, D. A.; Alvarez, R.; Blixt, O.; Taylor, M. E.; Weis, W. I.; Drickamer, K. *Nat. Struct. Mol. Biol.* **2004,** *11,* 591–598. doi:10.1038/nsmb784
23. Mitchell, D. A.; Fadden, A. J.; Drickamer, K. *J. Biol. Chem.* **2001,** *276,* 28939–28945. doi:10.1074/jbc.M104565200
24. Mori, K.; Kawasaki, T.; Yamashina, I. *Arch. Biochem. Biophys.* **1988,** *264,* 647–656. doi:10.1016/0003-9861(88)90331-1
25. Cambi, A.; Koopman, M.; Figdor, C. G. *Cell. Microbiol.* **2005,** *7,* 481–488. doi:10.1111/j.1462-5822.2005.00506.x
26. Schlesinger, L. S.; Torrelles, J. B.; Azad, A. K.; Henning, L. N.; Carlson, T. K. *Curr. Drug Targets* **2008,** *9,* 102–112. doi:10.2174/138945008783502467
27. van den Berg, L. M.; Gringhuis, S. I.; Geijtenbeek, T. B. H. *Ann. N. Y. Acad. Sci.* **2012,** *1253,* 149–158. doi:10.1111/j.1749-6632.2011.06392.x
28. Hall-Stoodley, L.; Costerton, J. W.; Stoodley, P. *Nat. Rev. Microbiol.* **2004,** *2,* 95–108. doi:10.1038/nrmicro821
29. Loris, R.; Tielker, D.; Jaeger, K.-E.; Wyns, L. *J. Mol. Biol.* **2003,** *331,* 861–870. doi:10.1016/S0022-2836(03)00754-X
30. Tielker, D.; Hacker, S.; Loris, R.; Strathmann, M.; Wingender, J.; Wilhelm, S.; Rosenau, F.; Jaeger, K.-E. *Microbiology (London, U. K.)* **2005,** *151,* 1313–1323. doi:10.1099/mic.0.27701-0
31. Lameignere, E.; Malinovská, L.; Sláviková, M.; Duchaud, E.; Mitchell, E. P.; Varrot, A.; Šedo, O.; Imberty, A.; Wimmerová, M. *Biochem. J.* **2008,** *411,* 307–318. doi:10.1042/BJ20071276
32. Marchetti, R.; Malinovska, L.; Lameignère, E.; Adamova, L.; de Castro, C.; Cioci, G.; Stanetty, C.; Kosma, P.; Molinaro, A.; Wimmerova, M.; Imberty, A.; Silipo, A. *Glycobiology* **2012,** *22,* 1387–1398. doi:10.1093/glycob/cws105
33. Mulvey, M. A.; Lopez-Boado, Y. S.; Wilson, C. L.; Roth, R.; Parks, W. C.; Heuser, J.; Hultgren, S. J. *Science* **1998,** *282,* 1494–1497. doi:10.1126/science.282.5393.1494
34. Krogfelt, K. A.; Bergmans, H.; Klemm, P. *Infect. Immun.* **1990,** *58,* 1995–1998.
35. Zhou, G.; Mo, W.-J.; Sebbel, P.; Min, G.; Neubert, T. A.; Glockshuber, R.; Wu, X.-R.; Sun, T.-T.; Kong, X.-P. *J. Cell Sci.* **2001,** *114,* 4095–4103.

36. Wellens, A.; Garofalo, C.; Nguyen, H.; Van Gerven, N.; Slättegård, R.; Hernalsteens, J.-P.; Wyns, L.; Oscarson, S.; De Greve, H.; Hultgren, S.; Bouckaert, J. *PLoS One* **2008,** *3,* e2040. doi:10.1371/journal.pone.0002040
37. Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhrer, M.; Hung, C.-S.; Pinkner, J.; Slättegård, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. *Mol. Microbiol.* **2005,** *55,* 441–455. doi:10.1111/j.1365-2958.2004.04415.x
38. Zihlmann, P.; Jiang, X.; Sager, C. P.; Fiege, B.; Jakob, R. P.; Siegrist, S.; Zalewski, A.; Rabbani, S.; Eriş, D.; Silbermann, M.; Pang, L.; Mühlethaler, T.; Sharpe, T.; Maier, T.; Ernst, B. *unpublished results*.
39. Old, D. C. *J. Gen. Microbiol.* **1972,** *71,* 149–157. doi:10.1099/00221287-71-1-149
40. Jégouzo, S. A. F.; Feinberg, H.; Dungarwalla, T.; Drickamer, K.; Weis, W. I.; Taylor, M. E. *J. Biol. Chem.* **2015,** *290,* 16759–16771. doi:10.1074/jbc.M115.660613
41. Feinberg, H.; Rowntree, T. J. W.; Tan, S. L. W.; Drickamer, K.; Weis, W. I.; Taylor, M. E. *J. Biol. Chem.* **2013,** *288,* 36762–36771. doi:10.1074/jbc.M113.528000
42. Crouch, E.; Hartshorn, K.; Horlacher, T.; McDonald, B.; Smith, K.; Cafarella, T.; Seaton, B.; Seeberger, P. H.; Head, J. *Biochemistry* **2009,** *48,* 3335–3345. doi:10.1021/bi8022703
43. Feinberg, H.; Castelli, R.; Drickamer, K.; Seeberger, P. H.; Weis, W. I. *J. Biol. Chem.* **2007,** *282,* 4202–4209. doi:10.1074/jbc.M609689200
44. Lee, R. T.; Ichikawa, Y.; Fay, M.; Drickamer, K.; Shao, M. C.; Lee, Y. C. *J. Biol. Chem.* **1991,** *266,* 4810–4815.
45. Sabin, C.; Mitchell, E. P.; Pokorná, M.; Gautier, C.; Utille, J.-P.; Wimmerová, M.; Imberty, A. *FEBS Lett.* **2006,** *580,* 982–987. doi:10.1016/j.febslet.2006.01.030
46. Feinberg, H.; Mitchell, D. A.; Drickamer, K.; Weis, W. I. *Science* **2001,** *294,* 2163–2166. doi:10.1126/science.1066371
47. Ng, K. K.-S.; Kolatkar, A. R.; Park-Snyder, S.; Feinberg, H.; Clark, D. A.; Drickamer, K.; Weis, W. I. *J. Biol. Chem.* **2002,** *277,* 16088–16095. doi:10.1074/jbc.M200493200
48. Zihlmann, P.; Silbermann, M.; Sharpe, T.; Jiang, X.; Mühlethaler, T.; Jakob, R. P.; Rabbani, S.; Sager, C. P.; Frei, P.; Maier, T.; Ernst, B. *submitted.*
49. Fitch, C. A.; Karp, D. A.; Lee, K. K.; Stites, W. E.; Lattman, E. E.; García-Moreno, E. B. *Biophys. J.* **2002,** *82,* 3289–3304. doi:10.1016/S0006-3495(02)75670-1
50. Schmidtke, P.; Luque, F. J.; Murray, J. B.; Barril, X. *J. Am. Chem. Soc.* **2011,** *133,* 18903–18910. doi:10.1021/ja207494u
51. Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E. *Drug Discovery Today* **2013,** *18,* 667–673. doi:10.1016/j.drudis.2013.02.007
52. Levitt, M.; Park, B. H. *Structure* **1993,** *1,* 223–226. doi:10.1016/0969-2126(93)90011-5
53. Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000,** *103,* 467–479. doi:10.1016/S0092-8674(00)00138-0
54. Scharenberg, M.; Jiang, X.; Pang, L.; Navarra, G.; Rabbani, S.; Binder, F.; Schwardt, O.; Ernst, B. *ChemMedChem* **2014,** *9,* 78–83. doi:10.1002/cmdc.201300349
55. Jiménez-Moreno, E.; Jiménez-Osés, G.; Gómez, A. M.; Santana, A. G.; Corzana, F.; Bastida, A.; Jiménez-Barbero, J.; Asensio, J. L. *Chem. Sci.* **2015,** *6,* 6076–6085. doi:10.1039/C5SC02108A
56. Laughrey, Z. R.; Kiehna, S. E.; Riemen, A. J.; Waters, M. L. *J. Am. Chem. Soc.* **2008,** *130,* 14625–14633. doi:10.1021/ja803960x
57. *Schrödinger,* Release 2015-4; Desmond Molecular Dynamics System, D. E. Shaw Research: New York, 2015.
58. Ladbury, J. E. *Chem. Biol.* **1996,** *3,* 973–980. doi:10.1016/S1074-5521(96)90164-7
59. Quiocho, F. A.; Wilson, D. K.; Vyas, N. K. *Nature* **1989,** *340,* 404–407. doi:10.1038/340404a0
60. Bryce, R. A.; Hillier, I. H.; Naismith, J. H. *Biophys. J.* **2001,** *81,* 1373–1388. doi:10.1016/S0006-3495(01)75793-1
61. Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. *J. Solution Chem.* **1981,** *10,* 563–595. doi:10.1007/BF00646936
62. Jeffrey, G. A. *An Introduction to Hydrogen Bonding;* Oxford University Press: Canada, Oxford, 1997.
63. Steiner, T. *Angew. Chem., Int. Ed.* **2002,** *41,* 48–76. doi:10.1002/1521-3773(20020104)41:1<48::AID-ANIE48>3.0.CO;2-U
64. Charifson, P. S.; Hiskey, R. G.; Pedersen, L. G.; Kuyper, L. F. *J. Comput. Chem.* **1991,** *12,* 899–908. doi:10.1002/jcc.540120717
65. Blades, A. T.; Jayaweera, P.; Ikonomou, M. G.; Kebarle, P. *J. Chem. Phys.* **1990,** *92,* 5900–5906. doi:10.1063/1.458360
66. *Gaussian 09,* Revision D.01; Gaussian Inc.: Wallingford, CT, 2009.
67. Binder, F. P. C.; Lemme, K.; Preston, R. C.; Ernst, B. *Angew. Chem., Int. Ed.* **2012,** *51,* 7327–7331. doi:10.1002/anie.201202555
68. Dunitz, J. D. *Science* **1994,** *264,* 670. doi:10.1126/science.264.5159.670
69. Kleeb, S.; Pang, L.; Mayer, K.; Eris, D.; Sigl, A.; Preston, R. C.; Zihlmann, P.; Sharpe, T.; Jakob, R. P.; Abgottspon, D.; Hutter, A. S.; Scharenberg, M.; Jiang, X.; Navarra, G.; Rabbani, S.; Smiesko, M.; Lüdin, N.; Bezençon, J.; Schwardt, O.; Maier, T.; Ernst, B. *J. Med. Chem.* **2015,** *58,* 2221–2239. doi:10.1021/jm501524q
70. Mayer, K.; Eris, D.; Schwardt, O.; Sager, C. P.; Rabbani, S.; Kleeb, S.; Ernst, B. *J. Med. Chem.* **2017,** *60,* 5646–5662. doi:10.1021/acs.jmedchem.7b00342
71. Sager, C. P.; Fiege, B.; Zihlmann, P.; Vannam, R.; Rabbani, S.; Jakob, R. P.; Preston, R. C.; Zalewski, A.; Maier, T.; Peczuh, M. W.; Ernst, B. *Chem. Sci.* **2018,** in press. doi:10.1039/c7sc04289b
72. Eris, D.; Preston, R. C.; Scharenberg, M.; Hulliger, F.; Abgottspon, D.; Pang, L.; Jiang, X.; Schwardt, O.; Ernst, B. *ChemBioChem* **2016,** *17,* 1012–1020. doi:10.1002/cbic.201600066
73. Nagae, M.; Ikeda, A.; Kitago, Y.; Matsumoto, N.; Yamamoto, K.; Yamaguchi, Y. *Proteins: Struct., Funct., Bioinf.* **2014,** *82,* 1512–1518. doi:10.1002/prot.24504
74. Dam, T. K.; Brewer, C. F. *Glycobiology* **2010,** *20,* 270–279. doi:10.1093/glycob/cwp186
75. van Kooyk, Y.; Geijtenbeek, T. B. H. *Nat. Rev. Immunol.* **2003,** *3,* 697–709. doi:10.1038/nri1182
76. Horiya, S.; MacPherson, I. S.; Krauss, I. J. *Nat. Chem. Biol.* **2014,** *10,* 990–999. doi:10.1038/nchembio.1685
77. Curtis, B. M.; Scharnowske, S.; Watson, A. J. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** *89,* 8356–8360. doi:10.1073/pnas.89.17.8356
78. Hartmann, M.; Lindhorst, T. K. *Eur. J. Org. Chem.* **2011,** 3583–3609. doi:10.1002/ejoc.201100407
79. Bernardi, A.; Jiménez-Barbero, J.; Casnati, A.; Castro, C. D.; Darbre, T.; Fieschi, F.; Finne, J.; Funken, H.; Jaeger, K.-E.; Lahmann, M.; Lindhorst, T. K.; Marradi, M.; Messner, P.; Molinaro, A.; Murphy, P. V.; Nativi, C.; Oscarson, S.; Penadés, S.; Peri, F.; Pieters, R. J.; Renaudet, O.; Reymond, J.-L.; Richichi, B.; Rojo, J.; Sansone, F.; Schäffer, C.; Turnbull, W. B.; Velasco-Torrijos, T.; Vidal, S.; Vincent, S.; Wennekes, T.; Zuilhof, H.; Imberty, A. *Chem. Soc. Rev.* **2013,** *42,* 4709–4727. doi:10.1039/C2CS35408J

80. Lusvarghi, S.; Ghirlando, R.; Wong, C.-H.; Bewley, C. A. *Angew. Chem., Int. Ed.* **2015,** *54,* 5603–5608. doi:10.1002/anie.201500157
81. Rose, C.; Parker, A.; Jefferson, B.; Cartmell, E. *Crit. Rev. Environ. Sci. Technol.* **2015,** *45,* 1827–1879. doi:10.1080/10643389.2014.1000761
82. Quiocho, F. A. *Annu. Rev. Biochem.* **1986,** *55,* 287–315. doi:10.1146/annurev.bi.55.070186.001443
83. Sauer, M. M.; Jakob, R. P.; Eras, J.; Baday, S.; Eriş, D.; Navarra, G.; Bernèche, S.; Ernst, B.; Maier, T.; Glockshuber, R. *Nat. Commun.* **2016,** *7,* No. 10738. doi:10.1038/ncomms10738
84. Herrendorff, R.; Hänggi, P.; Pfister, H.; Yang, F.; Demeestere, D.; Hunziker, F.; Frey, S.; Schaeren-Wiemers, N.; Steck, A. J.; Ernst, B. *Proc. Natl. Acad. Sci. U. S. A.* **2017,** *114,* E3689–E3698. doi:10.1073/pnas.1619386114

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Recent applications of click chemistry for the functionalization of gold nanoparticles and their conversion to glyco-gold nanoparticles

Vivek Poonthiyil[*1], Thisbe K. Lindhorst[1], Vladimir B. Golovko[2,3] and Antony J. Fairbanks[*2,4]

Review

Open Access

Address:
[1]Otto Diels Institute of Organic Chemistry, Christiana Albertina University of Kiel, Otto-Hahn-Platz 3/4, Kiel, 24098, Germany, [2]Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch, 8140, New Zealand, [3]The MacDiarmid Institute for Advanced Materials and Nanotechnology, Wellington, 6140, New Zealand and [4]Biomolecular Interaction Centre, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Email:
Vivek Poonthiyil[*] - vpoonthiyil@oc.uni-kiel.de; Antony J. Fairbanks[*] - antony.fairbanks@canterbury.ac.nz

* Corresponding author

Keywords:
azide–alkyne Huisgen cycloaddition; carbohydrates; click chemistry; glyco-gold nanoparticles; triazole

Beilstein J. Org. Chem. **2018,** *14,* 11–24.
doi:10.3762/bjoc.14.2

Received: 23 September 2017
Accepted: 20 December 2017
Published: 03 January 2018

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Glycoscience, despite its myriad of challenges, promises to unravel the causes of, potential new detection methods for, and novel therapeutic strategies against, many disease states. In the last two decades, glyco-gold nanoparticles have emerged as one of several potential new tools for glycoscientists. Glyco-gold nanoparticles consist of the unique structural combination of a gold nanoparticle core and an outer-shell comprising multivalent presentation of carbohydrates. The combination of the distinctive physicochemical properties of the gold core and the biological function/activity of the carbohydrates makes glyco-gold nanoparticles a valuable tool in glycoscience. In this review we present recent advances made in the use of one type of click chemistry, namely the azide–alkyne Huisgen cycloaddition, for the functionalization of gold nanoparticles and their conversion to glyco-gold nanoparticles.

Introduction

Metal nanoparticles (NPs), with their unique physicochemical properties, have drawn significant interest in recent years, and are expected to form the basis of many biological and technological innovations during the remainder of the 21st century [1]. Gold nanoparticles (AuNPs) are one of the most significant and stable classes of metal NPs [2] and have potential applications in optics [3], biology [4] and catalysis [5].

Carbohydrates are one of the classes of molecules that are essential for life. Although they are involved in many important biological processes, it is now well established that the binding interactions of a particular oligosaccharide, either with another carbohydrate or more commonly with carbohydrate-binding proteins (lectins), are generally weak. In order to augment these low affinity interactions, oligosaccharides usually bind lectins in a multivalent cooperative fashion. This avidity is significantly greater than the sum of the individual monomeric carbohydrate–protein interactions, and is sometimes referred to as the 'cluster glycoside' effect [6]. In order to study biological processes that involve these types of carbohydrate–protein interactions, it is therefore essential to present carbohydrates in a multivalent fashion. For that purpose, different scaffolds, such as peptides, proteins, lipids, and synthetic polymers, have all been used [7].

The search for better scaffolds for the presentation of multivalent carbohydrate structures led to the development of self-assembled monolayers (SAMs) of carbohydrates on the spherical surface of AuNPs. In 2001, the Penadés group reported the first synthesis of AuNPs with attached carbohydrates [8]. These systems, termed 'glyco-gold nanoparticles' (GAuNPs), were comprised of AuNPs with the surface Au atoms covalently attached to thiols of thiol-terminated oligosaccharides [8]. It was found that GAuNPs could be used as mimics of the glycocalyx to study both carbohydrate–carbohydrate and carbohydrate–protein interactions [9,10]. Other applications of GAuNPs, as sensors for various biomolecules and toxins, including the detection of pathogenic agents such as viruses and bacteria, have also been reported by various groups [11-16].

Since the first report by Penadés [8], numerous methods have been developed for the synthesis of GAuNPs. However, recent use of click chemistry for the functionalization of AuNPs and their conversion to GAuNPs has increased significantly. This short review, after giving a brief introduction to general methods for GAuNP synthesis, will focus on both potential advantages and issues of using click chemistry for the functionalization of AuNPs and their conversion to GAuNPs.

Review

Methods for the synthesis of GAuNPs

In general, there are three main methods that can be used to synthesize GAuNPs (Figure 1). The first one is a direct method, involving the reduction of $HAuCl_4$ in the presence of carbohydrate derivatives with a thiol end group, which is generally attached to the reducing terminus by a linker (Figure 1a) [8,14,17-27].

The second method is a ligand exchange reaction involving the replacement of the ligands on pre-formed AuNPs with thiol-linked carbohydrate derivatives (Figure 1b). The most frequently employed approach here is to first synthesize citrate-stabilized AuNPs (Cit-AuNPs) [28], and then to replace the citrate ligands with the desired thiol-linked carbohydrate derivatives [29,30]. Ligand exchange on the AuNP surface is driven by the higher binding affinity of Au for the thiol than for citrate, due to the significant energy difference between Au–S ($\approx$40 kcal·mol^{-1}) and Au–O_{COOH} ($\approx$2 kcal·mol^{-1}) interactions [31].

The third method involves the chemical reaction of functional groups of ligands attached to the surface of pre-formed AuNPs with suitably functionalized carbohydrates (Figure 1c). Various types of reaction, such as reductive amination [32], oxime formation [33], amidation [34], and perfluorophenyl azide (PFPA) photocoupling [35,36], have been used to functionalize the surface of AuNPs with carbohydrates. The detailed information regarding the synthesis and application of GAuNPs can be found in the reviews by Penadés and co-workers [9,26] and also in a recent review by Compostella et al. [10]. In this regard, azide–alkyne click chemistry is an attractive approach that could be used to synthesize GAuNPs.

The functionalization of AuNPs using the azide–alkyne Huisgen cycloaddition

AuNP surface modification using NCAAC

The azide–alkyne Huisgen cycloaddition (AAC) is a 1,3-dipolar cycloaddition between an organic azide and an alkyne that gives triazole products [37,38]. The non-catalysed azide–alkyne Huisgen cycloaddition (NCAAC) is very slow, and gives a mixture of 1,4- and 1,5-triazole regioisomers (Scheme 1) [39]. Interest in and applications of the AAC have surged over the past 15 or so years, since the introduction of Cu(I) catalysis, which led to significant improvements in both the regioselectivity and rates of the reaction [40,41]. The versatility of the Cu(I)-catalysed azide–alkyne Huisgen cycloaddition (CuAAC) has been demonstrated by its robustness, insensitivity to water and oxygen, and its applicability to a wide range of substrates [42-44]. Although the AAC has been used by many groups to modify the surface of AuNPs [45-48], until recently it has only rarely been used to synthesize GAuNPs.

In 2006, Fleming et al. used the NCAAC to attach a series of different species to AuNPs [45]. Small AuNPs (1.8 nm) were used as the substrates for the NCAAC because of their ease of synthesis, high solubility, and good ligand exchange properties. A two-phase Brust–Schiffrin method (BSM) [49] was first used to synthesize decanethiol-stabilized AuNPs. These particles were then reacted with 11-bromo-1-undecanethiol to replace

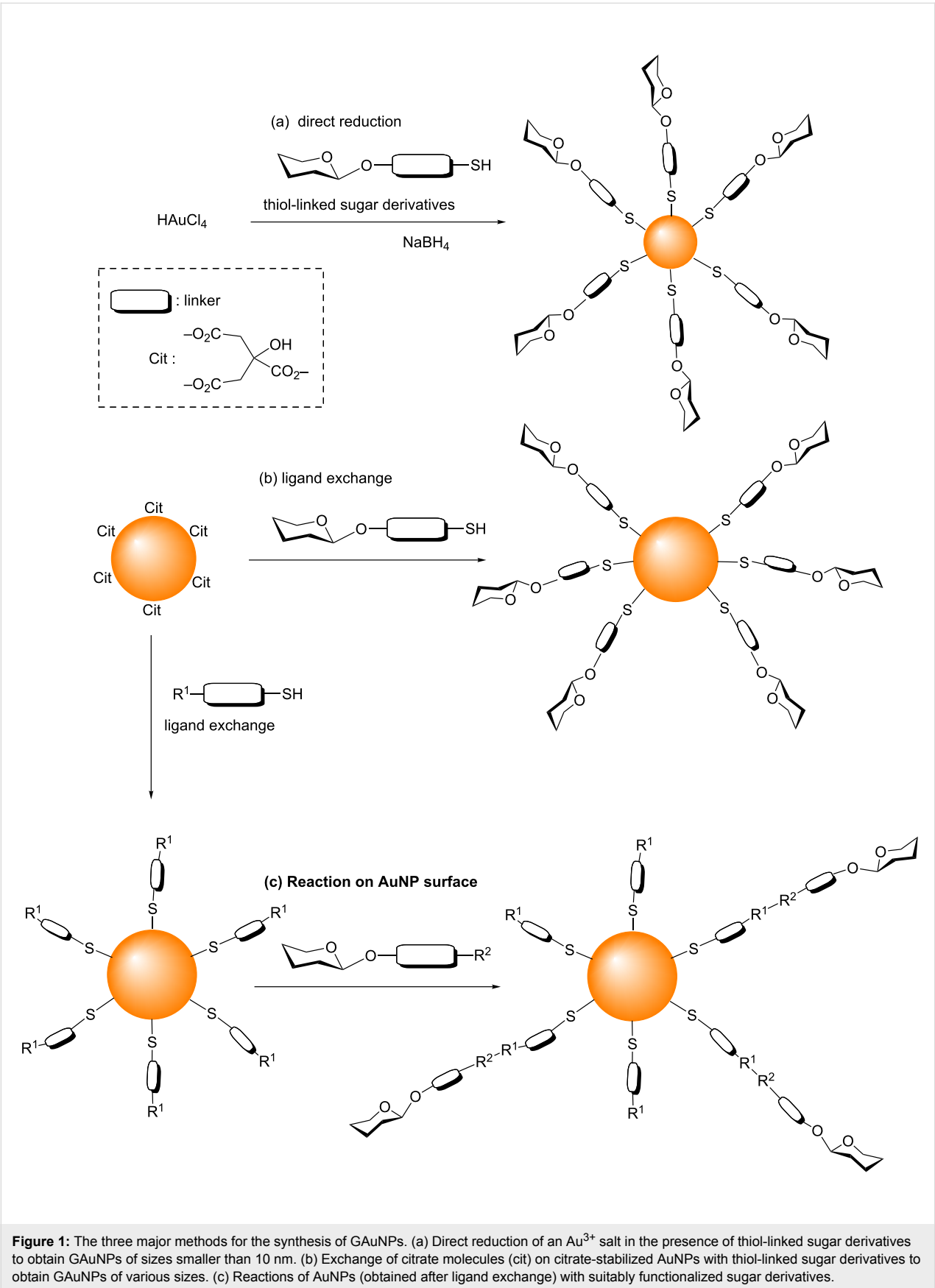
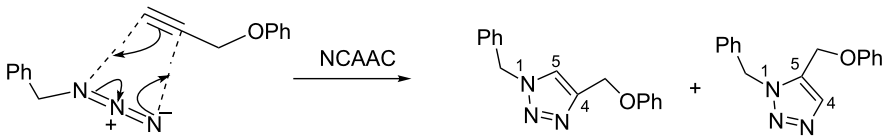


Figure 1: The three major methods for the synthesis of GAuNPs. (a) Direct reduction of an Au^{3+} salt in the presence of thiol-linked sugar derivatives to obtain GAuNPs of sizes smaller than 10 nm. (b) Exchange of citrate molecules (cit) on citrate-stabilized AuNPs with thiol-linked sugar derivatives to obtain GAuNPs of various sizes. (c) Reactions of AuNPs (obtained after ligand exchange) with suitably functionalized sugar derivatives.



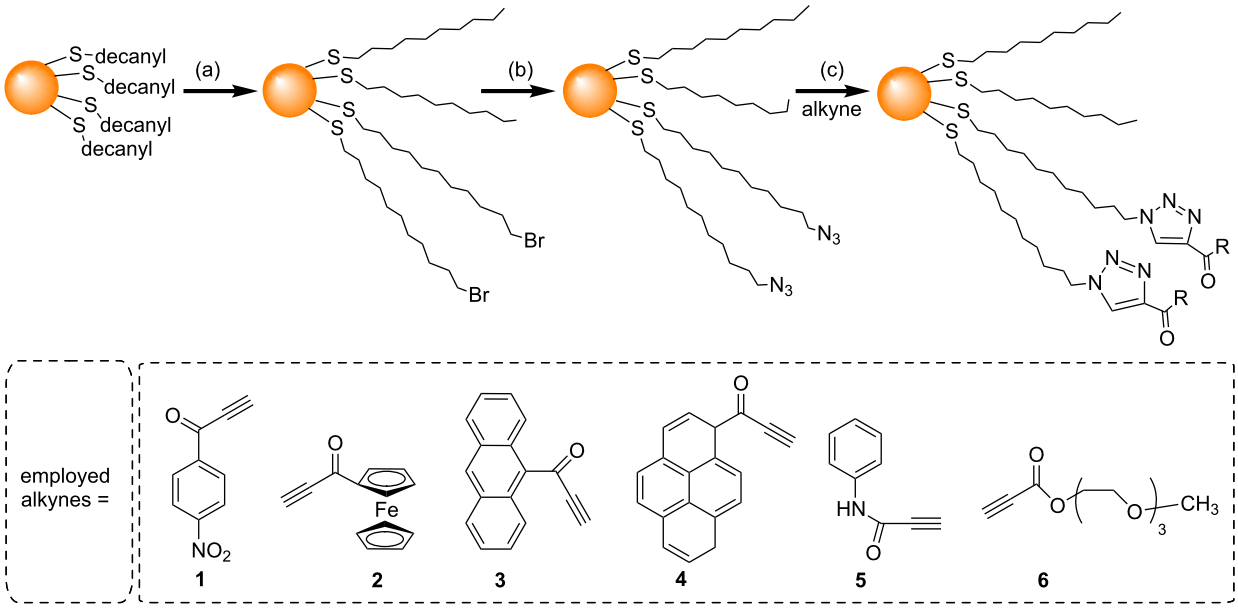
Scheme 1: The non-catalysed azide–alkyne Huisgen cycloaddition (NCAAC) between an organic azide (1,3-dipole) and an alkyne (dipolarophile) resulting in the formation of regioisomeric triazole products.

some of the decanethiol ligands with Br-terminated undecanethiol ligands (Scheme 2). Nucleophilic substitution by reaction with NaN_3 then resulted in AuNPs with mixed monolayers containing 52% N_3- and 44% CH_3-terminated alkanethiol ligands. A series of alkynes were synthesised, including derivatives of nitrobenzene (**1**), ferrocene (**2**), anthracene (**3**), pyrene (**4**), aniline (**5**), and polyethylene glycol (**6**) all of which contained a carbonyl group next to the alkyne to increase the rate of triazole formation [50]. NCAAC between the azide-decorated AuNPs and the alkyne derivatives (**1**–**6**) was then performed (Scheme 2). Although a small amount of the AuNPs underwent irreversible aggregation, the majority of the AuNPs (>90%) remained soluble, and could be separated from aggregates after the reaction. Although Fleming et al. successfully performed NCAAC on these AuNPs, the yields (i.e., the extent of the azide conversion to triazole) were low (22%, or 54% in one specific case) even after 60 hours [45,51].

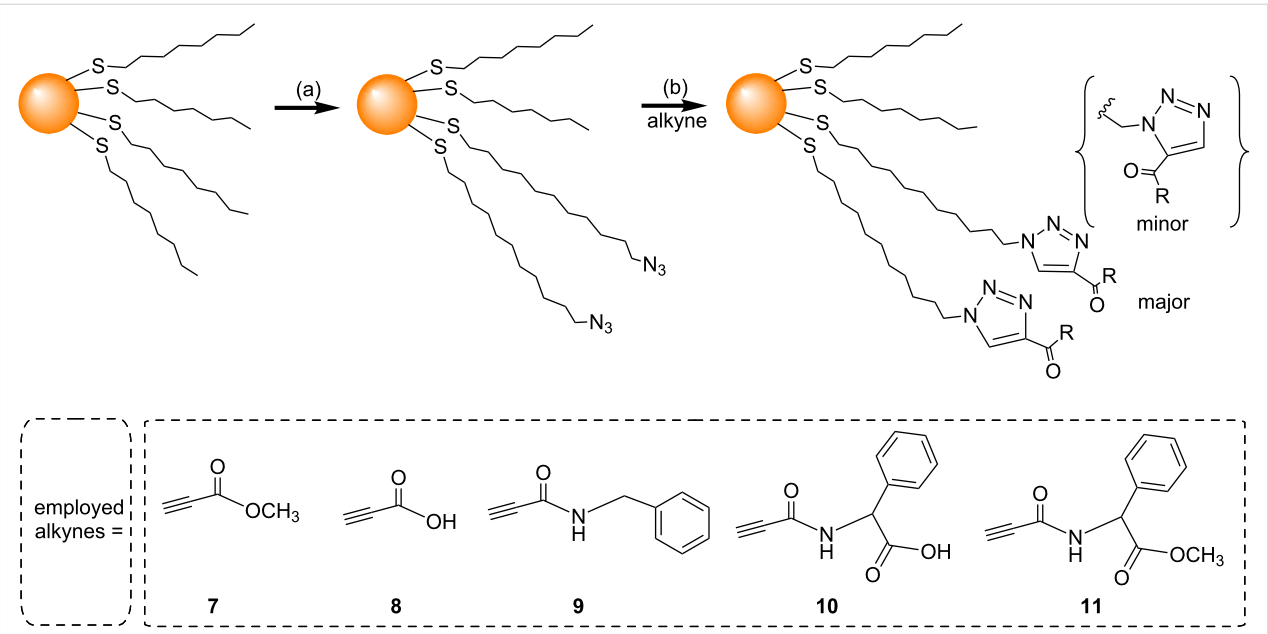
Following the work of Fleming et al., several groups have investigated the use of different conditions to try and increase the efficiency of the NCAAC on the surface of AuNPs. Limapichat et al. used other electron deficient alkynes (**7**–**11**) as substrates for the NCAAC, and observed that 75% of the azides on the AuNP surface underwent cycloaddition in 16 hours (Scheme 3) [52]. Ismaili et al. carried out the NCAAC with a number of terminal-acyl alkynes (**1**–**5** and **12**–**17**) under hyperbaric conditions (11000 atm pressure), and observed 80% or higher conversions within 15 to 24 hours (Scheme 4) [48].

AuNP surface modification using strain-promoted azide–alkyne cycloaddition

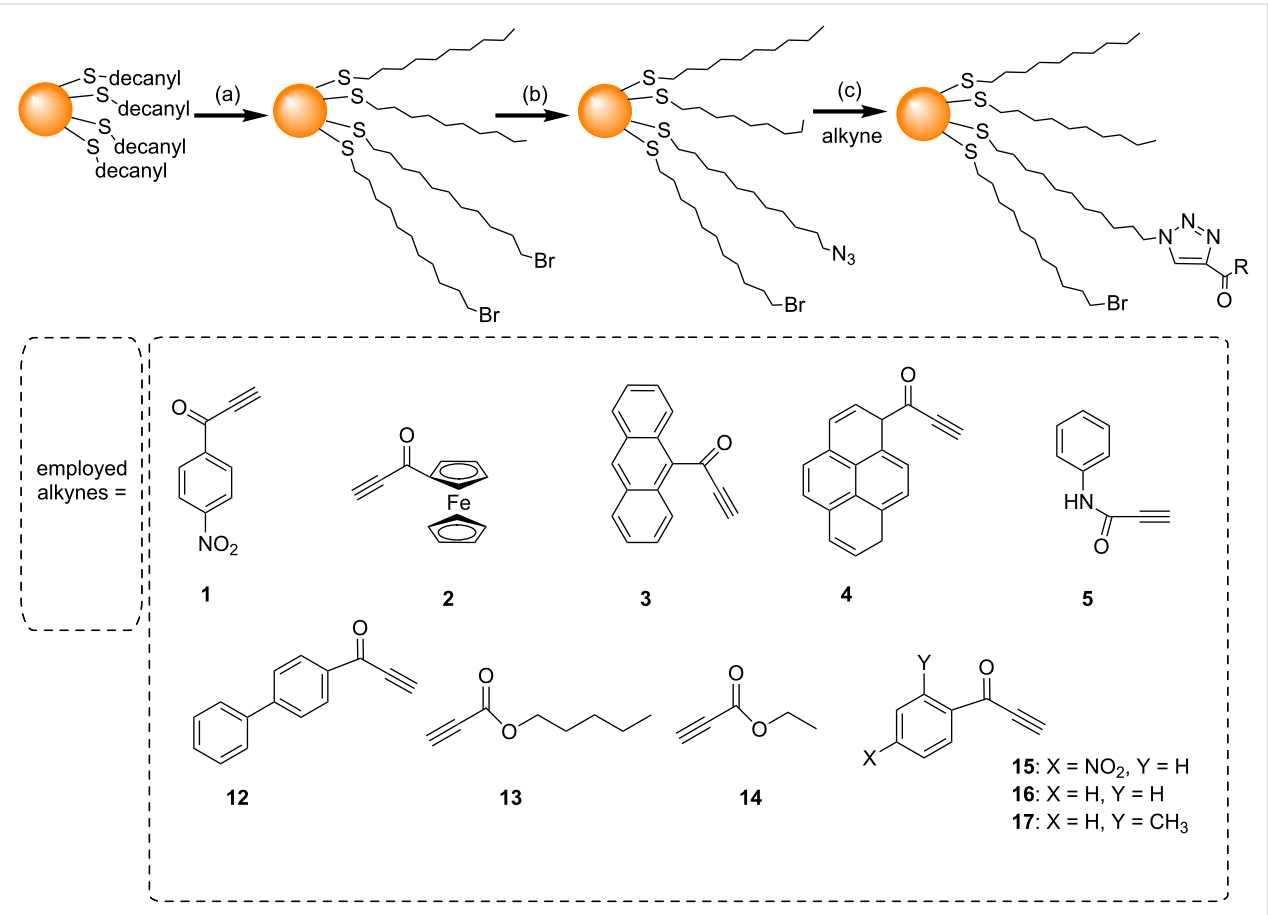
In 2014, Workentin and co-workers used the strain promoted azide–alkyne cycloaddition (SPAAC) [53-56] to modify AuNP surfaces [57]. Firstly 2.8 nm AuNPs functionalized with strained dibenzocyclooctyne derivatives (DBCO-AuNPs) were synthesized in two steps (Scheme 5). Herein, the treatment of methyl-terminated triethylene glycol monolayer-protected AuNPs (Me-EG_3-AuNPs) with ω-carboxy tetraethylene glycol thiols (HOOC-EG_4-SH) gave carboxy-functionalized AuNPs (HOOC-EG_4-AuNPs). Peptide coupling of these HOOC-EG_4-



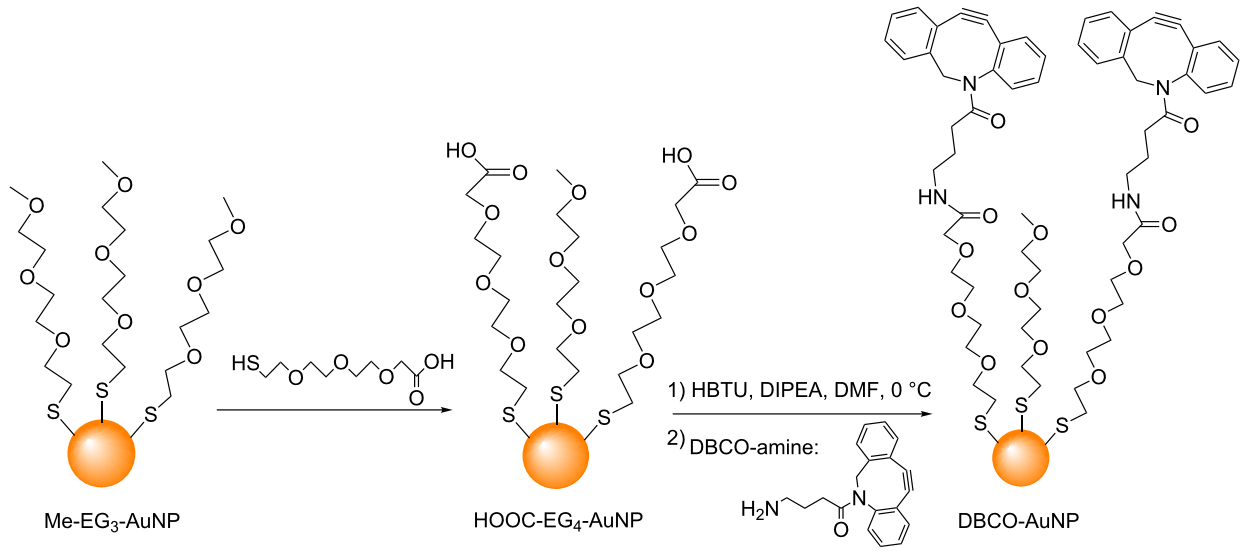
Scheme 2: Ligand exchange and NCAAC on an AuNP surface. Reagents and conditions: (a) $Br(CH_2)_{11}SH$ in DCM, 60 h, rt; (b) NaN_3, DCM/DMSO, 48 h; (c) R = propyn-1-one derivatives, 24–96 h in dioxane, or 1:1 hexane/dioxane [45].



Scheme 3: Azide functionalization and NCAAC on an AuNP surface using electron deficient alkynes. Reagents and conditions: (a) $HS(CH_2)_{11}N_3$, C_6H_6, rt, 7 h; (b) THF, rt, 16 h [52].



Scheme 4: NCAAC performed under hyperbaric conditions. Reagents and conditions: (a) $Br(CH_2)_{11}SH$ in C_6H_6, 48 h, rt; (b) NaN_3 in C_6H_6/DMSO, 48 h; (c) R = propyn-1-one derivatives, DCM, 11000 atm, 25 °C, 15–24 h [48].



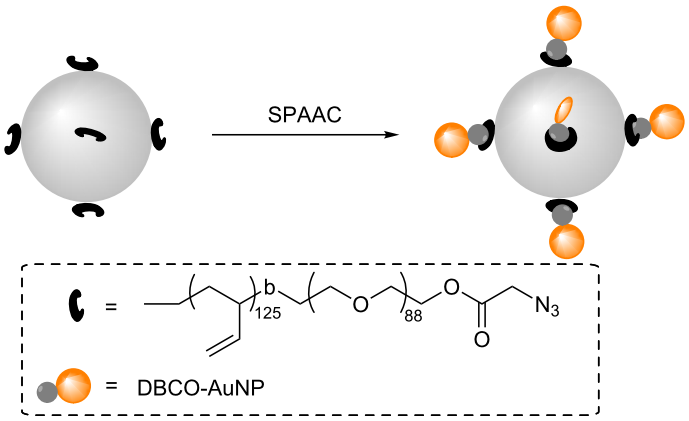
Scheme 5: The synthesis of AuNPs functionalized with strained alkyne derivatives. HBTU = *O*-benzotriazole-*N*,*N*,*N*',*N*'-tetramethyluroniumhexafluorophosphate; DIPEA = *N*,*N*-diisopropylethylamine [57].

AuNPs with a DBCO-amine then yielded the DBCO-AuNPs. When these DBCO-AuNPs were treated with azide-decorated polymersomes (a class of artificial vesicles) [58], the AuNPs were successfully attached to the surface of the polymersomes (Scheme 6). Workentin and co-workers have also reported the successful use of SPAAC to synthesize peptide-decorated AuNPs [59] and nanomaterial hybrids containing single walled carbon nanotubes and AuNPs [60].

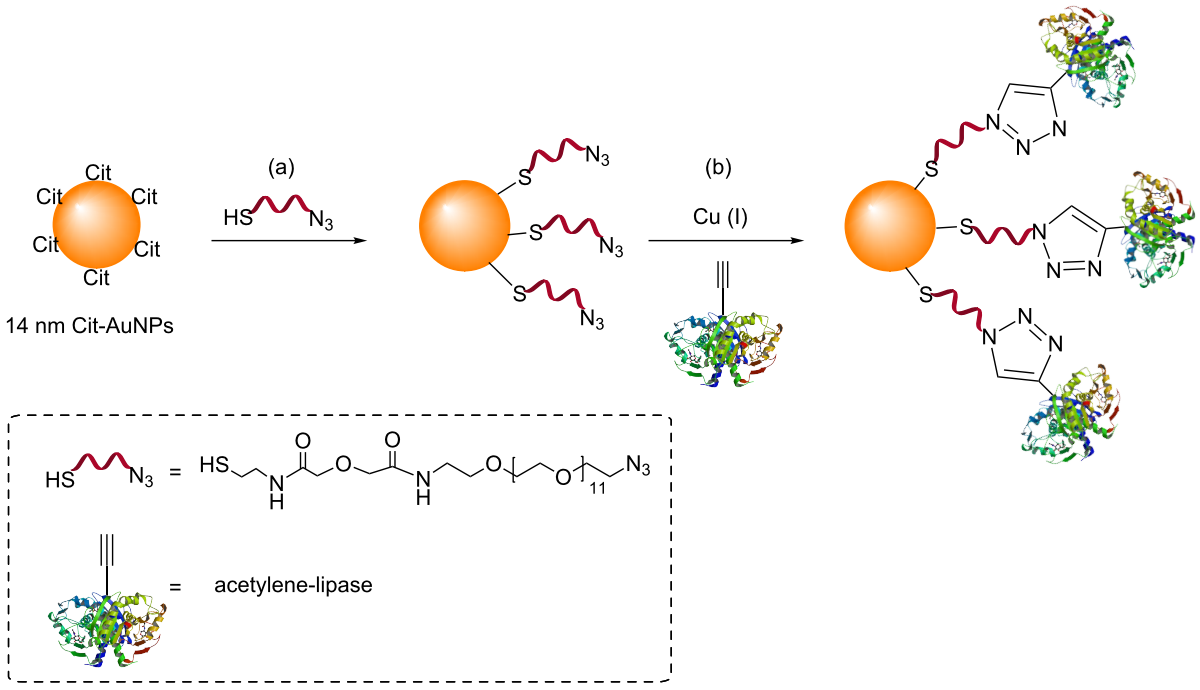
AuNP surface modification by CuAAC

The distinct advantages of CuAAC over NCAAC, such as improved regioselectivity and rates of the reaction, motivated several groups to use CuAAC for the surface modification of AuNPs. In 2006, Brennan et al. demonstrated that enzyme–AuNP conjugates could be synthesized by CuAAC [47]. Azide-functionalized AuNPs were first synthesized by treating standard 14 nm Cit-AuNPs [28] with an a queous solution of an azide-containing thiol ligand (Scheme 7).

An acetylene-functionalized *Thermomyces lanuginosus* lipase was then attached to these azide-functionalized water-soluble AuNPs by CuAAC (Scheme 7). It was found that the enzyme retained its activity after the click reaction. However, the vast excesses of both Cu (a one million-fold excess relative to the azide) and lipase needed, the long reaction time (3 days), the extensive purification procedure required, and the poor overall



Scheme 6: A schematic representation of the SPAAC between azide-functionalized polymersomes and strained alkyne-functionalized AuNPs (DBCO-AuNPs) in water [57].



Scheme 7: Functionalization of AuNPs with an azide containing thiol ligand, and subsequent attachment to an acetylene-functionalized lipase by CuAAC. Reagents and conditions: (a) H_2O, rt, 18 h; (b) H_2O, $CuSO_4$, ascorbic acid, rt, 3 d. [47].

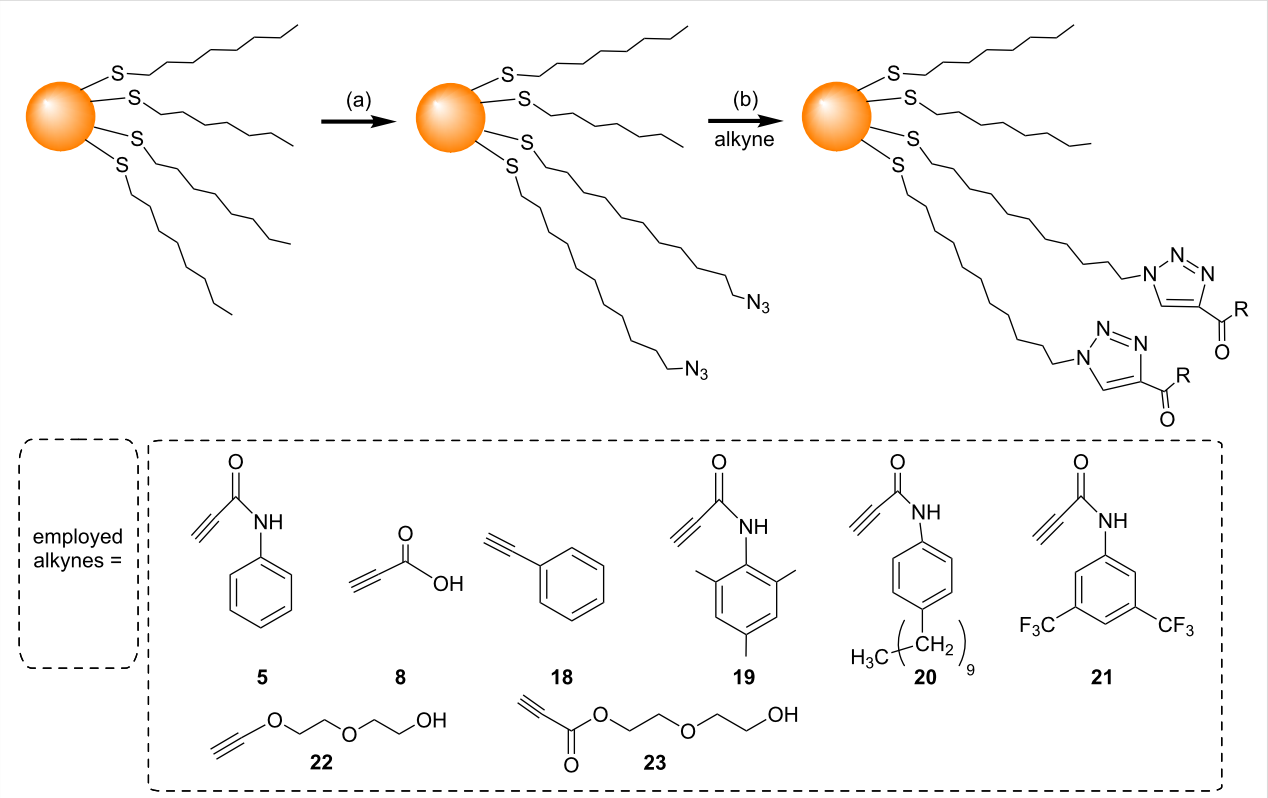
conversion of azide to triazole (less than 1%) limited any further use of this procedure.

In 2007, Sommer and Weck developed a simpler and more efficient method to perform CuAAC on the surface of AuNPs [61]. Herein microwave-assisted CuAAC was used to attach a variety of alkyne derivatives (**5**, **8**, and **18**–**23**) to azide-functionalized AuNPs (Scheme 8). The use of the microwave heating for the CuAAC reduced the reaction time to 5–10 minutes, and also gave almost quantitative conversion of the azides to triazoles. However, significant particle decomposition and/or aggregation were observed when the AuNPs were heated for more than 15 minutes in the microwave reactor.

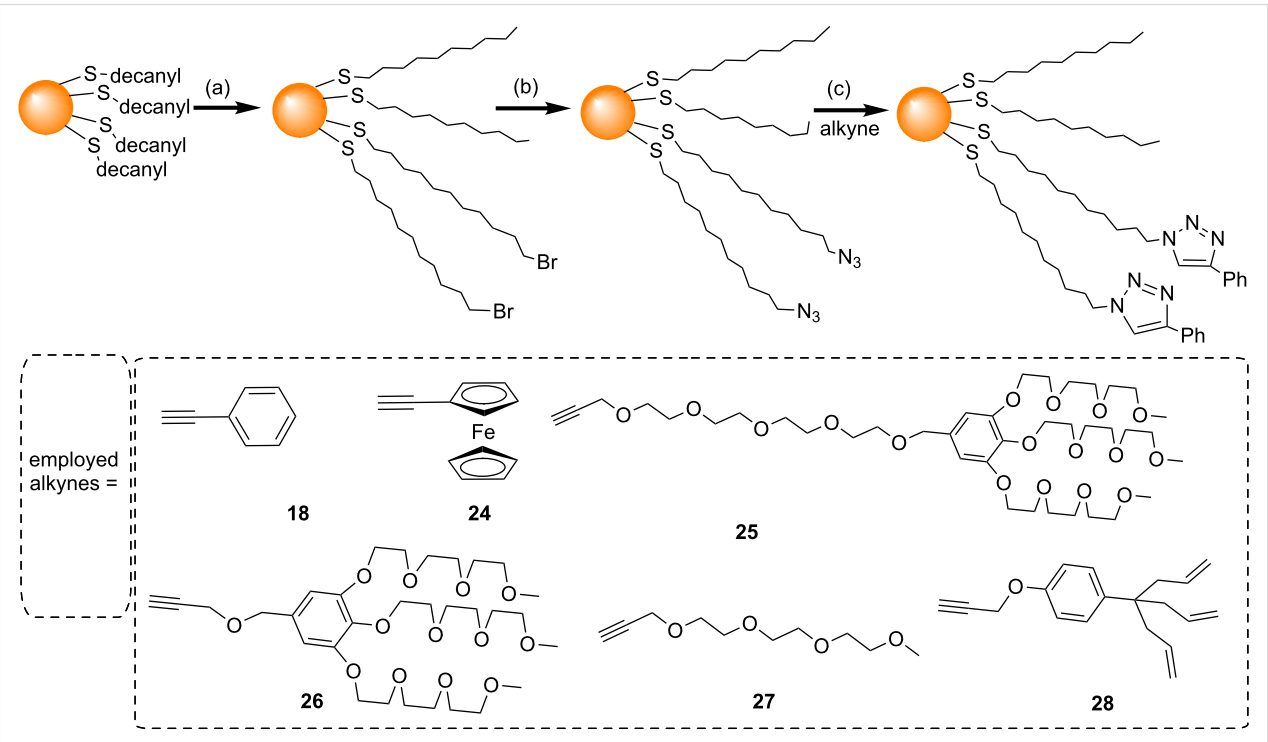
Astruc and co-workers reported several modifications to try and increase the efficiency of CuAAC reactions of AuNPs [62]. They reasoned that one important consideration that needed to be addressed to enable an efficient click reaction was the solubility of the reagents; in particular alkanethiol-functionalized AuNPs are generally only soluble in organic solvents, whereas water is required to dissolve the $CuSO_4$ catalyst. In order to circumvent this solubility problem, a homogenous water/THF solvent system was used, wherein a solution of the AuNPs in THF was added to either an aqueous solution containing water-soluble alkyne derivatives, or to a THF/water solution of organic soluble alkyne derivatives. The amount of ascorbic acid and Cu(I) was also increased to a stoichiometric amount with respect to the alkyne and azide. Finally the click reaction was performed under an inert atmosphere. The authors reported that if any of the above-mentioned conditions were not met, then the reaction gave a very poor yield of product. However, when all the conditions were fulfilled, the conversion of azide to triazole was virtually quantitative at room temperature. The reaction was performed with a variety of alkynes (**18** and **24**–**28**), and good results were obtained despite their variety of sizes and hydrophilicities (Scheme 9).

Astruc and co-workers have also reported that the use of copper(I) (hexabenzyl)tris(2-aminoethyl)amine bromide ([Cu(I)tren(CH_2Ph_6)]Br) instead of the $CuSO_4$–ascorbic acid system improves the efficiency of CuAAC for the functionalization of AuNPs with a wide variety of organic, organometallic, polymeric and dendronic alkynes of different sizes and hydrophilicities [63,64]. CuAAC worked with a catalytic amount of [Cu(I)tren(CH_2Ph_6)]Br under ambient conditions with good yields and without any particle aggregation.

Following these reports, several groups have used the CuAAC reaction of AuNPs as a means for the detection of copper(II) salts [65-67] and ascorbic acid [68], and also for protein quantification (i.e., for proteins capable of reducing Cu(II) to Cu(I)) [69]. The basis of these detection systems was that two sets of



Scheme 8: Surface modification of AuNPs using microwave-assisted CuAAC. Reagents and conditions: (a) $HS(CH_2)_{11}N_3$, C_6H_6, rt, 7 h; (b) dioxane/*t*-BuOH/H_2O or THF, $CuSO_4$, sodium ascorbate, microwave heating (1000 W), 5–10 minutes [61].



Scheme 9: AuNP functionalization and efficient CuAAC with a range of alkynes reported by Boisselier et al. [62]. Reagents and conditions: (a) $HS(CH_2)_{11}Br$, DCM, rt, 5 d; (b) NaN_3, DCM/DMSO, rt, 2 d; (c) $CuSO_4$, sodium ascorbate, THF/H_2O, 2 d, inert atmosphere.

AuNPs were synthesized, one of which was functionalized with azide-containing ligands and the other with alkyne-containing ligands. When these two were mixed in the presence of the required reagents and the corresponding analyte, a click reaction occurred causing aggregation of the AuNPs. The colour change and the surface plasmon resonance band shift induced by the particle aggregation thus served as the basis for the analyte detection.

The functionalization of AuNPs with carbohydrates using AAC

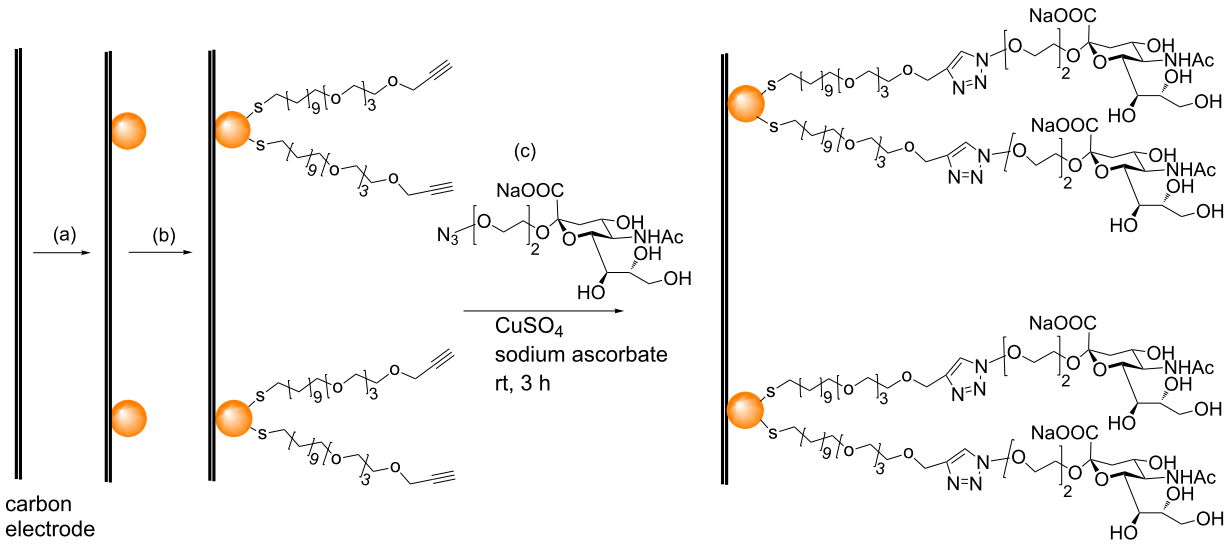
The functionalization of AuNPs with carbohydrates using CuAAC

Although several groups have used the CuAAC to attach thiol-containing ligands to various sugars and then subsequently attach these sugar-containing thiol ligands to AuNPs [70-73], there has so far only been one study that reported the use of the CuAAC to click sugars directly onto the surface of AuNPs. In 2008, Chikae et al. reported the use of CuAAC to react alkyne-terminated thiol-functionalized AuNPs that had been deposited on a carbon electrode with an azide-terminated sialic acid derivative [74]. Firstly, AuNPs were electrodeposited on a carbon electrode. Then a solution of an alkyne-terminated disulphide (4,7,10,13,38,41,44,47-octaoxa-25,26-dithiapentaconta-1,49-diyne) was 'dropped over' the AuNP-electrode system to cover the AuNP surfaces with alkyne-terminated SAMs (Scheme 10). Next, a CuAAC reaction was used to couple the alkyne-functionalized AuNPs to an azide-linked sialic acid derivative, to produce GAuNPs attached to the carbon electrode. This sialic acid-functionalized GAuNP-carbon electrode system was then used for the detection of amyloid-β peptides [74], whose aggregation is responsible for Alzheimer's disease [75].

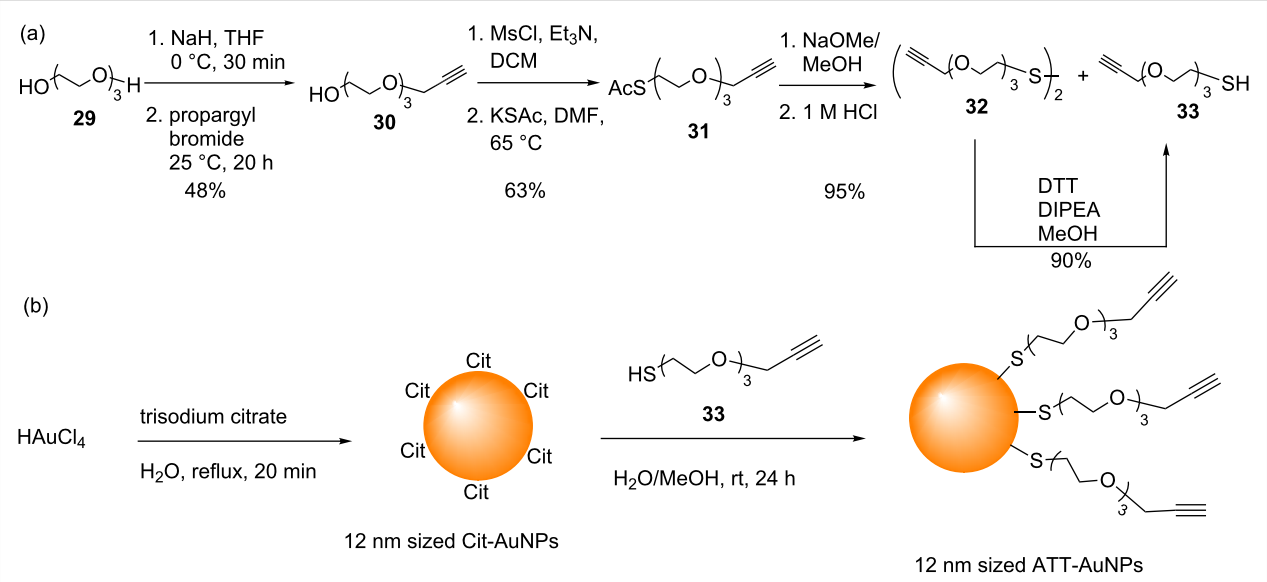
In 2014, Fairbanks and co-workers reported a one-pot aqueous compatible method for making various triazole-linked glycoconjugates via intermediate glycosyl azides, which then underwent CuACC with a wide variety of alkynes [76]. The scarcity of reports on the use of the CuAAC for the functionalization of AuNPs with carbohydrates and the simplicity of the one-pot formation of glycosyl azides and their subsequent reaction with alkynes motivated us to investigate the use of this reaction sequence for the synthesis of GAuNPs.

Firstly, the alkyne-terminated thiol (ATT) ligand **33** was synthesized as shown in Scheme 11a (see Supporting Information File 1 for full experimental data). Next, 12 nm ATT-AuNPs were synthesized by a ligand exchange reaction of 12 nm Cit-AuNPs (themselves synthesized by the Turkevich reaction) with the ATT **33** (Scheme 11b, see Supporting Information File 1 for full experimental data).

The particles obtained by this sequence were not soluble in either water or polar organic solvents, such as MeOH or MeCN, but they were soluble in non-polar solvents, such as DCM, $CHCl_3$, and THF. The broad peaks corresponding to the ligand ATT **33** protons in the ^1H NMR spectra of the purified ATT-AuNPs (Supporting Information File 1, Figure S1) confirmed the attachment of the ATT **33** to the AuNPs. Thermogravimetric analysis of ATT-AuNPs (Figure S2) and the size distribution of Cit-AuNPs and ATT-AuNPs (Figure S3) are also provided in Supporting Information File 1.



Scheme 10: Schematic illustration of: (a) AuNP deposition on a carbon electrode; (b) formation of alkyne-terminated SAMs on these AuNPs; (c) conversion of these AuNPs into GAuNPs by CuAAC [74].



Scheme 11: (a) Synthesis of the alkyne-terminated thiol (ATT) ligand **33**; (b) synthesis of 12 nm sized ATT-AuNPs by ligand exchange.

Whenever water-soluble ligands are used to perform exchange reactions on Cit-AuNPs, the wine-red colour of the AuNP solution (which corresponds to the dispersed state of the AuNPs as can be confirmed by TEM), and the SPR peak in the UV–vis spectrum are typically unchanged. However, in this case, when the water-insoluble ligand **33** was used, the solution turned purple (Supporting Information File 1, Figure S4), and the SPR peak shifted to a higher wavelength (523 nm to 541 nm) and became broader (Supporting Information File 1, Figure S5). Furthermore TEM revealed partial aggregation of the particles (Supporting Information File 1, Figure S6). However, despite this partial aggregation the ATT-AuNP solution was stable without any precipitation at least for three months when stored at 4 °C. Similar observations have been reported by Baranov et al. [77].

GlcNAc azide **34** was synthesized following the reported procedure (Supporting Information File 1) [76], and CuAAC of azide **34** and the AAT-AuNPs was attempted (Supporting Information File 1). Initially, only 1.5 mol % of $CuSO_4{\cdot}5H_2O$ (with respect to the ligands on the AAT-AuNPs) was used. However, ^{1}H NMR analysis of the AuNPs revealed that the particles had not reacted with the glycosyl azide. Following the report of Boisselier et al. [62], a stoichiometric amount of $CuSO_4{\cdot}5H_2O$ was then used, and the reaction was performed under a nitrogen atmosphere. Firstly a solution of AAT-AuNPs in THF was added to an aqueous solution of the crude glycosyl azide, and then ascorbic acid, and finally a solution of $CuSO_4{\cdot}5H_2O$ dissolved in water were added. However, as soon as the $CuSO_4{\cdot}5H_2O$ was added, the particles precipitated; thus the click reaction failed and no GAuNPs were obtained. In further experiments the CuAAC was attempted using a solution of purified GlcNAc azide **34**. Water and THF were used as the solvent in a 1:1 ratio to be in line with the conditions reported by Boisselier et al. [62]. However, even with these conditions precipitation of the particles could not be prevented. Although this did confirm that neither the reagents nor byproducts from the azide synthesis were responsible for the particle aggregation, ultimately the reaction was unsuccessful. We include this finding in this comprehensive account in order to draw conclusions from it.

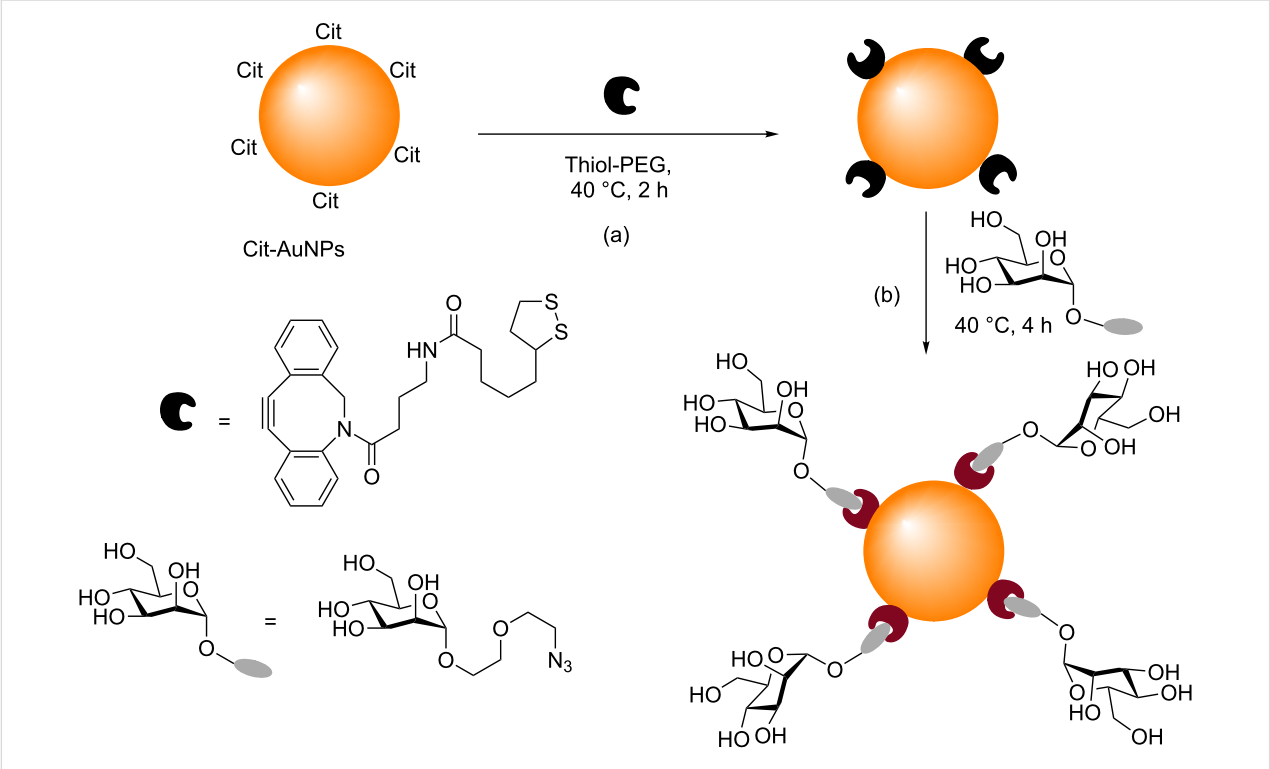
While several groups have demonstrated the successful use of CuAAC for the modification of AuNPs [47,61,62,78,79], at least three groups have reported that attempts to modify azide-functionalized AuNPs with alkyne derivatives by CuAAC either resulted in the reversible aggregation of the particles, or in negligible conversion [45,52,57]. For example, Fleming et al. reported attempts to increase the yield of the AAC using several different Cu-based catalyst systems [45]. As the particles (AuNPs functionalized with a mixture of decanethiol, Br-terminated undecanethiol, and azide-terminated undecanethiol) were insoluble in aqueous solutions, the most frequently used $CuSO_4$-ascorbic acid system could not be used. Thus catalysts soluble in organic solvents, such as CuI, CuBr/2,6-lutidine, and bromotris(triphenylphosphinato)copper(I) were investigated. However in all cases, rapid and extensive particle aggregation or decomposition was observed. Limapichat et al. also reported similar results when Cu catalysts were used to accelerate the cycloaddition reaction [52]. In order to demonstrate the advantages of Cu-free SPAAC reactions, Workentin and co-workers compared Cu-free and Cu-catalysed click reactions with small

water soluble AuNPs (particles functionalized with a mixture of Me-EG_3-SH and N_3-EG_4-SH). Their attempts to perform CuAAC between the azide-modified AuNPs and alkynes (2-propyn-1-amine hydrochloride or 1-ethynylpyrene) in the presence of $CuSO_4$ and sodium ascorbate resulted in particle decomposition [57]. However, when they performed SPAAC of the azide-modified AuNPs and dibenzocyclooctyne (DBCO)-amine, cycloaddition was complete after one hour, and gave the product in 60% yield. Hence, they suggested that the reaction of Cu(I) salts with the Au surface caused the particles to undergo aggregation/decomposition during the CuAAC [57]. It seems therefore that our attempts to synthesize GAuNPs using the one pot glycosyl azide/CuAAC reaction ran into the same limitations as reported by these three groups.

Boisselier et al. reported that by employing specific conditions, namely stoichiometric quantities of both $CuSO_4$ and sodium ascorbate, a 1:1 mixture of water/THF as the reaction solvent, and a nitrogen atmosphere, CuAAC could be used to modify the surface of AuNPs [62]. However, it is notable that these reactions involved 2.5 nm AuNPs. Since the properties of AuNPs are highly dependent on their size, it may be that although the conditions reported by Boisselier et al. work well for smaller sized particles, however, may not be enough to overcome the precipitation of the larger sized AuNPs (>10 nm) caused by Cu as observed by some groups. Unfortunately our attempts to synthesize smaller sized (≈2 nm) ATT-AuNPs, either using two-phase (water/toluene) [49], or one-phase (MeOH) Brust–Schiffrin methods (BSM) [80] both resulted in the formation of decomposed/aggregated particles. We postulate that perhaps reaction of $HAuCl_4$ with the terminal alkyne [81] of ATT **33** might have interfered with the Brust–Schiffrin reaction, and resulted in the formation of unstable AuNPs.

The functionalization of AuNPs with carbohydrates using SPAAC

An alternative method for the functionalization of AuNPs with carbohydrates using click chemistry has recently been reported by Tian and co-workers [82]. They used SPAAC in their one-pot stepwise preparation of GAuNPs, and then used those particles as supramolecular glycoprobes for the rapid serological recognition of a cancer biomarker. Firstly, ligand exchange was performed on Cit-AuNPs by reaction with a THF solution of a cyclooctyne disulfide and an aqueous solution of tetraethylene glycol–thiol (dilutor ligands), to produce particles decorated with cyclooctynes (Scheme 12). These AuNPs then underwent SPAAC when an aqueous solution of a mannose-derived azide was added, to produce mannose-functionalized GAuNPs (Scheme 12). In the presence of the mannose-specific, dimeric lectin LcA (*Lens culinaris lectin*), these GAuNPs underwent aggregation. The GAuNP aggregates that were formed were then used as a supramolecular glycoprobe for the rapid detec-



Scheme 12: Synthesis of (a) cyclooctyne-functionalized AuNPs and (b) GAuNPs using SPAAC [82].

tion of α-fetoprotein (AFP)-L3, a protein which binds strongly to LcA and is a serological biomarker for hepatocellular carcinoma (HCC). In this study AFP-L3 was captured on a microplate and the GAuNPs were added. The strong binding interaction between AFP-L3 and LcA caused disruption of the GAuNP-LcA aggregates, and a change in the optical density of the GAuNPs, which was measured with a microplate reader, enabling the detection of AFP-L3. Clearly this successful synthesis of GAuNPs by Tian and co-workers demonstrates that by employing SPAAC the Cu-induced aggregation/decomposition of AuNPs observed under CuAAC reactions as reported by some groups [45,52,57] can be avoided.

Conclusion

With the combined features of an Au core and a surface decorated with multiple copies of biologically relevant carbohydrates, GAuNPs have become valuable tools in glycoscience. The simplicity and the versatility of the azide–alkyne Huisgen cycloaddition has stimulated several recent attempts to employ this type of reaction for the production of GAuNPs. When the non-catalysed azide–alkyne Huisgen cycloaddition was used for the surface modification of AuNPs, either the yields (i.e., the extent of the azide conversion to triazole) were poor, or long reaction times or hyperbaric conditions were required. There are somewhat conflicting reports in the literature with regard to the use of Cu(I)-catalysed azide–alkyne cycloaddition with AuNPs. Indeed although several groups have reported the successful use of CuAAC for the modification of AuNPs, both our own investigations, and those of number of other groups, have found that AuNP precipitation occurred under CuAAC reaction conditions [45,52,57]. Moreover the immediate precipitation of AuNPs that was observed upon the addition of $CuSO_4{\cdot}5H_2O$ implies that it was the Cu catalyst that caused precipitation. The precise reasons for this AuNP aggregation are not yet clear. Also, it seems difficult to extract a definite reason to explain as to why the CuAAC with AuNPs works for some groups while it fails in some other groups. However, in order to circumvent the limitations of CuAAC, SPAAC can be used as an alternative, and this provides a reliable method for the functionalization of AuNPs with carbohydrates using the azide–alkyne Huisgen cycloaddition.

Supporting Information

Supporting Information File 1

Synthetic protocols and spectral and TEM characterisation for ATT **33** (Scheme 11), ATT-AuNPs (Scheme 11), GlcNAc azide **34**, and click reaction of ATT-AuNPs.
[http://www.beilstein-journals.org/bjoc/content/supplementary/1860-5397-14-2-S1.pdf]

ORCID® iDs

Vivek Poonthiyil - https://orcid.org/0000-0002-1154-1692
Vladimir B. Golovko - https://orcid.org/0000-0002-2679-8917
Antony J. Fairbanks - https://orcid.org/0000-0002-9975-3269

References

1. Thakor, A. S.; Jokerst, J.; Zavaleta, C.; Massoud, T. F.; Gambhir, S. S. *Nano Lett.* **2011,** *11,* 4029–4036. doi:10.1021/nl202559p
2. Zhao, P.; Li, N.; Astruc, D. *Coord. Chem. Rev.* **2013,** *257,* 638–665. doi:10.1016/j.ccr.2012.09.002
3. Elghanian, R.; Storhoff, J. J.; Mucic, R. C.; Letsinger, R. L.; Mirkin, C. A. *Science* **1997,** *277,* 1078–1081. doi:10.1126/science.277.5329.1078
4. Saha, K.; Agasti, S. S.; Kim, C.; Li, X.; Rotello, V. M. *Chem. Rev.* **2012,** *112,* 2739–2779. doi:10.1021/cr2001178
5. Turner, M.; Golovko, V. B.; Vaughan, O. P. H.; Abdulkin, P.; Berenguer-Murcia, A.; Tikhov, M. S.; Johnson, B. F. G.; Lambert, R. M. *Nature* **2008,** *454,* 981–983. doi:10.1038/nature07194
6. Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995,** *28,* 321–327. doi:10.1021/ar00056a001
7. Jayaraman, N. *Chem. Soc. Rev.* **2009,** *38,* 3463–3483. doi:10.1039/b815961k
8. de la Fuente, J. M.; Barrientos, A. G.; Rojas, T. C.; Rojo, J.; Cañada, J.; Fernández, A.; Penadés, S. *Angew. Chem.* **2001,** *113,* 2317–2321. doi:10.1002/1521-3757(20010618)113:12<2317::AID-ANGE2317>3.0.CO;2-U
9. Marradi, M.; Chiodo, F.; Garcia, I.; Penadés, S. *Chem. Soc. Rev.* **2013,** *42,* 4728–4745. doi:10.1039/c2cs35420a
10. Compostella, F.; Pitirollo, O.; Silvestri, A.; Polito, L. *Beilstein J. Org. Chem.* **2017,** *13,* 1008–1021. doi:10.3762/bjoc.13.100
11. Marin, M. J.; Schofield, C. L.; Field, R. A.; Russell, D. A. *Analyst* **2015,** *140,* 59–70. doi:10.1039/C4AN01466A
12. Zhao, W.; Brook, M. A.; Li, Y. *ChemBioChem* **2008,** *9,* 2363–2371. doi:10.1002/cbic.200800282
13. Schofield, C. L.; Field, R. A.; Russell, D. A. *Anal. Chem.* **2007,** *79,* 1356–1361. doi:10.1021/ac061462j
14. Otsuka, H.; Akiyama, Y.; Nagasaki, Y.; Kataoka, K. *J. Am. Chem. Soc.* **2001,** *123,* 8226–8230. doi:10.1021/ja010437m
15. Richards, S.-J.; Fullam, E.; Besra, G. S.; Gibson, M. I. *J. Mater. Chem. B* **2014,** *2,* 1490–1498. doi:10.1039/C3TB21821J
16. Niikura, K.; Nagakawa, K.; Ohtake, N.; Suzuki, T.; Matsuo, Y.; Sawa, H.; Ijiro, K. *Bioconjugate Chem.* **2009,** *20,* 1848–1852. doi:10.1021/bc900255x
17. Barrientos, Á. G.; de la Fuente, J. M.; Rojas, T. C.; Fernández, A.; Penadés, S. *Chem. – Eur. J.* **2003,** *9,* 1909–1921. doi:10.1002/chem.200204544
18. Lin, C.-C.; Yeh, Y.-C.; Yang, C.-Y.; Chen, G.-F.; Chen, Y.-C.; Wu, Y.-C.; Chen, C.-C. *Chem. Commun.* **2003,** 2920–2921. doi:10.1039/b308995a
19. Lin, C.-C.; Yeh, Y.-C.; Yang, C.-Y.; Chen, C.-L.; Chen, G.-F.; Chen, C.-C.; Wu, Y.-C. *J. Am. Chem. Soc.* **2002,** *124,* 3508–3509. doi:10.1021/ja0200903
20. Svarovsky, S. A.; Szekely, Z.; Barchi, J. J. *Tetrahedron: Asymmetry* **2005,** *16,* 587–598. doi:10.1016/j.tetasy.2004.12.003
21. Sundgren, A.; Barchi, J. J., Jr. *Carbohydr. Res.* **2008,** *343,* 1594–1604. doi:10.1016/j.carres.2008.05.003

22. de Paz, J.-L.; Ojeda, R.; Barrientos, Á. G.; Penadés, S.; Martín-Lomas, M. *Tetrahedron: Asymmetry* **2005,** *16,* 149–158. doi:10.1016/j.tetasy.2004.11.066
23. Carvalho De Souza, A.; Halkes, K. M.; Meeldijk, J. D.; Verkleij, A. J.; Vliegenthart, J. F. G.; Kamerling, J. P. *Eur. J. Org. Chem.* **2004,** 4323–4339. doi:10.1002/ejoc.200400255
24. Carvalho de Souza, A.; Vliegenthart, J. F. G.; Kamerling, J. P. *Org. Biomol. Chem.* **2008,** *6,* 2095–2102. doi:10.1039/b802235f
25. Chien, Y.-Y.; Jan, M.-D.; Adak, A. K.; Tzeng, H.-C.; Lin, Y.-P.; Chen, Y.-J.; Wang, K.-T.; Chen, C.-T.; Chen, C.-C.; Lin, C.-C. *ChemBioChem* **2008,** *9,* 1100–1109. doi:10.1002/cbic.200700590
26. Marradi, M.; Martín-Lomas, M.; Penadés, S. *Adv. Carbohydr. Chem. Biochem.* **2010,** *64,* 211–290. doi:10.1016/S0065-2318(10)64005-X
27. Schofield, C. L.; Mukhopadhyay, B.; Hardy, S. M.; McDonnell, M. B.; Field, R. A.; Russell, D. A. *Analyst* **2008,** *133,* 626–634. doi:10.1039/b715250g
28. Frens, G. *Nature (London), Phys. Sci.* **1973,** *241,* 20–22. doi:10.1038/physci241020a0
29. Poonthiyil, V.; Golovko, V. B.; Fairbanks, A. J. *Org. Biomol. Chem.* **2015,** *13,* 5215–5223. doi:10.1039/C5OB00447K
30. Poonthiyil, V.; Nagesh, P. T.; Husain, M.; Golovko, V. B.; Fairbanks, A. J. *ChemistryOpen* **2015,** *4,* 708–716. doi:10.1002/open.201500109
31. Chen, F.; Li, X.; Hihath, J.; Huang, Z.; Tao, N. *J. Am. Chem. Soc.* **2006,** *128,* 15874–15881. doi:10.1021/ja065864k
32. Halkes, K. M.; Carvalho De Souza, A.; Maljaars, C. E. P.; Gerwig, G. J.; Kamerling, J. P. *Eur. J. Org. Chem.* **2005,** 3650–3659. doi:10.1002/ejoc.200500256
33. Nagahori, N.; Abe, M.; Nishimura, S.-I. *Biochemistry* **2009,** *48,* 583–594. doi:10.1021/bi801640n
34. Telli, F. C.; Demir, B.; Barlas, F. B.; Guler, E.; Timur, S.; Salman, Y. *RSC Adv.* **2016,** *6,* 105806–105813. doi:10.1039/C6RA21976D
35. Wang, X.; Ramström, O.; Yan, M. *Anal. Chem.* **2010,** *82,* 9082–9089. doi:10.1021/ac102114z
36. Wang, X.; Ramström, O.; Yan, M. *J. Mater. Chem.* **2009,** *19,* 8944–8949. doi:10.1039/b917900c
37. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001,** *40,* 2004–2021. doi:10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5
38. Huisgen, R. *Angew. Chem., Int. Ed. Engl.* **1963,** *2,* 565–598. doi:10.1002/anie.196305651
39. Liang, L.; Astruc, D. *Coord. Chem. Rev.* **2011,** *255,* 2933–2945. doi:10.1016/j.ccr.2011.06.028
40. Tornøe, C. W.; Christensen, C.; Meldal, M. *J. Org. Chem.* **2002,** *67,* 3057–3064. doi:10.1021/jo011148j
41. Himo, F.; Lovell, T.; Hilgraf, R.; Rostovtsev, V. V.; Noodleman, L.; Sharpless, K. B.; Fokin, V. V. *J. Am. Chem. Soc.* **2005,** *127,* 210–216. doi:10.1021/ja0471525
42. Meldal, M.; Tornøe, C. W. *Chem. Rev.* **2008,** *108,* 2952–3015. doi:10.1021/cr0783479
43. Kolb, H. C.; Sharpless, K. B. *Drug Discovery Today* **2003,** *8,* 1128–1137. doi:10.1016/S1359-6446(03)02933-7
44. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew. Chem.* **2002,** *114,* 2708–2711. doi:10.1002/1521-3757(20020715)114:14<2708::AID-ANGE2708>3.0.CO;2-0
45. Fleming, D. A.; Thode, C. J.; Williams, M. E. *Chem. Mater.* **2006,** *18,* 2327–2334. doi:10.1021/cm060157b
46. Li, N.; Binder, W. H. *J. Mater. Chem.* **2011,** *21,* 16717–16734. doi:10.1039/c1jm11558h
47. Brennan, J. L.; Hatzakis, N. S.; Tshikhudo, T. R.; Dirvianskyte, N.; Razumas, V.; Patkar, S.; Vind, J.; Svendsen, A.; Nolte, R. J. M.; Rowan, A. E.; Brust, M. *Bioconjugate Chem.* **2006,** *17,* 1373–1375. doi:10.1021/bc0601018
48. Ismaili, H.; Alizadeh, A.; Snell, K. E.; Workentin, M. S. *Can. J. Chem.* **2009,** *87,* 1708–1715. doi:10.1139/V09-138
49. Brust, M.; Walker, M.; Bethell, D.; Schiffrin, D. J.; Whyman, R. *J. Chem. Soc., Chem. Commun.* **1994,** 801–802. doi:10.1039/C39940000801
50. Collman, J. P.; Devaraj, N. K.; Chidsey, C. E. D. *Langmuir* **2004,** *20,* 1051–1053. doi:10.1021/la0362977
51. Thode, C. J.; Williams, M. E. *J. Colloid Interface Sci.* **2008,** *320,* 346–352. doi:10.1016/j.jcis.2007.12.027
52. Limapichat, W.; Basu, A. *J. Colloid Interface Sci.* **2008,** *318,* 140–144. doi:10.1016/j.jcis.2007.09.054
53. Jewett, J. C.; Bertozzi, C. R. *Chem. Soc. Rev.* **2010,** *39,* 1272–1279. doi:10.1039/b901970g
54. Debets, M. F.; van Berkel, S. S.; Dommerholt, J.; Dirks, A. J.; Rutjes, F. P. J. T.; van Delft, F. L. *Acc. Chem. Res.* **2011,** *44,* 805–815. doi:10.1021/ar200059z
55. Evans, H. L.; Slade, R. L.; Carroll, L.; Smith, G.; Nguyen, Q.-D.; Iddon, L.; Kamaly, N.; Stöckmann, H.; Leeper, F. J.; Aboagye, E. O.; Spivey, A. C. *Chem. Commun.* **2012,** *48,* 991–993. doi:10.1039/C1CC16220A
56. Ning, X.; Guo, J.; Wolfert, M. A.; Boons, G.-J. *Angew. Chem., Int. Ed.* **2008,** *47,* 2253–2255. doi:10.1002/anie.200705456
57. Gobbo, P.; Mossman, Z.; Nazemi, A.; Niaux, A.; Biesinger, M. C.; Gillies, E. R.; Workentin, M. S. *J. Mater. Chem. B* **2014,** *2,* 1764–1769. doi:10.1039/C3TB21799J
58. Amos, R. C.; Nazemi, A.; Bonduelle, C. V.; Gillies, E. R. *Soft Matter* **2012,** *8,* 5947–5958. doi:10.1039/c2sm25172h
59. Wang, X.; Gobbo, P.; Suchy, M.; Workentin, M. S.; Hudson, R. H. E. *RSC Adv.* **2014,** *4,* 43087–43091. doi:10.1039/C4RA07574A
60. Gobbo, P.; Novoa, S.; Biesinger, M. C.; Workentin, M. S. *Chem. Commun.* **2013,** *49,* 3982–3984. doi:10.1039/c3cc41634h
61. Sommer, W. J.; Weck, M. *Langmuir* **2007,** *23,* 11991–11995. doi:10.1021/la7018742
62. Boisselier, E.; Salmon, L.; Ruiz, J.; Astruc, D. *Chem. Commun.* **2008,** 5788–5790. doi:10.1039/b812249k
63. Li, N.; Zhao, P.; Salmon, L.; Ruiz, J.; Zabawa, M.; Hosmane, N. S.; Astruc, D. *Inorg. Chem.* **2013,** *52,* 11146–11155. doi:10.1021/ic4013697
64. Zhao, P.; Grillaud, M.; Salmon, L.; Ruiz, J.; Astruc, D. *Adv. Synth. Catal.* **2012,** *354,* 1001–1011. doi:10.1002/adsc.201100865
65. Zhou, Y.; Wang, S.; Zhang, K.; Jiang, X. *Angew. Chem.* **2008,** *120,* 7564–7566. doi:10.1002/ange.200802317
66. Hua, C.; Zhang, W. H.; De Almeida, S. R. M.; Ciampi, S.; Gloria, D.; Liu, G.; Harper, J. B.; Gooding, J. J. *Analyst* **2012,** *137,* 82–86. doi:10.1039/C1AN15693D
67. Zhang, Z.; Li, W.; Zhao, Q.; Cheng, M.; Xu, L.; Fang, X. *Biosens. Bioelectron.* **2014,** *59,* 40–44. doi:10.1016/j.bios.2014.03.003
68. Zhang, Y.; Li, B.; Xu, C. *Analyst* **2010,** *135,* 1579–1584. doi:10.1039/c0an00056f
69. Zhu, K.; Zhang, Y.; He, S.; Chen, W.; Shen, J.; Wang, Z.; Jiang, X. *Anal. Chem.* **2012,** *84,* 4267–4270. doi:10.1021/ac3010567

70. Papp, I.; Sieben, C.; Ludwig, K.; Roskamp, M.; Böttcher, C.; Schlecht, S.; Herrmann, A.; Haag, R. *Small* **2010,** *6,* 2900–2906. doi:10.1002/smll.201001349
71. Marín, M. J.; Rashid, A.; Rejzek, M.; Fairhurst, S. A.; Wharton, S. A.; Martin, S. R.; McCauley, J. W.; Wileman, T.; Field, R. A.; Russell, D. A. *Org. Biomol. Chem.* **2013,** *11,* 7101–7107. doi:10.1039/c3ob41703d
72. Wei, J.; Zheng, L.; Lv, X.; Bi, Y.; Chen, W.; Zhang, W.; Shi, Y.; Zhao, L.; Sun, X.; Wang, F.; Cheng, S.; Yan, J.; Liu, W.; Jiang, X.; Gao, G. F.; Li, X. *ACS Nano* **2014,** *8,* 4600–4607. doi:10.1021/nn5002485
73. Martos-Maldonado, M. C.; Thygesen, M. B.; Jensen, K. J.; Vargas-Berenguel, A. *Eur. J. Org. Chem.* **2013,** 2793–2801. doi:10.1002/ejoc.201300205
74. Chikae, M.; Fukuda, T.; Kerman, K.; Idegami, K.; Miura, Y.; Tamiya, E. *Bioelectrochemistry* **2008,** *74,* 118–123. doi:10.1016/j.bioelechem.2008.06.005
75. Miura, Y.; Yasuda, K.; Yamamoto, K.; Koike, M.; Nishida, Y.; Kobayashi, K. *Biomacromolecules* **2007,** *8,* 2129–2134. doi:10.1021/bm0701402
76. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J. *Angew. Chem.* **2014,** *126,* 12101–12105. doi:10.1002/ange.201406694
77. Baranov, D.; Kadnikova, E. N. *J. Mater. Chem.* **2011,** *21,* 6152–6157. doi:10.1039/c1jm10183h
78. Zhang, M.-X.; Huang, B.-H.; Sun, X.-Y.; Pang, D.-W. *Langmuir* **2010,** *26,* 10171–10176. doi:10.1021/la100315u
79. Kim, Y.-P.; Daniel, W. L.; Xia, Z.; Xie, H.; Mirkin, C. A.; Rao, J. *Chem. Commun.* **2010,** *46,* 76–78. doi:10.1039/B915612G
80. Brust, M.; Fink, J.; Bethell, D.; Schiffrin, D.; Kiely, C. *J. Chem. Soc., Chem. Commun.* **1995,** 1655–1656. doi:10.1039/c39950001655
81. Hashmi, A. S. K. *Chem. Rev.* **2007,** *107,* 3180–3211. doi:10.1021/cr000436x
82. He, X.-P.; Hu, X.-L.; Jin, H.-Y.; Gan, J.; Zhu, H.; Li, J.; Long, Y.-T.; Tian, H. *Anal. Chem.* **2015,** *87,* 9078–9083. doi:10.1021/acs.analchem.5b02384

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Aminosugar-based immunomodulator lipid A: synthetic approaches

Alla Zamyatina

Review

Open Access

Address:
Department of Chemistry, University of Natural Resources and Life Sciences, Muthgasse 18, 1190 Vienna, Austria

Email:
Alla Zamyatina - alla.zamyatina@boku.ac.at

Keywords:
glycoconjugate; glycolipids; glycosylation; immunomodulation; lipopolysaccharide; TLR4

Beilstein J. Org. Chem. **2018,** *14,* 25–53.
doi:10.3762/bjoc.14.3

Received: 15 June 2017
Accepted: 23 October 2017
Published: 04 January 2018

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

The immediate immune response to infection by Gram-negative bacteria depends on the structure of a lipopolysaccharide (LPS, also known as endotoxin), a complex glycolipid constituting the outer leaflet of the bacterial outer membrane. Recognition of picomolar quantities of pathogenic LPS by the germ-line encoded Toll-like Receptor 4 (TLR4) complex triggers the intracellular pro-inflammatory signaling cascade leading to the expression of cytokines, chemokines, prostaglandins and reactive oxygen species which manifest an acute inflammatory response to infection. The “endotoxic principle” of LPS resides in its amphiphilic membrane-bound fragment glycophospholipid lipid A which directly binds to the TLR4·MD-2 receptor complex. The lipid A content of LPS comprises a complex mixture of structural homologs varying in the acylation pattern, the length of the (*R*)-3-hydroxyacyl- and (*R*)-3-acyloxyacyl long-chain residues and in the phosphorylation status of the β(1→6)-linked diglucosamine backbone. The structural heterogeneity of the lipid A isolates obtained from bacterial cultures as well as possible contamination with other pro-inflammatory bacterial components makes it difficult to obtain unambiguous immunobiological data correlating specific structural features of lipid A with its endotoxic activity. Advanced understanding of the therapeutic significance of the TLR4-mediated modulation of the innate immune signaling and the central role of lipid A in the recognition of LPS by the innate immune system has led to a demand for well-defined materials for biological studies. Since effective synthetic chemistry is a prerequisite for the availability of homogeneous structurally distinct lipid A, the development of divergent and reproducible approaches for the synthesis of various types of lipid A has become a subject of considerable importance. This review focuses on recent advances in synthetic methodologies toward LPS substructures comprising lipid A and describes the synthesis and immunobiological properties of representative lipid A variants corresponding to different bacterial species. The main criteria for the choice of orthogonal protecting groups for hydroxyl and amino functions of synthetically assembled β(1→6)-linked diglucosamine backbone of lipid A which allows for a stepwise introduction of multiple functional groups into the molecule are discussed. Thorough consideration is also given to the synthesis of 1,1′-glycosyl phosphodiesters comprising partial structures of 4-amino-4-deoxy-β-L-arabinose modified *Burkholderia* lipid A and galactosamine-modified *Francisell*a lipid A. Particular emphasis is put on the stereoselective construction of binary

glycosyl phosphodiester fragments connecting the anomeric centers of two aminosugars as well as on the advanced P(III)-phosphorus chemistry behind the assembly of zwitterionic double glycosyl phosphodiesters.

Introduction

The mammalian innate immune system possesses an efficient and incredibly complex evolutionary ancient machinery responsible for host defence against pathogens. The receptors of the innate immune system can detect particular components present in bacteria, viruses or fungi which are designated as "pathogen associated molecular patterns" (PAMPs) [1]. These receptors, termed pattern recognition receptors (PRRs), are able of sensing and responding to PAMPs. The major surface antigen of Gram-negative bacteria, a complex heterogeneous glycolipid lipopolysaccharide (LPS, or endotoxin) [2,3], is recognised by a receptor complex composed of Toll-like Receptor 4 (TLR4) and a co-receptor protein myeloid differentiation factor 2 (MD-2) which are expressed by mammalian immune cells such as macrophages, monocytes and dendritic cells [4]. LPS represents the major virulence factor of Gram-negative bacteria and is essential for bacterial survival. LPS constitutes the outer leaflet of the outer membrane of Gram-negative bacteria (Figure 1A) and possesses a complex micro-heterogeneous structure distinguished by three regions: the lipid A [5], the core oligosaccharide [6] and the O-antigen [7] (Figure 1B). The TLR4·MD-2 receptor complex senses picomolar amounts of LPS and initiates the biosynthesis of diverse mediators of inflammation (such as tumor necrosis factor-TNF-α, interleukin 6 (IL-6) and IL-8) thereby triggering a downstream pro-inflammatory signaling cascade aimed at the clearance of infection [8]. Thus, LPS-induced TLR4·MD-2-mediated signaling largely contributes to the development of inflammation and initiation of the beneficial defensive host response which is essential for bacterial clearance and managing the Gram-negative bacterial disease.

However, under circumstances of an upregulated inflammation, the TLR4 activation results in the excessive production of the pro-inflammatory mediators [9] leading to overstimulation of the innate immune system and systemic inflammatory response syndrome (SIRS) which eventually results in a life-threatening sepsis syndrome and lethal septic shock [10,11] (the 10th leading cause of death in developed countries, 40–60% mortality rate) [12,13]. The membrane-bound portion of LPS, a glycophospholipid lipid A (Figure 1C), constitutes the "endotoxic principle" of LPS [14,15]. In depth studies demonstrated that the lipid A moiety of *E. coli* LPS causes a similar scope of

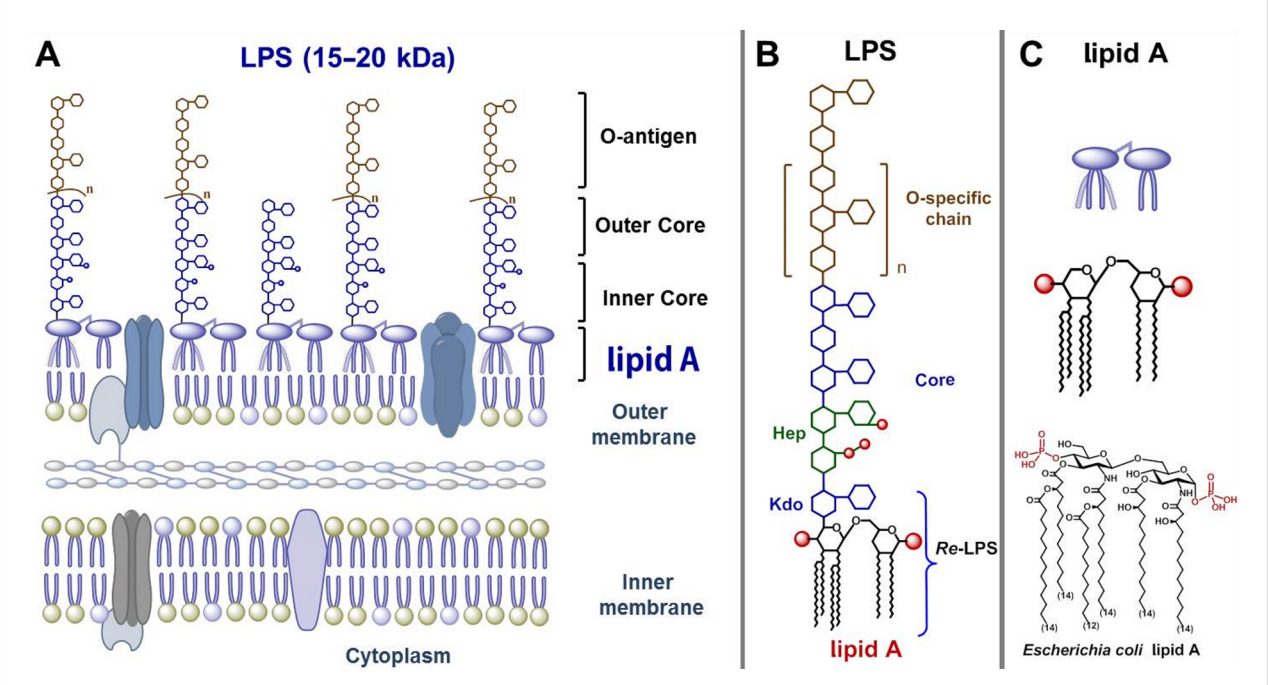


Figure 1: (A) Gram-negative bacterial membrane with LPS as major component of the outer membrane; (B) structural constituents of LPS: lipid A, inner/outer core and O-specific chain.

sepsis-associated effects as its parent LPS which confirmed the proposed key role of lipid A in Gram-negative sepsis syndrome [15].

The chemical structure of lipid A is based on the β(1→6)-linked 1-,4′-bisphosphorylated diglucosamine backbone which is typically tetra- till heptaacylated at the amino groups (positions 2 and 2') and hydroxyl groups (positions 3 and 3') by (*R*)-3-hydroxy- or/and (*R*)-3-acyloxyacyl fatty acids of variable lengths usually comprising 12–16 carbon atoms [16,17]. The endotoxic activity of lipid A depends on numerous factors such as acylation and phosphorylation pattern [18], the length of lipid chains, and the tertiary 3D structure of the MD-2 bound βGlcN(1→6)GlcN backbone [19,20]. The most profoundly studied lipid A of *Escherichia coli* and *Neisseria meningitidis* contains six acyl chains (C_{14}–C_{12}) differently distributed across the diglucosamine backbone and two phosphate groups – one at the anomeric position of the proximal GlcN residue and the second at position 4' of the distal GlcN moiety (Figure 2). These lipid A variants are highly endotoxic and represent the most effective stimulators of the intracellular pro-inflammatory signaling. However, partial activation of the TLR4·MD-2 complex by certain lipid A substructures (such as 1-*O*-dephosphorylated *Salmonella minnesota* lipid A – a licenced vaccine adjuvant monophosphoryl lipid A, MPLA – leads to the induction of a different cytokine profile that weakens toxicity but preserves the beneficial adjuvant effects of endotoxin. Other Gram-negative bacteria can produce lipid A variants which are either less endotoxic or inactive (e.g., cannot be recognised by the TLR4·MD-2 complex) such as tetraacylated 1-*O*-monophosphorylated *Helicobacter pylori* lipid A (Figure 2) [21]. Underacylated lipid A of some Gram-negative organisms exhibit TLR4 antagonist activity, for example, pentaacyl lipid A from *Rhodobacter sphaeroides* [22] or C_{14}-tetraacylated biosynthetic precursor of *E. coli* lipid A, lipid IVa [23] (Figure 2).

Many Gram-negative bacteria, particularly those with mammalian and environmental reservoirs, can produce modified forms of LPS in response to growth conditions, especially in response to a shift in growth temperature (e.g, 37 °C in human host vs 25 °C in a disease vector). These modifications include, in the first line, a cleavage of one or more acyl chains from the lipid A portion of LPS which results in the production of underacylated LPS variants which are "overseen" by the innate immune system of the host. For instance, *Yersinia pestis* produces tetraacylated lipid A in mammalian host compared to the hexaacylated lipid A in the insect vector which renders the bacterium resistant to the hosts innate immune system [24]. Lipid A modifications result in the "remodeling" of the bacterial membrane which alters the outer membrane integrity and antigen presentation, decreases susceptibility to antimicrobial peptides and enhances pathogenicity [25]. In some LPS, the lipid A phosphates are post-translationally modified by substitution with the compounds that reduce the net negative charge of LPS, such as phosphoethanolamine in *E. coli* and *Salmonella* [2,26], ethanolamine in *Helicobacter pylori*, 4-amino-4-deoxy-β-L-arabinose (β-L-Ara4N) [27,28] in *E. coli* [29], *Burkholderia* [27] and *Yersinia pestis* [30] or galactosamine in *Francisella* [2,26], and glucosamine in *Bordetella* species [31] (Figure 2). Covalent attachment of aminosugar to the phosphate groups of lipid A alters the TLR4-mediated host immunity and accounts for the modulation of the pro-inflammatory signaling. Additionally, this modification is associated with an amplified bacterial virulence since it confers resistance to the endogenous cationic antimicrobial peptides (CAMPs) and antibiotics [25,32-34].

Activation of the innate immune response by lipid A/LPS requires a consecutive interaction of lipid A with lipopolysaccharide-binding protein (LPB) [35], glycosylphosphatidylinositol-anchored surface protein CD14 (a differentiation antigen of monocytes) [36,37], followed by a soluble accessory protein MD-2 [38] and TLR4·MD-2 complex [39] (Figure 3) [40-44]. TLR4 is a germ-line encoded transmembrane protein composed of an ectodomain comprising leucin-rich-repeat motifs and a cytoplasmic domain responsible for the initiation of the pro-inflammatory signaling cascade. The lipid A portion of hexaacyl LPS (e.g., in *E. coli* LPS) is recognized and bound by a co-receptor protein MD-2 which is physically asssociated with TLR4. The binding of lipid A initiates dimerization of two copies of the TLR4·MD-2–LPS complexes which results in the formation of a hexameric $[TLR4{\cdot}MD\text{-}2{-}LPS]_2$ complex (Figure 3A). LPS-induced homodimerization of TLR4·MD-2–LPS complexes provokes the recruitment of adaptor proteins to the cytoplasmic TIR (Toll/interleukin-1 receptor) domains of TLR4 which eventually results in the induction of the intracellular pro-inflammatory signaling and activation of the host innate immunity (Figure 3B) [42,45,46].

Compounds which compete with LPS in binding to the same site on MD-2 are capable of inhibiting the induction of the signal transduction pathway by preventing the LPS-induced receptor complex dimerization (Figure 3C). Application of natural or synthetic TLR4 antagonists represents one of the most effective approaches for down-regulation of the TLR4-mediated signaling. So far, several lipid A variants which can block the LPS-binding site on human (h)MD-2 have been identified: tetraacylated lipid IVa [47] and a non-pathogenic lipid A from *R. sphaeroides* [22,48], which served as structural basis for the synthetic antisepsis drug candidate eritoran [49,50]. Inadequate regulation of the TLR4-mediated signaling was recognized as

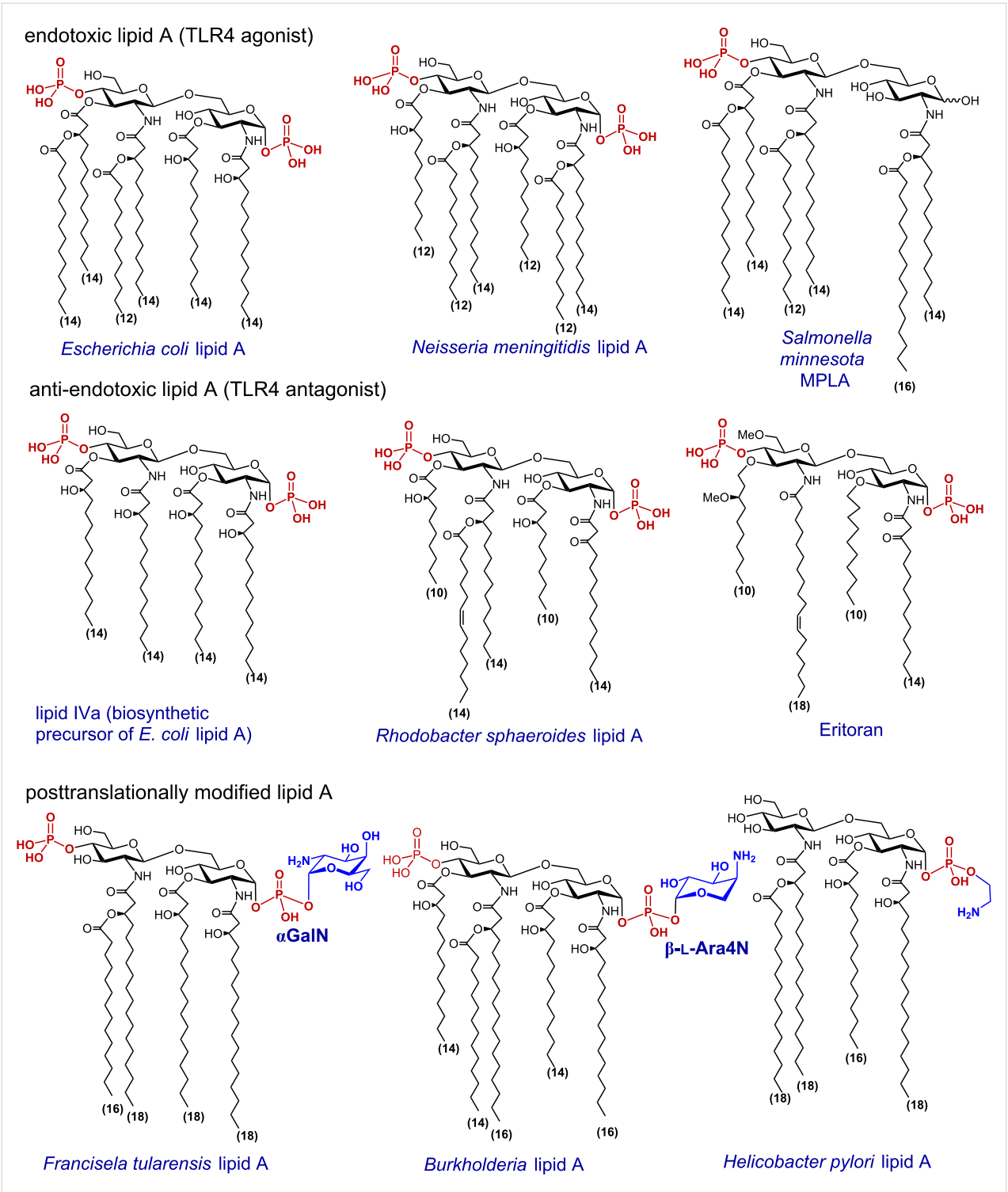


Figure 2: Structures of representative TLR4 ligands: TLR4 agonists (*E. coli* lipid A, *N. meningitidis* lipid A and MPLA) and TLR4 antagonists (lipid IVa, *R. sphaeroides* lipid A and eritoran (E5564)); examples of post-translationally modified lipid A from *Francisella*, *Burkholderia* and *Helicobacter*.

crucial factor in the pathogenesis of chronic inflammatory, autoimmune and infectious diseases [51-53]. A number of studies also suggested a possible implication of TLR4 in cardiovascular disorders [54] and Alzheimer desease – associated pathology [55]. Therapeutic down-regulation of the TLR4 signaling is believed to be beneficial for treatment of numerous chronic and acute inflammatory diseases such as asthma [51], arthritis [52], influenza [50], and cancer [56]. Furthermore,

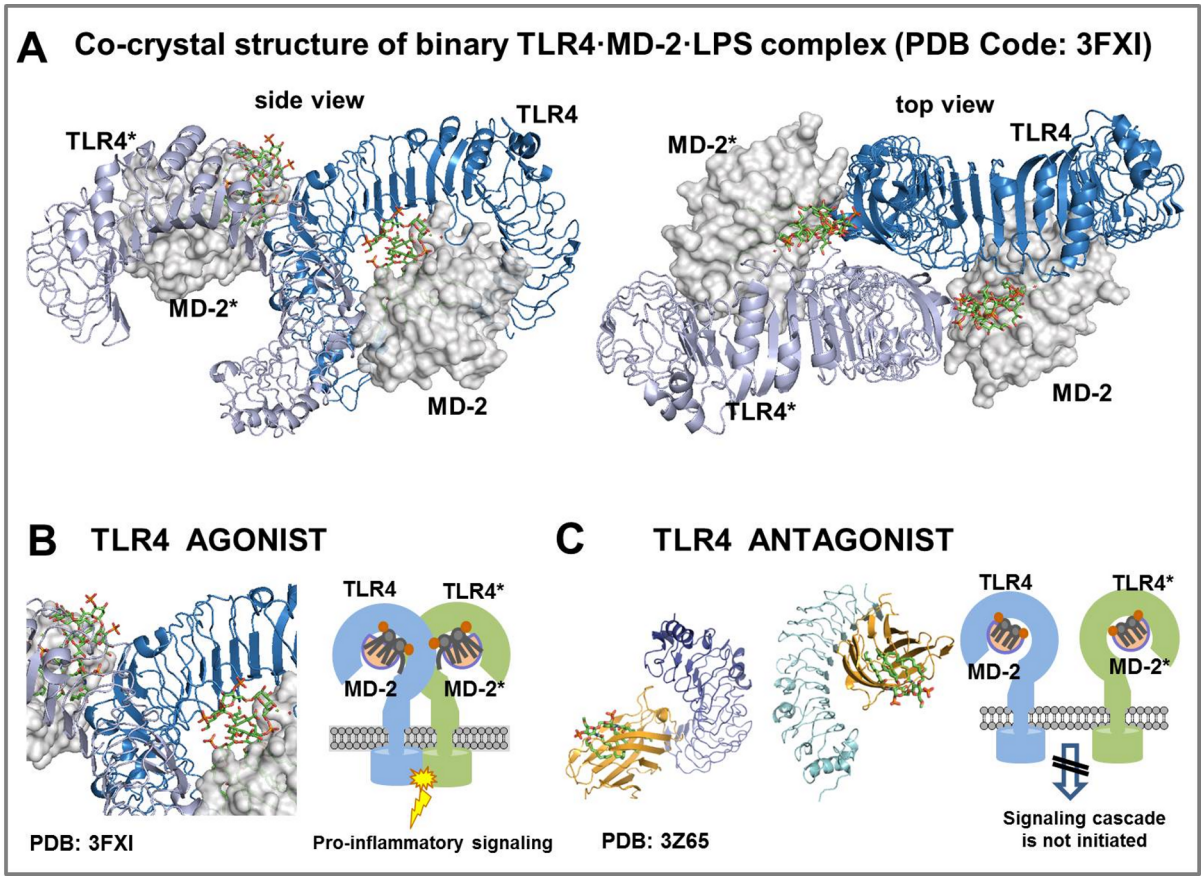


Figure 3: (A) Co-crystal structure of the homodimeric *E. coli* Ra-LPS·hMD-2·TLR4 complex (PDB code: 3FXI); (B) schematic representation of the *E. coli* lipid A induced activation of the MD-2·TLR4 complex (C) schematic representation of the interaction of TLR4 antagonist eritoran with MD-2·TLR4 complex. Images were generated with PyMol, ChemDraw and PowerPoint.

TLR4 has been shown to link the innate and adaptive immunity [57,58], underscoring stimulation of the TLR4·MD-2 complex by non-toxic TLR4-specific ligands as an apparent tactic for development of novel vaccine adjuvants [59-61].

X-ray structural analyses of the MD-2·TLR4 complexes with bound variably acylated lipid A uncovered markedly different modes of interaction of agonist and antagonist TLR4 ligands. Commonly, the binding of hexaacylated bisphosphorylated lipid A (such as lipid A from *E. coli*) by the TLR4·MD-2 complex results in an efficient activation of the innate immune response, while underacylated lipid A variants (such as tetraacylated lipid IVa [47], or a synthetic lipid A analogue eritoran) can block the endotoxic action of LPS [62,63]. All four acyl chains of antagonists eritoran and lipid IVa are fully inserted into the hydrophobic binding pocket of hMD-2 which results in an efficient binding without initiation of intracellular signaling (Figure 4A) [47,62]. In contrast, upon binding of hexaacylated *E. coli* LPS by the MD-2·TLR4 complex, only five long-chain acyl residues of lipid A are interpolated into the binding pocket of MD-2, whereas the sixth 2*N*-acyl lipid chain is exposed onto the surface of the co-receptor protein, constituting the core hydrophobic interface (together with the Phe126 loop of MD-2) for the interaction with the second TLR4*·MD-2*-LPS complex (Figure 4B) [42,64]. Thus, lipid A directly participates in the formation of an active multimeric ligand–receptor complex, whereas the tightness and efficiency of dimerization strongly depends on specific structural characteristics such as the acylation pattern and the number of negative charges (e.g., phosphate groups) in the molecule [65-67].

It has been just recently shown that TLR4 is not a sole receptor protein accountable for cellular responses induced by LPS. A number of pro-inflammatory effects such as autophagy, endocytosis and oxidative burst are induced by the LPS-mediated activation of an atypical inflammasome which is governed by the cytosolic enzyme caspase-11 and its human homologue caspase-4 [68]. Inflammasomes are protein complexes that are assembled in the cytosol of macrophages in response to the extracellular stimuli such as LPS [69]. The caspase-4/11 de-

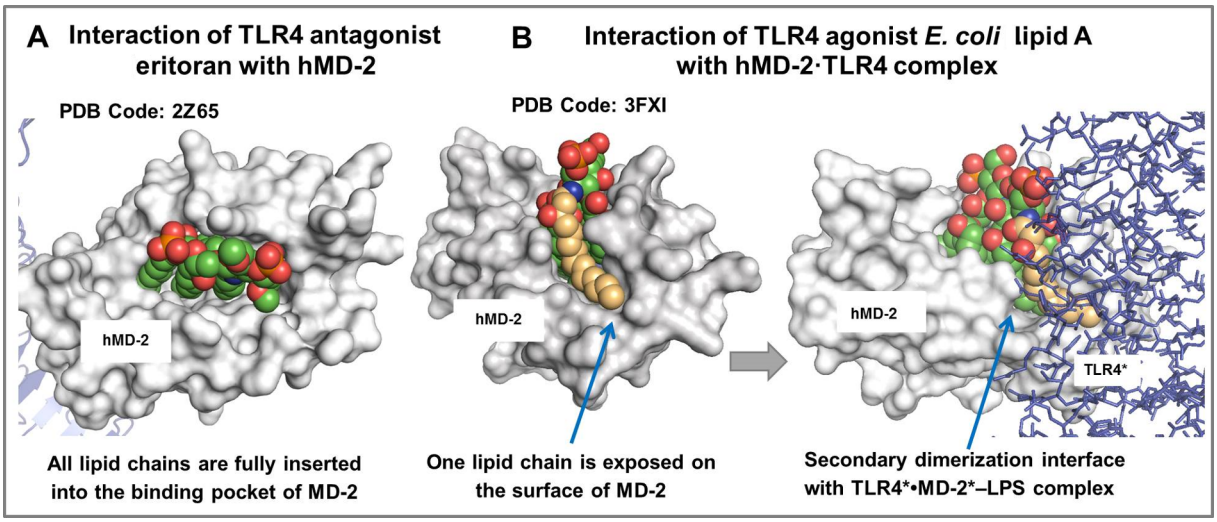


Figure 4: Co-crystal structures of (A) hybrid TLR4·hMD-2 with the bound antagonist eritoran (PDB: 2Z65, TLR4 is not shown); (B) homodimeric *E. coli Ra*-LPS·hMD-2·TLR4 complex (PDB code: 3FXI, TLR4 is not shown, only lipid A portion is shown for clarity). Images were generated with PyMol.

pendent inflammasomes are activated by the intracellular Gram-negative bacteria and largely contribute to development of endotoxic shock [70,71]. Biochemical studies revealed that caspase-4/11, which mediate inflammatory cell death by pyroptosis, are LPS receptors themselves [72,73].

Due to considerable micro-heterogeneity of the LPS isolates from wild-type or laboratory-adapted Gram-negative bacteria, the clinical and cellular studies as well as structure–activity relationship investigations using native LPS are complicated and difficult to evaluate. The lipid A content of LPS generally comprises a complex mixture of structural homologs having a variable number of the long-chain acyl residues of different chain lengths. The structural heterogeneity of lipid A preparations obtained through LPS isolation from bacterial cultures makes it difficult to get an unbiased correlation of specific structural features of lipid A and its TLR4-mediated activities. Moreover, possible contaminations with other pro-inflammatory bacterial components complicate the assessment of inflammatory pathways triggered by LPS in human and rodent immune cells. As example, not TLR4 but TLR2 (which mediates the host innate immune response to Gram-positive bacteria) was formerly reported to be responsible for the recognition of LPS belonging to certain bacterial strains. The micro-heterogeneity and contamination problem can be solved by application of synthetically prepared structurally defined lipid A variants of highest chemical and biological purity. To obtain clear structure–activity relationships data on lipid A–TLR4 interaction as well as unambiguous correlation of the lipid A acylation and phosphorylation pattern to its capacity in induction of different (i.e., MyD88-dependent and TRIF-dependent) signaling pathways, numerous well-defined lipid A substructures were synthesized. This review summarizes synthetic approaches developed in the past decade toward diverse LPS partial structures from different bacterial species including lipid A. The review provides comprehensive insight into the divergent and complex chemistry hidden under seemingly simple transformations needed for the assembly of lipid A, such as glycosylation towards fully orthogonally protected β(1→6)-linked diglucosamine backbone, sequential protective groups manipulation combined with successive instalment of multiple functional groups, *N*- and *O*-acylation with the long chain β-hydroxy fatty acids, anomeric phosphorylation and the synthesis of binary glycosyl phosphodiesters involving two amino sugars. Explicit structure–activity relationships data obtained with synthetic lipid A derivatives would also help to design novel therapeutic approaches for sepsis and inflammation.

Review

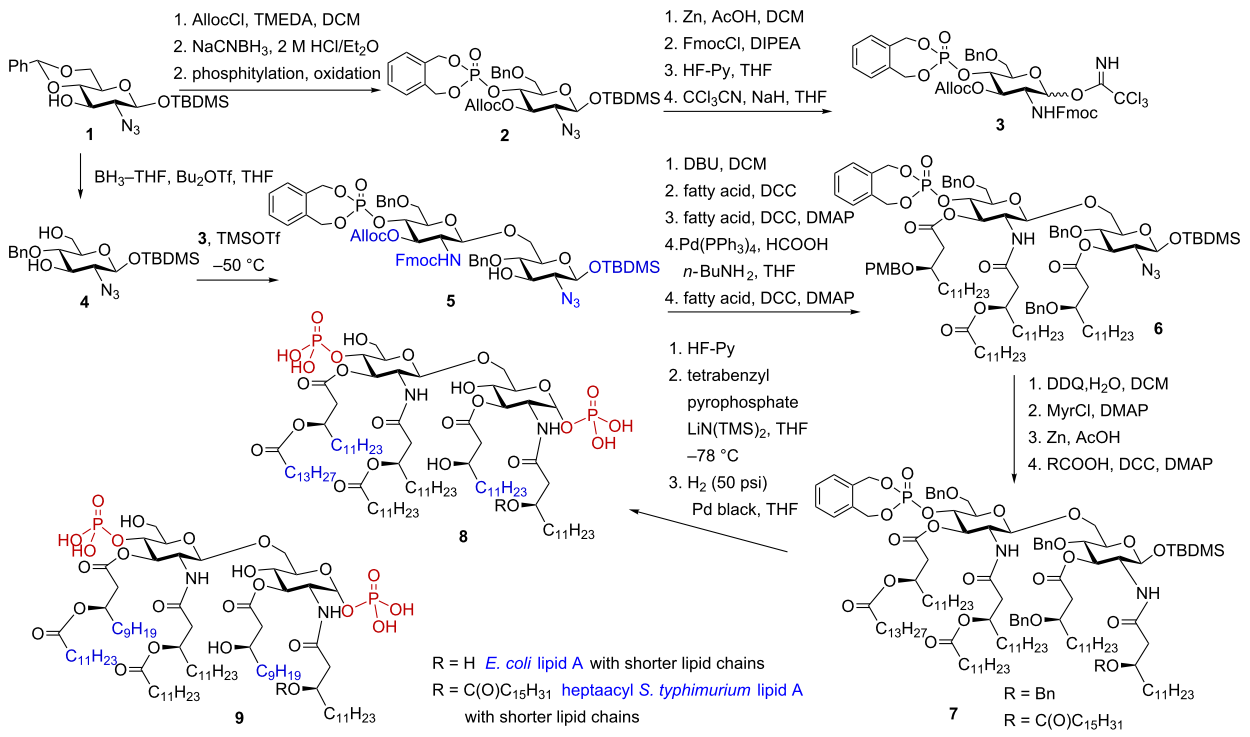
1. Synthesis of *E. coli, N. meningitidis, S. typhimurium* and *H. pylori* LPS partial structures comprising lipid A

1.1. Synthesis of *E. coli* and *S. typhimurium* lipid A

E. coli and *S. typhimurium* lipid A's count to the most powerful activators of the TLR4-mediated innate immune signaling and are responsible for the broad spectra of the inflammatory endotoxic effects in the infected host. To gain deeper insight into molecular basis of lipid A – TLR4 complex interaction and to determine the structural requirements for the efficient TLR4 activation, the hexaacylated lipid A corresponding to *E. coli* LPS, its analogue having 2 × CH_2 shorter acyl chains at positions 3

and 3’ as well as heptaacylated *S. typhimurium* lipid A and the corresponding analogue with shorter lipid chains at C-3 and C-3’ were synthesised via a highly convergent synthetic route [74]. In contrast to previously developed approaches which employed donor and acceptor monosaccharide molecules that were already functionalized with the lipid chains and phosphate groups [75,76], the new synthetic route used orthogonally protected monosaccharide precursors **3** and **4** (Scheme 1).

The glycosyl donor **3** was synthesised starting from azide **1** by first protecting the 3-OH group with an allyloxycarbonyl (Alloc) protecting group followed by regioselective reductive opening of the 4,6-*O*-benzylidene acetal using $NaCNBH_3$ and HCl in diethyl ether, and successive phosphitylation of the liberated 4’-OH functionality with *N,N*-diethylaminophosphepane (*N,N*-diethyl-1,5-dihydro-2,3,4-benzodioxaphosphepin-3-amine) in the presence of 1*H*-tetrazole followed by in situ oxidation with *m*-chloroperoxybenzoic acid (*m*-CPBA) to give fully protected 4’-phosphate **2**. The azido group in **2** was reduced, the resulting amine was converted to the *N*-Fmoc carbamate; the anomeric TBDMS ether was cleaved by treatment with HF in pyridine followed by reaction of the anomeric lactol with trichloroacetonitrile in the presence of a catalytic amount of NaH to provide trichloroacetimidate **3**. The glycosyl acceptor **4** was prepared from the same precursor **1** by regioselective reductive opening of benzylidene acetal using the borane−THF complex in the presence of Bu_2BOTf. Regioselective TMSOTf-catalysed glycosylation of the diol **4** by the imidate donor **3** resulted in the formation of a single product, the β(1→6)-linked disaccharide **5**. After the 2’-*N*-Fmoc group in **5** was removed with DBU to provide a free amino group, the 2’-NH_2 and 3-OH groups could be differentiated in the next acylation step by using DCC as activating agent for the *N*-acylation, and Steglich reaction conditions (DCC and DMAP) for the *O*-acylation. Following removal of the Alloc protecting group was readily performed by treatment with $Pd(PPh_3)_4$ in the presence of formic acid and butylamine to provide 3’-OH – containing precursor ready for the acylation by the long-chain acyloxyacyl acid. To avoid migration of the phosphotriester group from position 4’ to position 3’ and the formation of the acyloxy-chain elimination byproducts under DCC–DMAP-promoted acylation conditions, a two-step procedure for the acylation of 3’-OH group was applied. Acylation with the (*R*)-3-(*p*-methoxy)benzyloxytetradecanoic acid was initially performed to provide **6**, the (*p*-methoxy)benzyl ether was removed with DDQ and the liberated OH group was acylated with myristoyl chloride. Reduction of the 2-azido group by treatment with Zn in acetic acid followed by acylation of the amino group under standard conditions gave hexaacylated intermediate **7**.



Scheme 1: Synthesis of *E. coli* and *S. typhimurium* lipid A and analogues with shorter acyl chains.

The α-glycosyl phosphate was stereoselectively introduced by first, cleavage of the anomeric TBS ether by treatment with HF in pyridine, followed by phosphorylation using tetrabenzyl pyrophosphate in the presence of lithium bis(trimethyl)silylamide [76] in THF at −78 °C. Final deprotection by catalytic hydrogenolysis over Pd-black provided target lipid A derivatives **8** and **9** corresponding either to *E. coli* (R = H) and *S. typhimurium* (R = -C(O)$C_{15}H_{31}$) LPS with shorter acyl chains.

1.2. Synthesis of *N. meningitidis* LPS partial structures including lipid A

There has been significant controversy in reports concerning the induction of the pro-inflammatory responses by *N. meningitidis* LPS and the differentiation of the intracellular TLR4-mediated signaling pathways (MyD88 vs TRIF) by its lipid A compared to *E. coli* lipid A. Indeed, differences in the acylation pattern (non-symmetric [4 + 2] for *E. coli* and symmetric [3 + 3] for *N. meningitidis)* and the length of acyloxyacyl lipid chains substituting positions 2' and 3' of the diglucosamine backbone (shorter for lipid A of *N. meningitidis*) could be responsible for such discrepancy. However, significant heterogeneity of biological preparations used for cellular in vitro experiments as well as the presence of possible biologically active contaminations in the isolated samples put the consistency of immunobiological evaluation at risk. Moreover, to decipher the mode of interaction of LPS with the TLR4 system, the analysis of cytokine induction profile generated by meningococcal Kdo- (3-deoxy-D-manno-oct-2-ulosonic acid) lipid A compared to synthetic unsubstituted *N. meningitidis* lipid A was essential. To achieve these aims, a facile synthesis of meningococcal lipid A and Kdo-lipid A was elaborated. By the time the synthesis was performed, the crystal structure of the homodimeric TLR4·MD-2·LPS complex was not yet solved and the information on the biological activity obtained with synthetic molecules was fundamental for the understanding the structural basis of endotoxin-protein interaction.

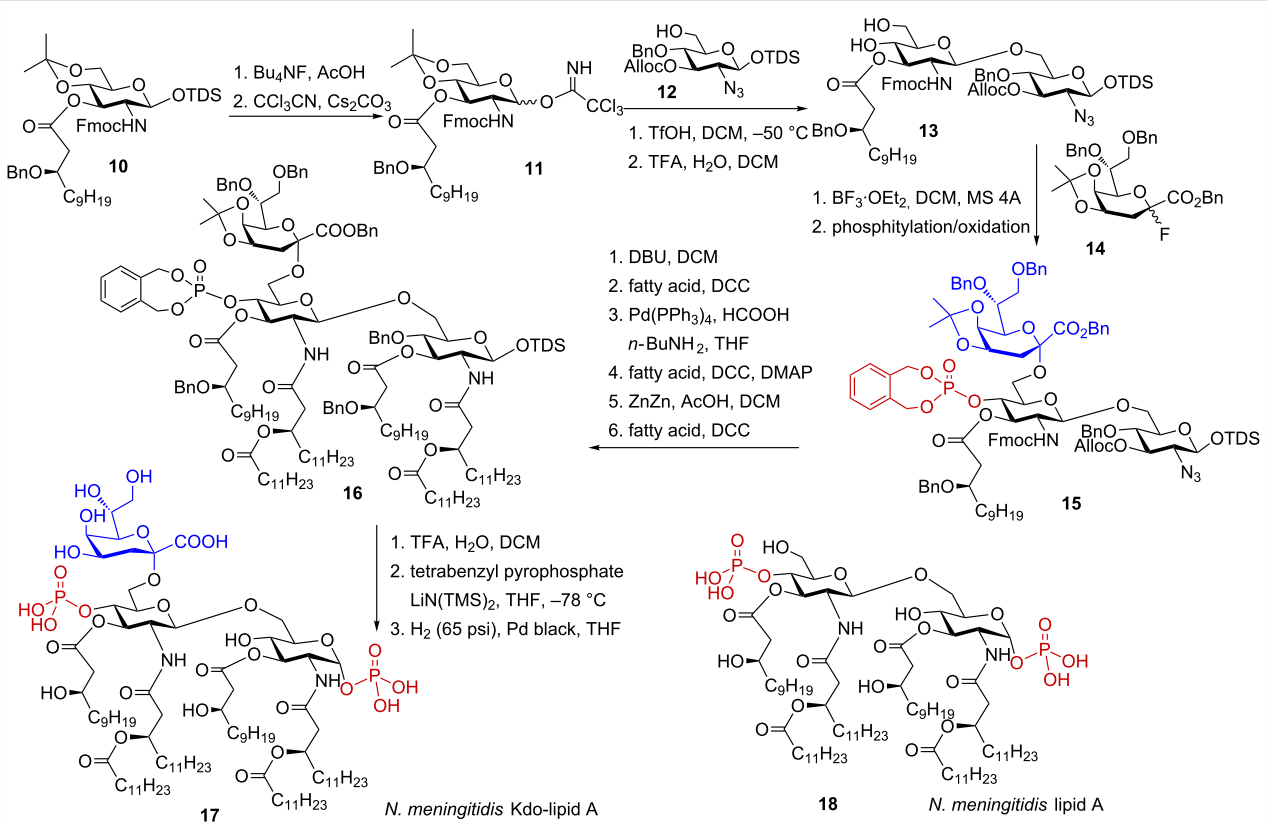
Preparation of Kdo-lipid A represents an even greater synthetic challenge than lipid A per se. The synthesis of *E. coli* type Kdo_2-lipid A (*Re*-LPS) was performed earlier [77] and was previously reviewed [76]. The synthesis of *N. meningitidis* Kdo-lipid A entailed initial preparation of donor and acceptor molecules constituting the diglucosamine backbone [78]. Accordingly, the *N*-Fmoc protected thexyldimethylsilyl (TDS) derivative **10** was anomerically deprotected by treatment with tetrabutylammonium fluoride buffered with acetic acid, and the resulting lactol was converted to the imidate donor **11** which was coupled to the orthogonally protected acceptor, an azide **12**, using triflic acid as promotor (Scheme 2). Subsequent hydrolytic cleavage of the isopropylidene group furnished diol **13**. Regioselective boron trifluoride diethyl etherate-promoted glycosylation of the 6-OH group in **13** with Kdo-fluoride donor **14** afforded an inseparable mixture of α- and β-anomeric products (α/β = 9:1) [78]. Phosphitylation of the remaining OH group in position 4' and facile separation of the anomeric α/β mixture furnished the anomerically pure trisaccharide **15**.

Next, three acyl residues were introduced at positions 2', 3' and 3 by successive deprotection–acylation sequence. The *N*-Fmoc protecting group was removed using DBU and the resulting free amino group was acylated with (*R*)-3-dodecanoyltetradecanoic acid in the presence of DCC as activating agent. Subsequently, the Alloc group was cleaved by treatment with $Pd(PPh_3)_4$ in the presence of $BuNH_2$ and HCOOH and the resulting 3-OH group was acylated using DCC in the presence of DMAP as activating agent. Succeeding reduction of the azido function with zinc in acetic acid followed by acylation of the liberated amino group with the long-chain acyloxyacyl fatty acid furnished fully acylated **16**. In the next steps, the isopropylidene acetal and anomeric TDS ether were removed by treatment with aqueous TFA and the anomeric hydroxyl group was regio- and stereoselectively phosphorylated using tetrabenzyl diphosphate in the presence of lithium bis(trimethylsilyl)amide [76] to provide glycosyl phosphotriester as exclusively α-anomer. Global deprotection was accomplished by catalytic hydrogenolysis over Pd-black to give meningococcal Kdo-lipid A **17**. A lipid A derivative **18** lacking Kdo residue at position 6' was prepared in a similar fashion.

Functional studies revealed that meningococcal Kdo-lipid A **17** was a much more potent inducer of the innate immune responses than lipid A **18** and stimulated the expression of TNF-α and IFN-β to a similar extent as its parent LPS. Thus, it could be confirmed, that lipid A having at least one Kdo residue attached at position 6' of the diglucosamine backbone represents the minimum structural requirement needed for the full activation of the LPS-sensing receptor TLR4. Comparison of activities of synthetic meningococcal and enteric lipid A revealed that the former was more potent in the induction of expression of the pro-inflammatory cytokines which could be attributed to the differences in the acylation pattern in both molecules. Importantly, it was demonstrated that neither of synthetic lipid A derivative had a bias towards MyD88- or TRIF-dependent immune responses [78].

1.3. Synthesis of fluorescent-labeled lipid A analogues

For studying the structural basis and the dynamics of TLR4-lipid A interplay, the application of labeled synthetic lipid A derivatives as versatile probes for tracking ligand–receptor interactions was exploited. However, the hydrophobic character and

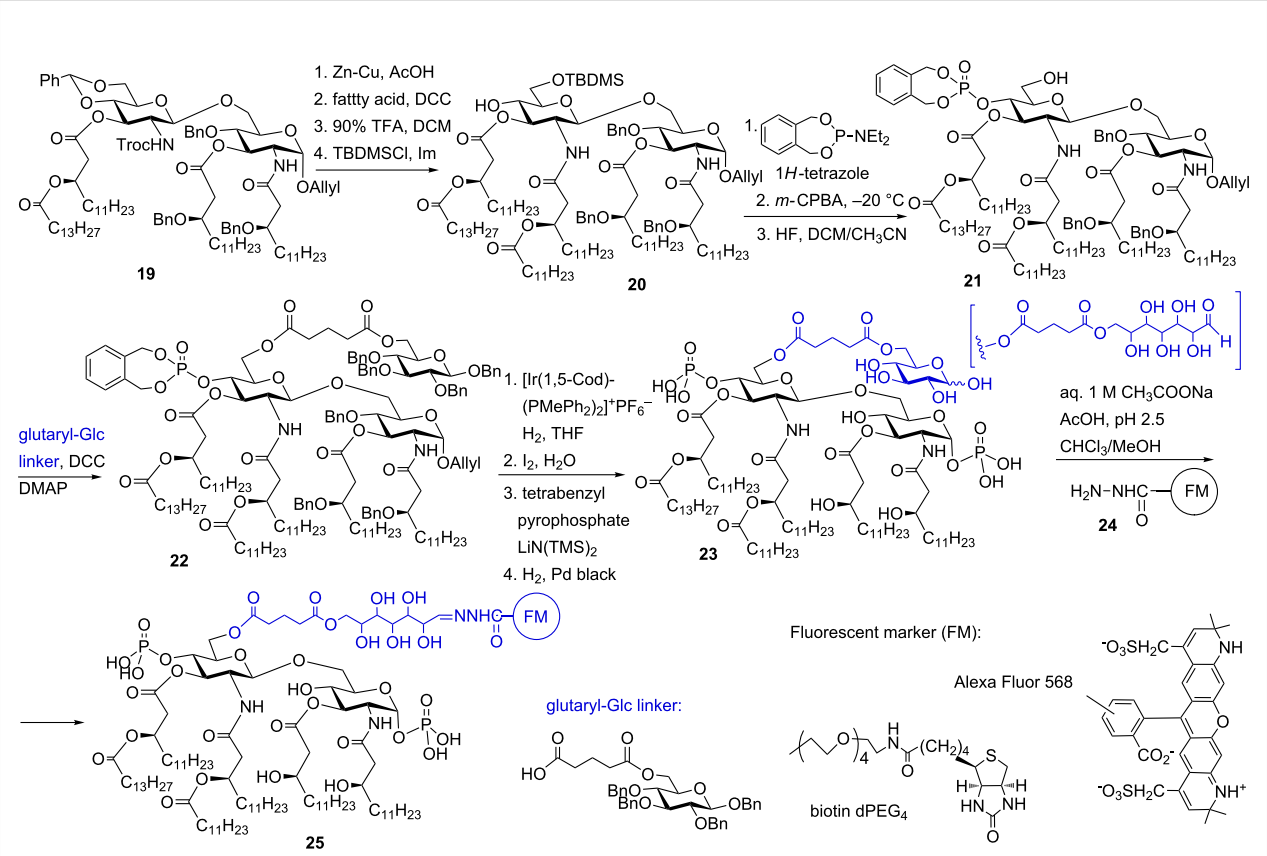


Scheme 2: Synthesis of *N. meningitidis* Kdo-lipid A.

the large size of most fluorescent labels which could potentially compete with lipid A for the LPS binding site at the TLR4 complex, posed an additional challenge. The only optional hydroxyl group which could qualify as the site of attachment of a fluorescent label without hindering the biological activity would be position 6’ of the diglucosamine backbone of lipid A. When attached to position 6’ via a linker molecule, the fluorescent label would not interfere with the binding of lipid A to the MD-2·TLR4 complex, such that the full TLR4-mediated activity would be preserved. Accordingly, the 6’-*O*-glycine-linked BODIPY (4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene)-labeled lipid A was previously synthesized [79]. However, this compound revealed only a weak fluorescence in aqueous solution owing to enhanced amphiphilicity of the hybrid molecule inflicted by the hydrophobic character of the fluorescent label and the formation of aggregates which resulted in self-quenching.

To circumvent these problems, a longer hydrophilic linker and a less hydrophobic fluorescent group were required. An elegant solution consisted in the application of glucose attached at position 6’ via a glutaryl group as a long-chain hydrophilic linker in combination with biotin or the hydrophilic fluorescent label AlexaFluor. The appropriately protected tetraacylated disaccharide **19** was subjected to treatment with Zn in AcOH which reductively cleaved the *N*-Troc group (Scheme 3). After *N*-acylation by (*R*)-3 acyloxyacyl fatty acid and hydrolytic cleavage of 4’,6’-*O*-benzylidene acetal group with 90% aqueous TFA, the liberated 6’-hydroxy group was regioselectively protected as TBDMS ether to furnish **20**. 1*H*-Tetrazole-catalysed phosphitylation of the 4’-OH group with *N,N*-diethylaminophosphepane followed by oxidation of the intermediate phosphite with *m*-CPBA to furnish the corresponding phosphate, and subsequent deprotection of the 6’-*O*-TBDMS ether gave the hexaacylated phosphotriester **21**.

The glutaryl-glucose linker (prepared from *O*-benzyl-protected glucose and glutaric anhydride) was introduced at the free 6’-OH group using DCC and DMAP to give **22**. The anomeric allyl group was cleaved by standard procedure, the phosphorylation of the 1-OH group was performed by 1-*O*-lithiation and subsequent treatment with tetrabenzyl pyrophosphate to furnish exclusively α-configured fully protected glycosyl phosphotriester. Global deprotection by catalytic hydrogenolysis over Pd-black gave *E. coli* lipid A functionalized with the glutaryl-Glc linker **23** which served as a key precursor for the preparation of fluorescent- or biotin-labeled compounds using labeling reagents having a hydrazide group.



Scheme 3: Synthesis of fluorescently labeled *E. coli* lipid A.

A hydrophilic fluorescence group Alexa Fluor 568 and polyethylene glycol-linked biotin were introduced using hydrazone formation reaction between the aldehyde group of the glutaryl-Glc linker and the hydrazide group of the labeling reagent. In addition to labeled *E. coli* type lipid A **25**, the labeled tetraacylated lipid IVa was also prepared. Importantly, the bioactivity of labeled compounds was fully preserved (the labeled *E. coli* type lipid A **25** performed as strong TLR4 agonist and the labeled tetraacylated lipid IVa acted, as expected, as TLR4 antagonist) and the fluorescence intensity of **25** and its tetraacylated counterpart was comparable with the fluorescence of the labeling reagent alone. Aggregation-mediated fluorescence quenching was not observed which confirmed the advantage of application of highly hydrophilic linker molecules and non-hydrophobic labeling reagents for amphiphilic glycoconjugates such as lipid A.

1.4. Synthesis of *Helicobacter pylori* Kdo-lipid A substructures

A *Helicobacter pylori* infection of the gastric mucosa causes chronic gastritis in humans and plays a pivotal role in the progression and pathogenesis of peptic ulcer diseases. Persistent infection with *H. pylori* is implicated in the development of gastric carcinoma [80]. *H. pylori* colonizes about 50% of the world's population and can asymptomatically persist for decades within a single host. The infection with *H. pylori* inevitably results in a chronic inflammatory response, whereas *H. pylori* LPS-dependent activation of monocytes and gastric epithelial cells leads to the production of several pro-inflammatory cytokines and reactive oxygen species (ROS) [81]. The mechanism by which *H. pylori* induces chronic inflammation and injury of gastric tissue is not fully understood. *H. pylori* produces a unique LPS molecule notable for strikingly low endotoxicity which is attributed to the structure of its lipid A moiety [81]. *H. pylori* uses two constitutive lipid A-mediated evasion strategies: repulsion of CAMPs (which are present at high concentrations in the gastric mucosa) and evasion of detection by the TLR4 system. Similarly to enteric *E. coli* LPS, *H. pylori* produces hexa-acylated lipid A, however, it displays a tetra- and triacylated lipid A molecule lacking the 4'-phosphate group on the bacterial surface [82,83]. Reduced number of acyl chains and the absence of the phosphate group at position 4' prevent detection of LPS by the TLR4. Thus, owing to post-translational modifications performed by several enzymes, the

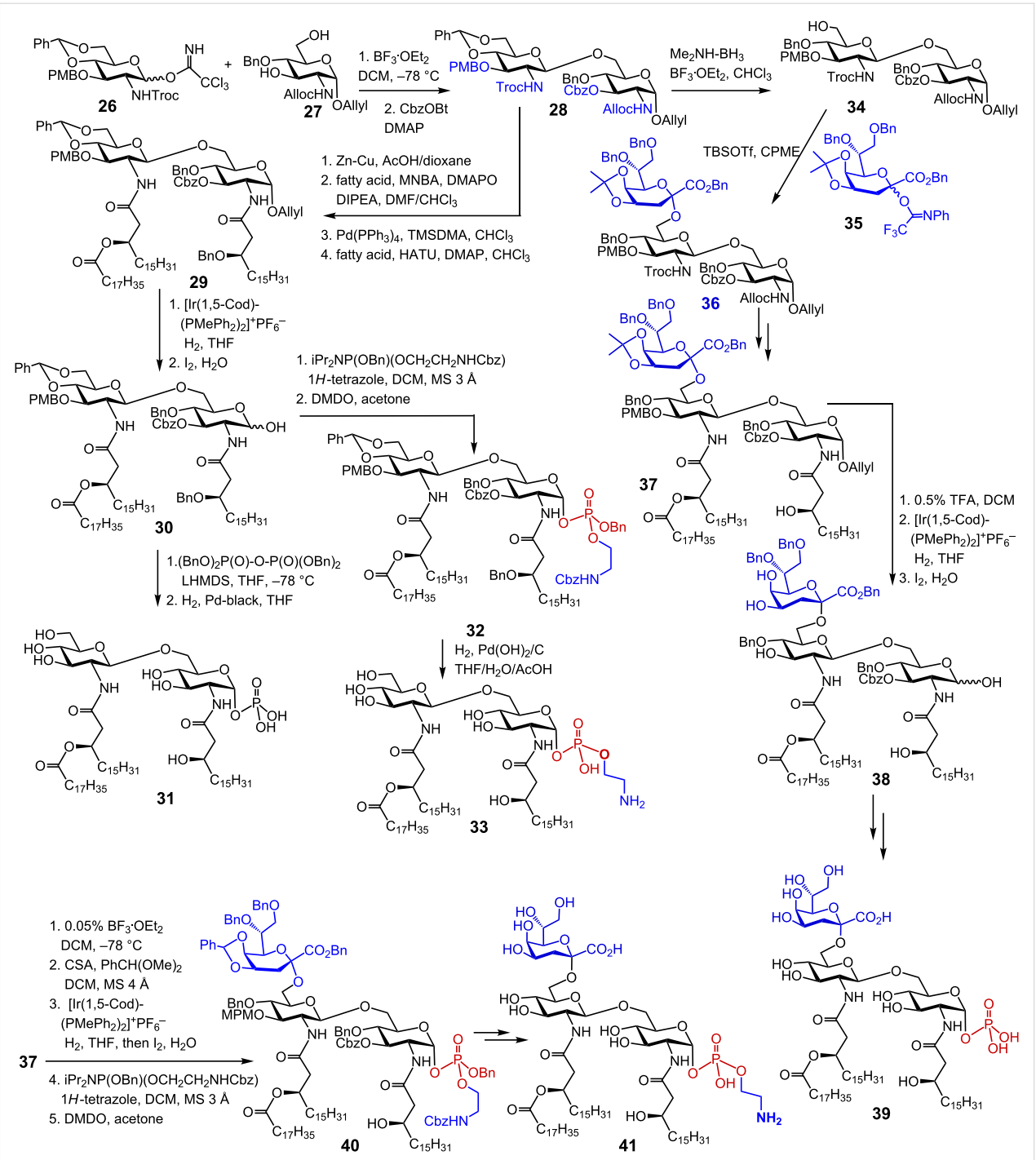
lipid A of *H. pylori* is poorly recognized by the innate immune system of the host [84]. The 1-phopshate group of *H. pylori* lipid A is further masked with ethanolamine that reduces the net negative charge and induces resistance to CAMPs (Figure 2). The unique structure of *H. pylori* lipid A plays a pivotal role in evading the host immune response by the bacterium [84]. Synthetically prepared structurally defined homogeneous *H. pylori* lipid A should help to identify the factors responsible for chronic inflammation during *H. pylori* infection.

The syntheses of *H. pylori* lipid A structures wherein the anomeric position was not modified with phosphoethanolamine were previously undertaken [85,86]. The syntheses of more sophisticated *H. pylori* lipid A substructures substituted by one Kdo residue at position 6' and/or modified with ethanolamine at the glycosidic phosphate were accomplished just recently [21,87,88]. The synthetic strategy relied on the initial preparation of fully orthogonally protected βGlcN(1→6)GlcN disaccharide which was then stepwise functionalized with a variable number of the long-chain (*R*)-3-acyloxy- and (*R*)-3-acyloxyacyl residues, 1-*O*-phosphate or 1-*O*-phosphoethanolamine groups and a 6'-linked Kdo moiety [21,88]. The synthesis commenced with the preparation of donor **26** and acceptor **27** molecules, which were coupled using $BF_3{\cdot}OEt_2$ as promotor to furnish fully protected β(1→6) diglucosamine (Scheme 4). Subsequently, the 3-OH functionality was protected with a carboxybenzyl group to give the key disaccharides **28**. The *N*-Troc group was reductively cleaved with Zn/Cu in acetic acid followed by acylation of the liberated 2'-amino group with the corresponding fatty acid using 2-methyl-6-nitrobenzoic anhydride (MNBA) as activating reagent in the presence of the nucleophilic catalysts 4-(dimethylamino)pyridine *N*-oxide (DMAPO) [89]. Next, the 2-*N*-Alloc group was cleaved by treatment with $Pd(PPh_3)_4$ and dimethylaminotrimethylsilane (TMSDMA) [90], followed by protection of the liberated 2-amino group by reaction with (*R*)-3-benzyloxycarboxylic acid using *O*-(7-azabenzotriazol-1-yl)-*N,N,N′,N′*-tetramethyluronium hexafluorophosphate (HATU) and DMAP as coupling reagents which furnished triacylated precursor **29**.

The 1-*O*-allyl group was then isomerized in the presence of an Ir complex and the resulting prop-1-enyl group was then removed by aqueous iodine to yield hemiacetal **30** which was stereoselectively phosphorylated by reaction with lithium hexamethyldisilazide (LHMDS), and subsequent treatment with tetrabenzyl pyrophosphate. Final deprotection by catalytic hydrogenation furnished lipid A **31**. Alternatively, the lactol **30** was phosphitylated by application of the phosphoramidite procedure with (benzyloxy)[(*N*-Cbz-3-aminopropyl)oxy](*N,N*-diisopropylamino)phosphine in the presence of 1*H*-tetrazole and subsequent oxidation with dimethyldioxirane (DMDO) [91] to furnish protected lipid A derivative **32**. Global deprotection by hydrogenation over $Pd(OH)_2/C$ in the presence of acetic acid afforded ethanolamine-modified *H. pylori* lipid A **33**.

To get deeper insight into the immunomodulatory potential of *H. pylori* lipid A, an access to synthetic *H. pylori* Kdo-lipid A was necessary. The presence of the Kdo moiety was shown to be decisive for the expression of full TLR4-mediated activity of lipid A. Previously, an efficient glycosylation strategy toward *E. coli* Kdo-lipid A using Kdo fluorides was developed by the same group. Glycosylation with Kdo fluoride required an excess of Lewis acid as promotor which was incompatible with the acid-labile protecting groups present in the key diglucosamine precursor. Therefore, a new *N*-phenyltrifluoroacetimidate Kdo donor **35** was developed (Scheme 4) [21]. The disaccharide acceptor **34** was prepared by regioselective reductive opening of 4′,6′-*O*-benzylidene acetal in **28** with $Me_2NH{\cdot}BH_3$ and $BF_3{\cdot}OEt_2$ in chloroform as solvent. The glycosylation of **34** with Kdo donor **35** was performed in CPME ether in the presence of TBSOTf as promotor to result in the stereoselective formation of trisaccharide **36**. Alternative microfluidic conditions applied by the authors ensured even better stereoselectivity and higher yields [21]. Sequential protective group manipulation and *N*-acylation procedure furnished the lipid A precursor **37**. The isopropylidene and anomeric allyl groups in **37** were removed and the anomeric position in **38** was regioselectively phosphorylated in a stereoselective manner by 1-*O*-lithiation with LHMDS, and subsequent treatment with tetrabenzyl pyrophosphate at −78 °C. Protecting groups were removed by hydrogenolysis on Pd-black to give *H. pylori* lipid A **39**. For the synthesis of Kdo-lipid A **41** entailing a phosphoethanolamine group at the anomeric position, the isopropylidene group in **37** had to be exchanged for the benzylidene group to avoid an application of acidic hydrolysis conditions for final deprotection of the labile glycosyl phosphodiester. After removal of the 1-*O*-allyl group using standard conditions, the anomeric lactol was phosphorylated via phosphoramidite procedure to furnish fully protected trisaccharide phosphodiester **40**, which was deprotected by hydrogenolysis on $Pd(OH)_2/C$ in $THF/H_2O/AcOH$ to give *H. pylori* lipid A **41**.

The availability of pure homogeneous synthetic compounds allowed for extensive immunobiological studies which revealed the unique functional properties of *H. pylori* lipid A. Triacylated lipid A variants efficiently inhibited the expression of IL-1β, IL-6 and IL-8 induced by *E. coli* LPS in human peripheral whole blood cells and the Kdo-containing lipid A substructures revealed the highest antagonist activity. On the other hand, all synthetic *H. pylori* lipid A and Kdo-lipid A showed IL-18 and IL-12 inducing activity, whereas the presence of Kdo decreased the potencies. Thus, it was shown that underacylated *H.*



Scheme 4: Synthesis of *H. pylori* lipid A and Kdo-lipid A.

pylori lipid A could disrupt the TLR4-mediated NF-κB signaling by inhibiting the LPS-triggered release of IL-6 and IL-8 and, at the same time, could activate other signaling pathways resulting in the induction of IL-12 and IL-18. This unique immunomodulating feature of *H. pylori* lipid A was linked to bacterial ability to dampen the acute immune reaction of the host and promote chronic inflammation.

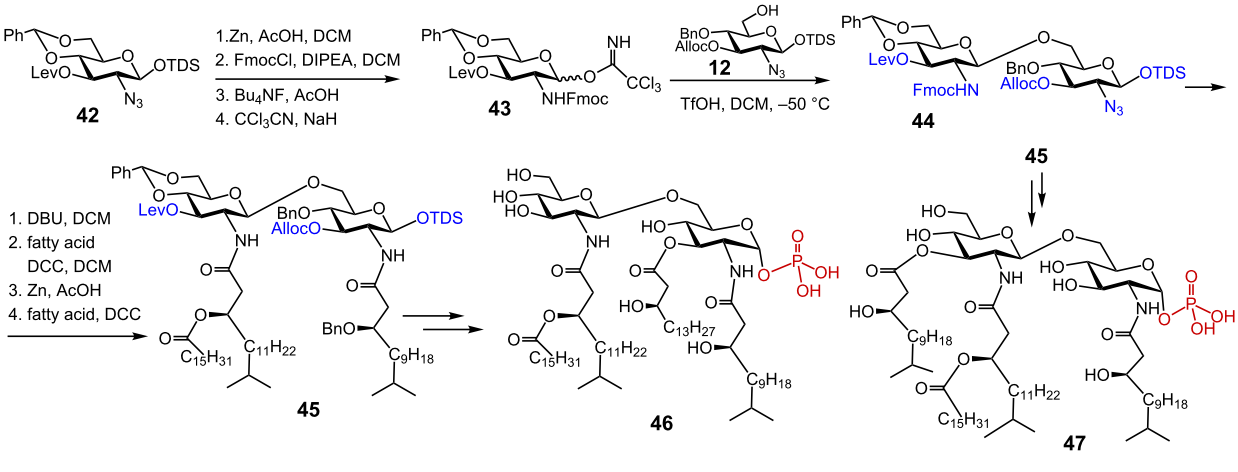
2. Synthesis of lipid A containing unusual lipid chains or lacking 1-phosphate group

2.1. Synthesis of variably acylated *Porphyromonas gingivalis* lipid A

Porphyromonas gingivalis is a major bacterial pathogen strongly implicated in periodontal disease (periodontitis) that is the primary cause of tooth loss in adults worldwide. Increasing

evidence suggest that *P. gingivalis* contributes to augmented systemic level of inflammation by invading the gingiva and modulating the innate inflammatory responses of the host which links periodontitis to various systemic diseases such as diabetes and cardiovascular disorders. The LPS of *P. gingivalis*, and particularly its lipid A, is recognized as major PAMP implicated in the pathogenesis of the periodontal disease. *P. gingivalis* LPS has been shown to stimulate the persistent production of IL-1, IL-6, and IL-8 in gingival fibroblasts which are thought to contribute to tissue destruction in gingivitis. On the other hand, it was demonstrated that *P. gingivalis* abolishes the expression of IL-8 in gingival epithelial cells which obstructs the host's capacity to recruit neutrophils to the sites of infection. Moreover, monocytes and human endothelial cells exhibit a low responsiveness to *P. gingivalis* LPS compared to *E. coli* LPS. *P. gingivalis* LPS was even shown to directly compete with *E. coli* LPS at the TLR4 complex in human endothelial cells, thus acting as TLR4-dependent antagonist of *E. coli* LPS. These discrepancies could be explained by a significant amount of structural heterogeneity displayed by *P. gingivalis* LPS containing both three-, tetra- and pentaacylated lipid A species [92]. The effects of these isoforms of *P. gingivalis* LPS on the expression of IL-6, IL-8 and TNF-α in human gingival fibroblasts are vastly diverse which contributes to periodontal pathogenesis [93,94]. Another structural peculiarity of the lipid A of *P. gingivalis* consists in the presence of the unusual branched fatty acid residues: *R*-(3)-hydroxy-13-methyltetradecanoate and *R*-(3)-hydroxy-15-methylhexadecanoate, which are non-symmetrically distributed across the diglucosamine backbone. Strong controversies in assessment of biological activities of *P. gingivalis* lipid A based on the LPS isolates [95-97] prompted chemical synthesis of structurally defined variably acylated *P. gingivalis* lipid A substructures [98,99].

Tetraacylated lipid A substructures representing the major lipid A of *P. gingivalis* were synthesised through a highly convergent approach employing a fully orthogonally protected key disaccharide **44** [98] (Scheme 5). A combination of temporary 3'-*O*-levulinoyl (Lev), 3-*O*-allyloxycarbonyl (Alloc) and 1-*O*-hexyldimethylsilyl (TDS) protecting groups with permanent benzyl/benzylidene acetal protections for hydroxyl groups and application of 9-fluorenylmethoxycarbamate (Fmoc) and azido protecting groups for masking the NH_2 functionalities allowed for the stepwise instalment of functional groups (phosphates and fatty acids) into the diglucosamine **44**. For the assembly of key disaccharide **44**, the azido group in **42** was exchanged for the *N*-Fmoc group by reduction with Zn in AcOH and reaction with FmocCl; anomeric TDS ether was cleaved and the resulting lactol was converted into the imidate donor **43** (Scheme 5). Glycosylation of the free 6-OH group in the acceptor azide **12** with the imidate donor **43** furnished fully orthogonally protected βGlcN(1→6)GlcN **44**. Next, the 2'-*N*-Fmoc group in **44** was removed by treatment with DBU and the first unusual branched acyloxyacyl residue was installed. For the preparation of (*R*)-3-hydroxy-13-methyltetradecanoic and (*R*)-3-hexadecanoyloxy-15-methylhexadecanoic acids an efficient cross-metathesis has been employed [98]. Reduction of the 2-azido group with Zn in acetic acid, followed by acylation with the respective 3-*O*-benzyl protected fatty acid provided the key intermediate **45**. Sequential protecting group manipulation (3'-*O*-Lev, 3-*O*-Alloc and 1-*O*-TDS) combined with acylation and regioselective anomeric phosphorylation furnished, after global deprotection, variably acylated *P. gingivalis* lipid A substructures **46** and **47**. The synthetic compounds did not stimulate the NF-κB signaling pathway, but efficiently inhibited the LPS-induced production of TNF-α in human monocytes. The acylation pattern was found to be decisive for the expression of the



Scheme 5: Synthesis of tetraacylated lipid A corresponding to *P. gingivalis* LPS.

antagonist activity since 2',3,2-triacylated lipid A **46** was a more potent antagonist than its 2',3',2-triacylated counterpart **47**.

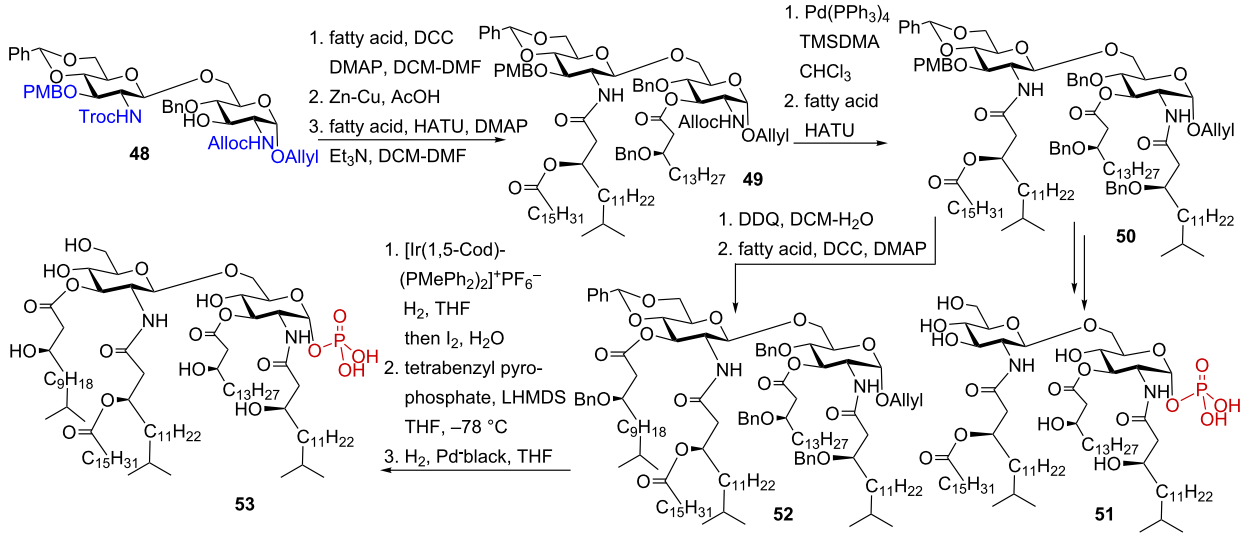
Synthesis of the *P. gingivalis* pentaacyl lipid A was based on the initial preparation of the orthogonally protected glucosamine disaccharide **48** [99]. Initial acylation of the free OH group in position 3, followed by sequential manipulation of the amino-protecting groups (2'-*N*-Troc and 2-*N*-Alloc) and acylation with the corresponding branched (*R*)-3-benzyloxyacyl and (*R*)-3-acyloxyacyl fatty acids furnished the lipid A precursor **50** (Scheme 6). Cleavage of the 1-*O*-allyl protecting group and stereoselective phosphorylation of the anomeric position via 1-*O*-lithiation with LHMDS, and subsequent treatment with tetrabenzyl pyrophosphate gave tetraacylated *P. gingivalis* lipid A **51**. For the synthesis of pentaacyl lipid A **53**, the 3'-*O*-*p*-methoxybenzyl group in **50** was cleaved by treatment with DDQ, and the liberated hydroxyl group was reacted with branched β-benzyloxy fatty acid to furnish fully acylated precursor **52**. After the cleavage of the 1-*O*-allyl group, the resulting lactol was phosphorylated to provide exclusively α-configured anomeric phosphotriester, which, after final deprotection by hydrogenolysis, gave pentaacyl lipid A **53**.

Immunobiological studies revealed that synthetic tri- and tetraacylated *P. gingivalis* lipid A substructures efficiently inhibited cytokine production induced by *E. coli* LPS, whereas the pentaacylated compound was less efficient in antagonizing LPS-mediated inflammatory responses. Interestingly, tetraacylated **51** selectively induced the expression of IL-18 which could be characteristic for LPS from bacteria causing asymptomatic chronic infection and persistent inflammation.

2.2. Synthesis of monophosphoryl lipid A (MPLA) as potential vaccine adjuvant

In contrast to the attenuated or whole killed vaccines which contain bacterial cell wall components and nucleic acids serving as naturally occurring adjuvants, the subunit vaccines lack these components. In the last decade much attention has been focused on the development of adjuvants that can render subunit vaccines more efficient by boosting the adaptive immune response. In this respect, TLR agonists deserved special consideration, since the induction of the innate immune signaling with PAMPs was shown to greatly enhance the adaptive immune responses [100].

Monophosphoryl lipid A (MPLA), an efficient and safe vaccine adjuvant registered for the use in Europe [59] is derived from the LPS of *Salmonella minnesota* R595 by following chemical modifications: elimination of the core oligosaccharide, hydrolysis of the 1-phosphate from the reducing end glucosamine, and removal of the acyl chain from position 3 of the disaccharide backbone [59]. Lower toxicity of the TLR4 ligand MPLA compared to its parent LPS/lipid A was linked to the absence of the phosphate group in position 1 of the diglucosamine backbone [101,102]. The absence of the 1-phosphate group on the MPLA molecule weakens the efficiency of the homodimerization of two TLR4·MD-2-ligand complexes which results in a weaker cytokine inducing capacity, diminished immune activation and lower endotoxic activity, while retaining immunogenicity [103]. MPLA differs from *E. coli* lipid A not only by the absence of the 1-phosphate group, but also in the acylation pattern. MPLA was reported to induce the innate immune response via a TRIF-mediated signaling pathway (in contrast to enteric lipid A which activates MyD88 pathway) [104]. A recent study demonstrated



Scheme 6: Synthesis of pentaacylated *P. gingivalis* lipid A.

that both TLR4 and MyD88 signaling have a significant effect on the adaptive immune responses in MPLA-adjuvanted vaccines [105]. To gain deeper understanding of the mechanisms underlying beneficial non-toxic immune response induced by MPLA and to reveal the major structural requirements responsible for adjuvant activity, monophosphoryl lipid A and several analogues differing in the acylation pattern have been synthetically prepared [106,107].

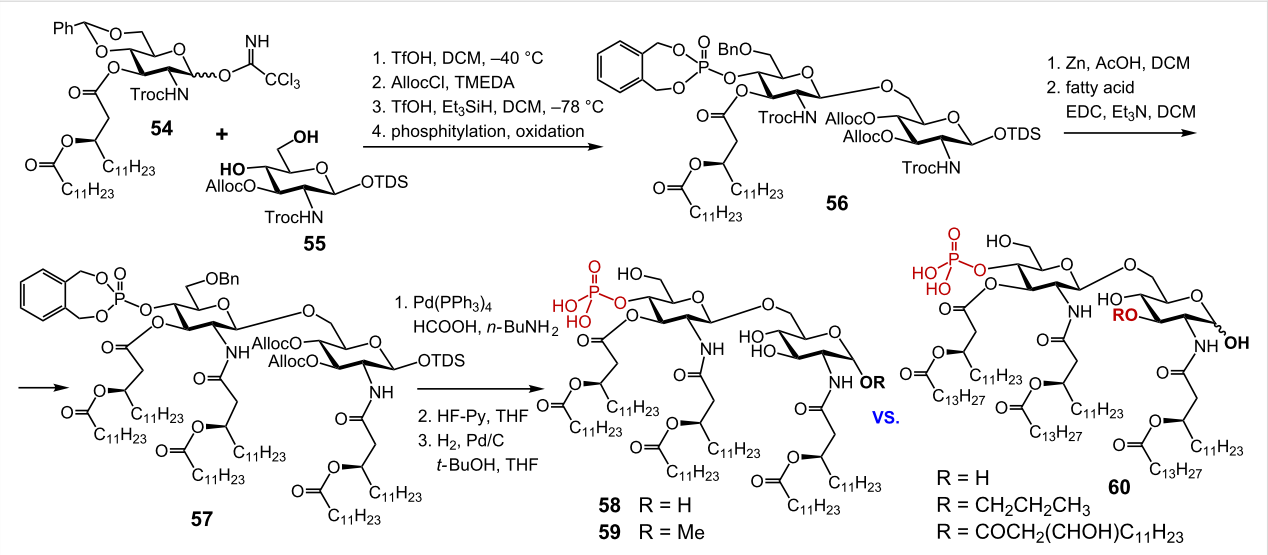
The synthesis of MPLA equipped with shorter secondary acyl chains **58** was achieved via regioselective glycosylation of the primary hydroxy group at position 6 in the *N*-Troc-protected acceptor **55** by the imidate donor **54** (Scheme 7) [106]. The (*R*)-3-dodecanoyloxytetradecanoyl residue was preinstalled in position 3 of the GlcN donor molecule. Acylation by an acyloxyacyl fatty acid at the latter stage of the synthesis could result in phosphate migration and/or elimination of the secondary acyl chain. TfOH-mediated 1,2-*trans* glycosylation smoothly provided β(1→6)-linked diglucosamine, the free OH group in position 3 was protected as Alloc carbonate and the benzylidene acetal protecting group was regioselectively reductively opened to furnish 6'-*O*-benzyl ether. The liberated 4'-OH group was phosphorylated via phosphoramidite procedure to furnish **56**. Next, both 2- and 2'-*N*-Troc groups were reductively cleaved using Zn in acetic acid and the resulting 2'- and 2-amino groups were acylated with (*R*)-3-dodecanoyltetradecanoic acid to give **57**. Three types of protecting groups – allyloxycarbonyl (Alloc), hexyldimethylsilyl (TDS) and benzyl – were sequentially removed to provide the target compound **58**. A monophosphoryl lipid A analogue **59** wherein the anomeric center of the proximal GlcN moiety is modified as methyl glycoside was prepared in a similar fashion.

It was expected that the small methyl group substituting the anomeric OH functionality would not compromise biological activity. Both MPLA analogues **58** and **59** were less efficient in eliciting TNF-α in mouse macrophages compared to a commercially available *S. minnesota* MPLA preparation, whereas methyl glycoside **59** showed somewhat higher pro-inflammatory activity. Interestingly, attachment of varying 3-*O*-substitutions at position 3 of the reducing GlcN moiety in MPLA analogue **60** did not enhance the adjuvant activity [107].

Importantly, synthetic MPLA derivatives having variable acylation pattern were successfully utilized as build-in-adjuvants in fully synthetic self-adjuvanting glycoconjugate cancer vaccines [108-110].

2.3. Synthesis of lipid A from *Rhizobium sin-1*

The *Rhizobiaceae* family refers collectively to the group of Gram-negative nitrogen-fixing plant endosymbiont bacteria. Lipid A of *Rhizobium* displays several significant structural differences when compared with *E. coli* lipid A: it lacks phosphate groups, but contains a galacturonic acid residue at the 4′-position and an aminogluconate moiety in place of the usual glucosamine 1-phosphate unit [111]. *Rhizobium* lipid A is esterified with a peculiar long chain fatty acid, 27-hydroxyoctacosanoate, which is not found in enteric Gram-negative bacteria [112]. The biosynthesis of lipid A in *R. leguminosarum* proceeds under the action of the same enzymes as in *E. coli* to generate the conserved phosphate containing precursor, Kdo_2-lipid IVa. Several additional enzymes, namely 1-phosphatase and 1-oxidase, catalyze further conversion of Kdo_2-lipid IVa into *R. leguminosarum* lipid A. The 1-phosphatase cleaves the 1-phosphate group to generate glucosamine which is subse-



Scheme 7: Synthesis of monophosphoryl lipid A (MPLA) and analogues.

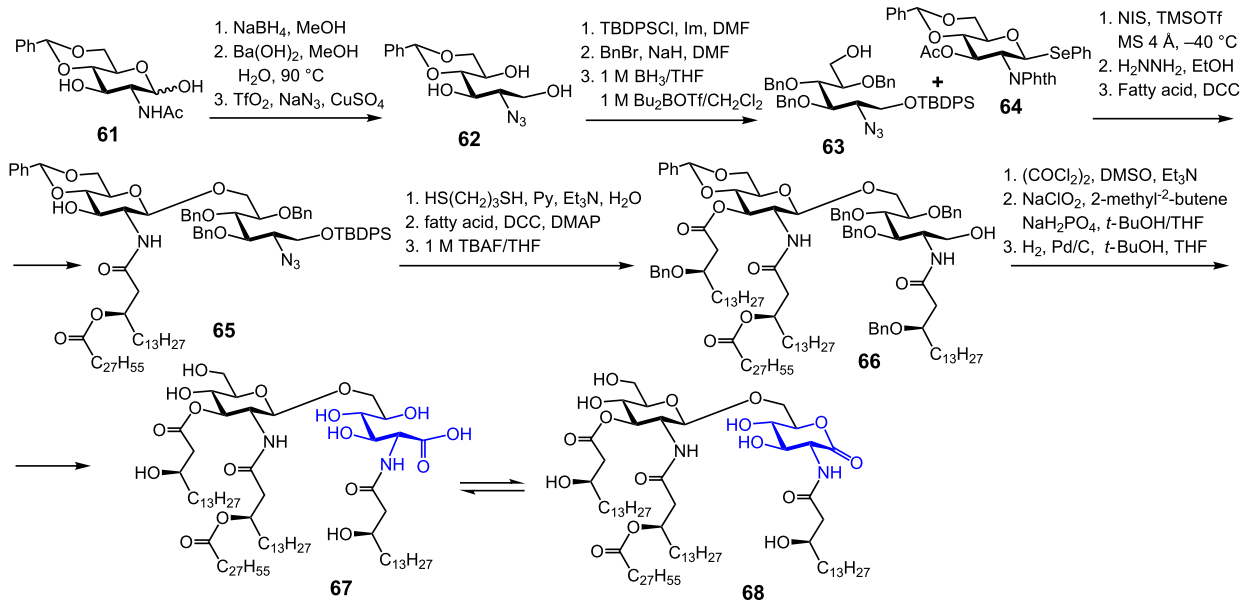
quently converted to 2-amino-2-deoxygluconate in an oxygen dependent manner via the action of an oxidase located in the outer membrane [113,114].

The unique *Rhizobium* lipid A lacks the structural features which are necessary for the TLR4-mediated stimulation of the innate immune system in animals. This might conceivably help bacteroids to evade the innate immune response in plants during symbiosis in root cells. Additionally, certain *Rhizobium sin-1* lipid A isolates were shown to inhibit the LPS induced toxic effects in human immune cells [115]. To determine the structural features which are responsible for the LPS antagonizing properties of the heterogeneous *Rhizobium sin-1* lipid A preparations, the synthesis of several defined *Rhizobium* lipid A structures has been undertaken [116-119].

A convergent synthetic approach towards *Rhizobium* lipid A substructures, 2-aminogluconate **67** and 2-aminogluconolactone **68**, included initial preparation of the alditol **62** (Scheme 8) [118]. To this end, GlcN hemiacetal **61** was reduced by treatment with $NaBH_4$, the acetamido group was removed with barium hydroxide, and the resulting amine was transformed into azide **62**. The primary alcohol in **62** was regioselectively protected as silyl ether, followed by benzylation and reductive opening of the benzylidene acetal to give the acceptor monosaccharide **63**. NIS/TMSOTf-promoted glycosylation of **63** with glycosyl donor **64** furnished desired β(1→6) disaccharide which was subjected to treatment with hydrazine hydrate to remove the phthalimido group. Subsequent acylation of the liberated NH_2 group provided **65**. A successive protective group manipulation/acylation sequence furnished tetraacylated **66**.

The oxidation of the primary alcohol in **66** to form the corresponding carboxylic acid was achieved by a two-step procedure involving oxidation under Swern conditions to give an intermediate aldehyde that was immediately subjected to a second oxidation with $NaClO_2$ and sodium dihydrogen phosphate to afford the 2-aminogluconate. In a final step, the benzyl ethers and the benzylidene acetal protecting group were removed by hydrogenolysis over Pd/C to give **67**. After the 2-aminogluconolactone **68** was separately synthesized, the NMR spectra of **67** and **68** were found to be identical indicating the co-existence of both forms in neutral conditions. Thus, it was demonstrated that *Rhizobium* lipid A exists in an equilibrium between open- and closed-ring forms, namely, as a mixture of 2-aminogluconate **67** and 2-aminogluconolactone **68**.

In an effort to develop more potent TLR4 antagonists, the synthesis of pentaacylated *R. sin-1* lipid A as well as its analogue modified by an ether-linked lipid chain in position 3 was undertaken [116,117]. High-yielding chemoselective coupling of the thioglycoside acceptor **69** with selenoglycoside donor **64** gave the disaccharide **70** (Scheme 9). Sequential removal of the amino-protecting groups (phthalimido group with ethylenedi-



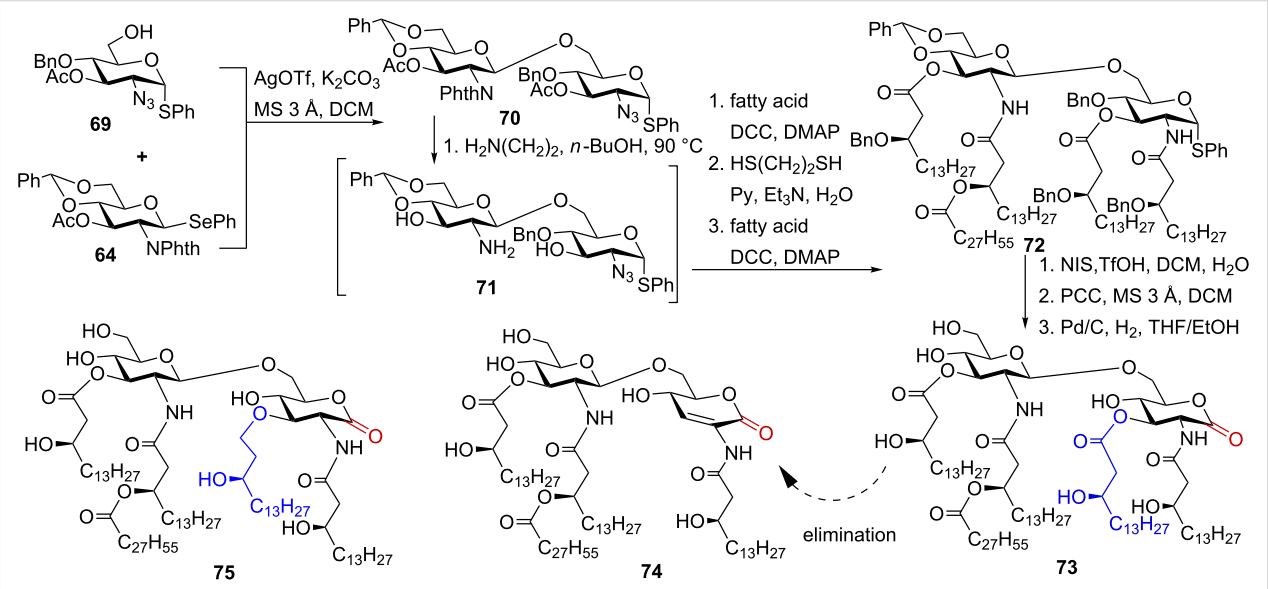
Scheme 8: Synthesis of tetraacylated *Rhizobium* lipid A containing aminogluconate moiety.

amine in refluxing butanol to furnish **71**, and the azido group by reduction with propane-1,3-dithiol) and subsequent acylation with respective fatty acids provided pentaacyl compound **72**. Hydrolysis of the thiophenyl moiety was performed by treatment with *N*-iodosuccinimide (NIS) and a catalytic amount of trifluoromethanesulfonic acid in wet dichloromethane, the benzyl ethers and benzylidene acetal were removed by catalytic hydrogenation on Pd/C to give *Rhizobium* lipid A **73**.

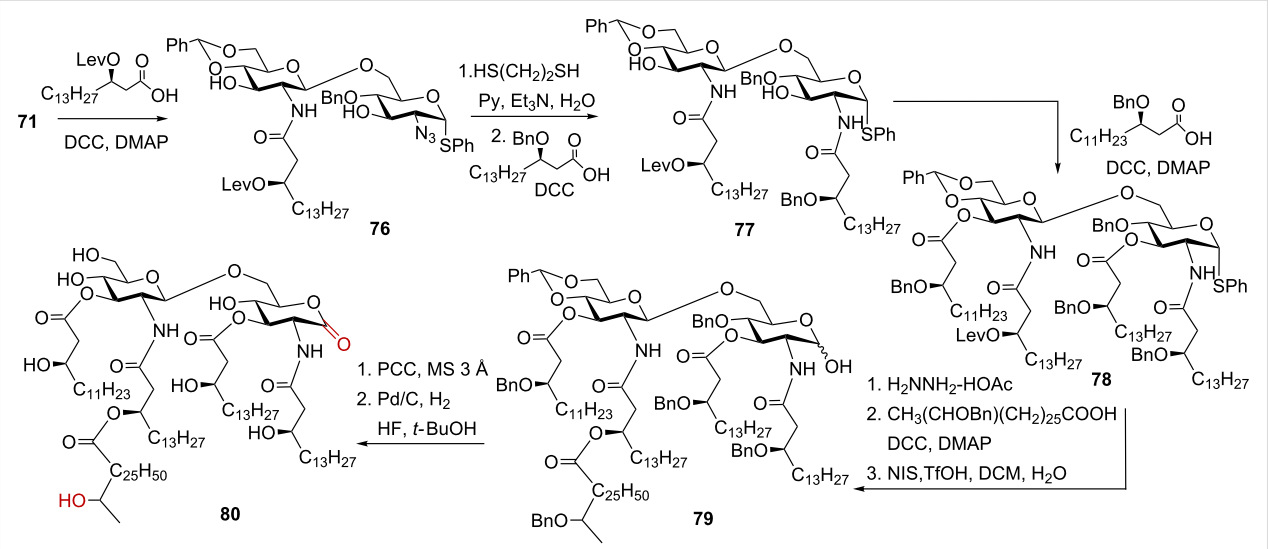
Biological evaluation of the synthetic *R. sin-1* lipid A **73** was complicated by its chemical lability owing to extensive elimination which gave the enone derivative **74**. To circumvent this problem, the β-hydroxy ester at C-3 of the proximal GlcN unit in **73** was replaced by an ether lipid chain to furnish *R. sin-1* lipid A analogue **75** [117].

Cellular activation studies revealed that synthetic *R. sin-1* lipid A was 100-fold less potent than its parent LPS in inducing TNF-α and IFN-β in murine macrophages. Interestingly, the difference in the TLR4 activation potencies between LPS and lipid A was much more pronounced for *E. coli* LPS (LPS was 10000-fold more active than the corresponding lipid A) than for *R. sin-1* LPS and lipid A (100-fold). No cytokine release was measured for 3-ether analogue **75**, however, **75** was nearly as active as **73** in inhibiting TNF-α and IP-10 production induced by *E. coli* LPS in human monocytes [117]. Thus, *R-sin 1* lipid A **73** and **75** antagonized the expression of cytokines resulting from both MyD88- and TRIF-dependent signaling pathways in human monocytic cell line indicating that the exchange of 3-ester linkage for the 3-ether linkage has only marginal impact on the TLR4 antagonizing activity. However, this difference exerted a dramatic effect on the species specific activation of cellular responses in murine macrophages wherein compound **73** induced the release of pro-inflammatory cytokines and the *R-sin 1* lipid A analogue **75** was inactive.

To determine the impact of hydroxylation of the long-chain 27-hydroxyoctacosanoic acid moiety for antagonist properties of *R-sin 1* lipid A, a lipid A containing this unique acyl residue was synthesised (Scheme 10). 27-Hydroxyoctacosanoic acid was prepared by employing a cross-metathesis between the ω-unsaturated ester and 3-butene-2-ol in the presence of Grubbs' second generation catalyst [119]. An appropriately protected disaccharide **71** having free amino group in position 2' was acylated by 3-*O*-levulinoyl protected (*R*)-3-hydroxyhexadecanoic acid [120] which, after the cleavage of levulinoyl protecting group, was esterified with benzyl ether protected 27-hydroxyoctacosanoic acid. Such a two-step approach facilitated the installment of the 27-hydroxyoctacosanoic residue into the lipid A moiety, and allowed for the synthesis of a series of differently acylated lipid A derivatives [119]. The azido group in monoacylated **76** was reduced with 1,3-propane dithiol, and the resulting amine was regioselectively acylated to give **77**. The free 3- and 3'-OH groups were acylated with (*R*)-3-benzyloxytetradecanoic acid under Steglich conditions to provide **78**, followed by cleavage of the levulinoyl ester and installment of the secondary ω-hydroxy acyl chain to furnish, after deprotection of the anomeric center, the hemiacetal **79**. The mixture of anomeric lactols was oxidized with pyridinium chlorochromate (PCC) to furnish the corresponding lactone, followed by hydrogenolysis on Pd/C to provide the target *R-sin 1* lipid A **80**.



Scheme 9: Synthesis of pentaacylated *Rhizobium* lipid A and its analogue containing ether chain.



Scheme 10: Synthesis of pentaacylated *Rhizobium* lipid A containing 27-hydroxyoctacosanoate lipid chain.

3. Synthesis of aminosugar modified lipid A: the assembly of binary glycosyl phosphodiesters

3.1. Synthetic challenges in the assembly of 1,1'-glycosyl phosphodiesters

Most naturally occurring glycosyl phosphodiesters entail the phosphoester linkage connecting one anomeric and one solely non-anomeric hydroxyl group. The assembly of such phosphodiesters is universally carried out using P(V)-based phosphotriester method, or P(III)-based phosphoramidite or H-phosphonate approaches [121-123]. In rare cases, however, the phosphodiester linkage can link the anomeric centers of two aminosugars as in the lipid A moieties of *Burkholderia, Bordetella* and *Francisella* LPS. The stereoselective assembly of 1,1′-glycosyl phosphodiesters represents a demanding synthetic challenge with respect to the necessity for the double anomeric stereocontrol and the inherent lability of the glycosyl phosphate intermediates. Generally, two major approaches can be applied for the synthesis of double glycosyl phosphodiesters, specifically, the phosphoramidite and the H-phosphonate procedures which are notorious for the mildness of the reaction conditions and the high reactivity of the P(III)-based intermediates. A three-coordinated phosphoramidite or a tetra-coordinated H-phosphonate species possess an electrophilic phosphorus centre which can instantly react with various nucleophiles. The benefits of the phosphoramidite methodology involve the mildness of the phosphitylation and oxidation conditions, while the chemical instability of the intermediary glycosyl phosphoramidites and glycosyl phosphites belongs to the drawbacks. For instance, isolation of the extraordinary labile glycosyl phosphoramidite intermediates in anomerically pure form looks rather unfeasible. The benefits of the H-phosphonate procedure rely on the stability of the glycosyl H-phosphonate monoesters which can be readily isolated by silica gel column chromatography, as well as on the absence of a protecting group at the phosphorus atom. Yet, the classic pivaloyl chloride (PivCl)-mediated H-phosphonate coupling reaction can result in the formation of a number of byproducts, and in the hydrolysis of the target 1,1′-glycosyl phosphodiester upon harsh conditions of aqueous iodine-mediated oxidation of the intermediate P(III) H-phosphonate phosphodiesters into the P(V) species. Fortunately, expedient modification of the H-phosphonate technique in terms of application of alternative coupling and oxidative reagents renders it to the method of choice for the assembly of binary glycosyl phosphodiesters.

3.2. Synthesis of partial structure of galactosamine-modified *Francisella* lipid A and a neoglycoconjugate based thereof

Francisella is a highly infectious Gram-negative zoonotic bacterium and the causative agent of tularemia, an extremely contagious lethal pulmonary disease in mammals [124]. Despite clinical and biosecurity importance (*F. tularensis* is classified as a bioterrorism agent [125]), the molecular basis for the pathogenesis of a *F. tularensis* infection remains largely unknown. The major lipid A of *Francisella* has an unusual tetraacylated structure composed of a common β(1→6)-linked diglucosamine backbone which lacks the 4′-phosphate group and the 3′-acyl chain characteristic for enteric lipid A; and contains an α-D-GalN residue that is glycosidically linked to the 1-phosphate group [126]. *Francisella* LPS does not trigger the pro-inflammatory signaling cascade since it cannot be recognised by the TLR4·MD-2 complex owing to the hypoacylated structure of its lipid A and the absence of the 4′-phosphate group [127].

Posttranslational modification of the anomeric phosphate group of lipid A in *Francisella* with α-GalN confers resistance to CAMPs and is associated with augmentation of bacterial virulence [26,128-130]. The full biological consequence of the GalN modification in *Francisella* lipid A is still poorly understood, although it was shown that *F. novicida* mutants which are deficient in GalN modification have attenuated pathogenicity in mice and are capable of stimulating the innate immune response [131].

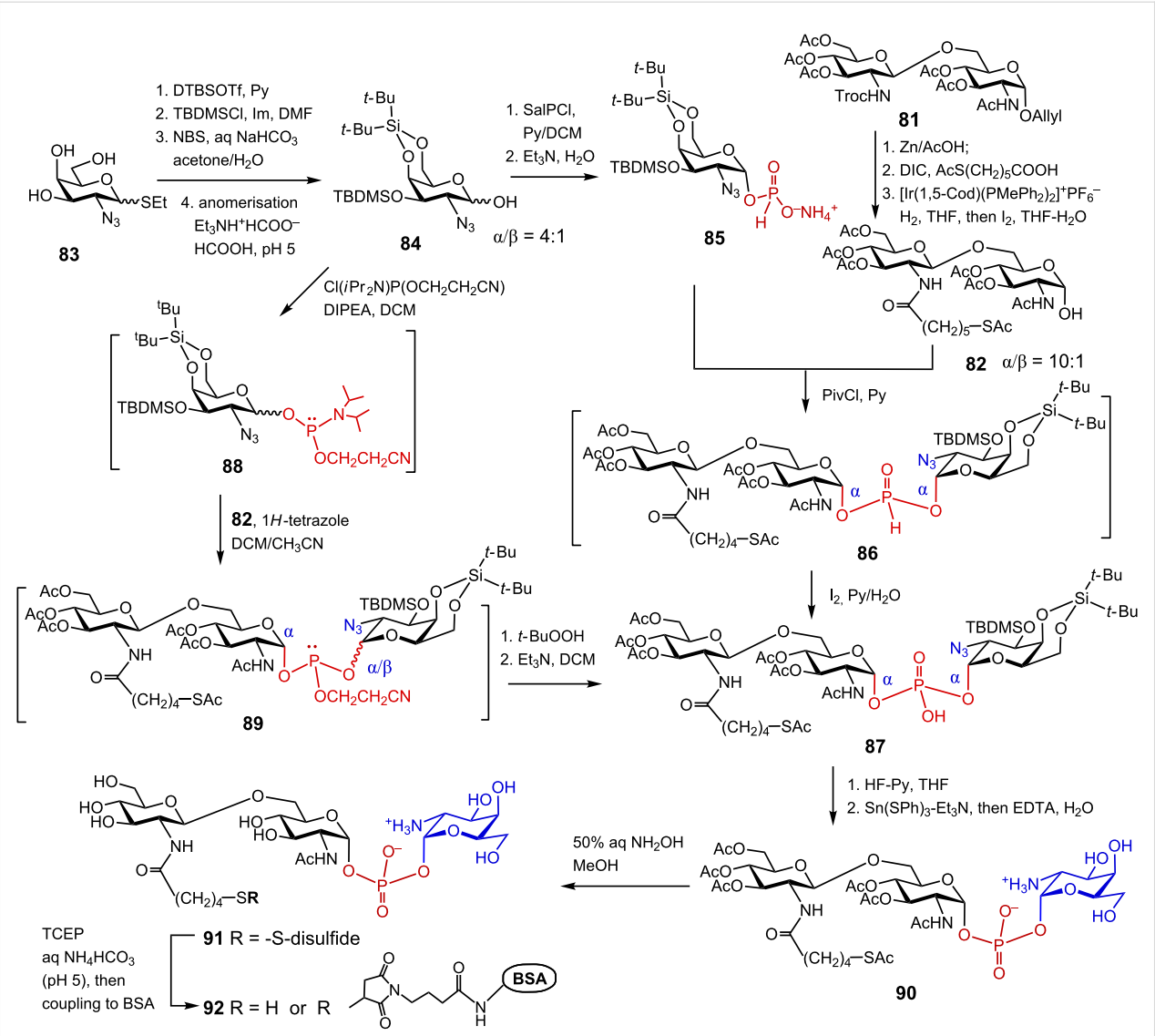
As a consequence of a unique system of the LPS remodelling enzymes [132-134], *Francisella* produces truncated LPS structure which is composed to 90% from a lipid A portion alone and is not substituted by the core sugars and polymeric O-antigen [126,135]. In this instance, the diglucosamine backbone of *Francisella* lipid A modified by α-D-GalN at the glycosidic phosphate group comprises the antigen-presenting entity of *Francisella* LPS. To assess the antigenic potential of the GalN modification in *Francisella* lipid A, a lipid A-based epitope βGlcN(1→6)-αGlcN(1→P←1)-αGalN **91**, which is conserved in all *Francisella* strains, and a corresponding neoglycoconjugate **92** were synthesised [136]. These compounds could be applied for the generation of diagnostic antibodies or utilized in immunoaffinity assays for detection of *Francisella* infection by direct antigen manifestation in clinical samples [137].

The β(1→6)-linked diglucosamine **81** was prepared by a TMSOTf-assisted glycosylation of the allyl glycoside of the per-acetylated GlcN acceptor having a free 6-OH group by the 2*N*-Troc protected GlcN-based trichloroacetimidate donor [136]. Reductive cleavage of the 2′*N*-Troc protecting group followed by *N,N*′-diisopropylcarbodiimide (DIC)-mediated acylation with 6-thioacetylhexanoic acid afforded a desired β(1→6)-linked disaccharide equipped with a masked spacer group. Cleavage of the 1-*O*-Allyl group by first isomerization to a propenyl group and subsequent aqueous I_2-mediated hydrolysis provided anomeric α-lactol **82** (α/β = 10:1) entailing an acetyl-protected sulfhydryl-containing spacer (Scheme 11).

For the synthesis of the *Francisella* lipid A backbone having a unique structure which encloses a double glycosyl phosphodiester functionality linking the anomeric centers of two aminosugars, the expediency of the H-phosphonate and phosphoramidite approaches was explored [136]. The synthesis of anomerically pure α-GalN-derived H-phosphonate **85** was performed via regioselective instalment of the 4,6-*O*-*tert*-butylsilylene (DTBS) group into the triol **83**, followed by reaction of the free 3-OH group with TBDMS chloride in the presence of imidazole to furnish a fully protected GalN derivative (Scheme 11). The latter was anomerically deprotected via *N*-bromosuccinimide (NBS)-mediated hydrolysis of the thioethyl glycoside to furnish hemiacetal **84**. The DTBS group exerted a remote α-directing effect [138] which facilitated an enhancement of the α/β ratio in the anomeric lactol **84**. The orthogonally protected GalN hemiacetal **84** (α/β = 3:1) was subjected to phosphitylation reaction with 2-chloro-1,3,2-benzodioxaphosphorin-4-one (salicylchlorophosphite, SalPCl) [139,140]. Since the stereoselectivity of phosphitylation by the P(III)-based reagents commonly reflects the α/β ratio in the starting hemiacetal, the proportion of the α-configured lactol in **84** was additionally enhanced by in situ anomerisation with triethylammonium formate–formic acid buffer (pH 5). The reaction of **84** (α/β = 4:1) with SalPCl in the presence of pyridine afforded glycosyl H-phosphonate **85** which was isolated in pure α-anomeric form as ammonium salt [136]. A pivaloyl chloride (PivCl)-mediated coupling of the H-phosphonate **85** and peracetylated β(1→6) diglucosamine hemiacetal **82** furnished double glycosyl H-phosphonate diester **86**. Oxidation of the intermediate H-phosphonate diester **86** with aqueous I_2 afforded anomerically pure binary glycosyl phosphodiester **87** entailing αGlcN(1→P←1)αGalN fragment. Application of a nearly pure α-anomeric form of the diglucosamine lactol **82** (α/β = 10:1) and high efficiency of the H-phosphonate coupling allowed for a highly pleasing 85% yield of the glycosyl phosphodiester **87**.

To explore the applicability of the phosphoramidite procedure, the anomeric *N,N*-diisopropyl-2-cyanoethyl phosphoramidite **88** was prepared in situ by treatment of GalN hemiacetal **84** with *N,N*-diisopropyl-2-cyanoethylchlorophosphite in the presence of DIPEA [141]. 1*H*-Tetrazole-mediated coupling of the latter to lactol **82** (α/β = 10:1) afforded a mixture of the intermediate anomeric phosphite triesters **89**. After oxidation with *tert*-butylhydroperoxide and treatment with Et_3N to remove the cyanoethyl protecting group from the phosphotriester by β-elimination, the target phosphodiester **87** was obtained in a 24% yield. Due to the intrinsic lability of the glycosyl phosphoramidite and glycosyl phosphite intermediates, four sequential transformations were performed as "one-pot" procedure without isolation of individual anomers which ultimately resulted in a poor overall yield.

The progress of a phosphorylation reaction involving phosphorus P(III)-intermediates can be easily monitored by ^{31}P NMR spectroscopy. Thus, the H-phosphonate monoester like **85** usually displays a doublet at δ: 4–8 ppm with the coupling constant $^2J_{PH}$ = 630–650 Hz. After the coupling reaction of the H-phosphonate with the nucleophilic component (hemiacetal **82**), the H-phosphonate diester **86** is expected to have a slightly downfield ^{31}P NMR shift δ: 6–12 ppm and a larger coupling constant of $^2J_{PH}$ = 730–750 Hz. As soon as the H-phosphonate **86** is oxidised to furnish a P(V) phosphodiester **87**, the phosphorus chemical shift usually appears at around δ: 0 ppm.



Scheme 11: Synthesis of zwitterionic 1,1′-glycosyl phosphodiester: a partial structure of GalN-modified *Francisella* lipid A and a neoglycoconjugate based thereof.

The phosphoramidites like **88** have a characteristic ^{31}P NMR chemical shift δ: 150 ppm (two signals corresponding to the *R*- and *S*-diastereomers at phosphorus), whereas the phosphite triesters like **89** display two ^{31}P NMR resonances (R_p- and S_p-diastereomers) at δ: 138–142 ppm.

Sequential deprotection of **87** had to be performed under explicitly mild reaction conditions to avoid hydrolysis of the labile double glycosyl phosphodiester functionality. The desilylation of the GalN moiety was accomplished by treatment with diluted HF·Py solution which furnished the corresponding triol. The presence of the terminal thiol precluded application of the Pd-catalysed hydrogenation for the reduction of azido group, so that the Staudinger reaction conditions (using PPh_3 or PMe_3) in THF/aq NaOH [142] were initially attempted. The Staudinger reaction did not result in a desired transformation and the alternative procedures for the reduction of azido group were investigated. The best results were achieved upon application of the tin(II) complex $[Et_3NH][Sn(SPh)_3]$ [143,144] which quantitatively reduced the 2-azido group in the GalN moiety to yield zwitterionic compound **90**. The use of an excess of the tin(II) reagent caused partial hydrolysis of the GalN fragment in the phosphodiester **90**, unless the tin(II) reagent was trapped by a chelating agent, ethylenediaminetetraacetic acid (EDTA) immediately after the reduction was completed. Final deacetylation was performed under mild basic conditions to afford a zwitterionic phosphodiester **91**. After reduction of the disulfide bond in **91** with tris(2-carboxyethyl)phosphine (TCEP) [145], the result-

ing thiol was coupled to a maleimide-activated BSA which provided βGlcN(1→6)-αGlcN(1→P←1)-αGalN containing neoglycoconjugate **92**. The epitope can be potentially attached to different surfaces via its thiol-terminated spacer and utilized in diagnostic immuno-assays as capture antigen.

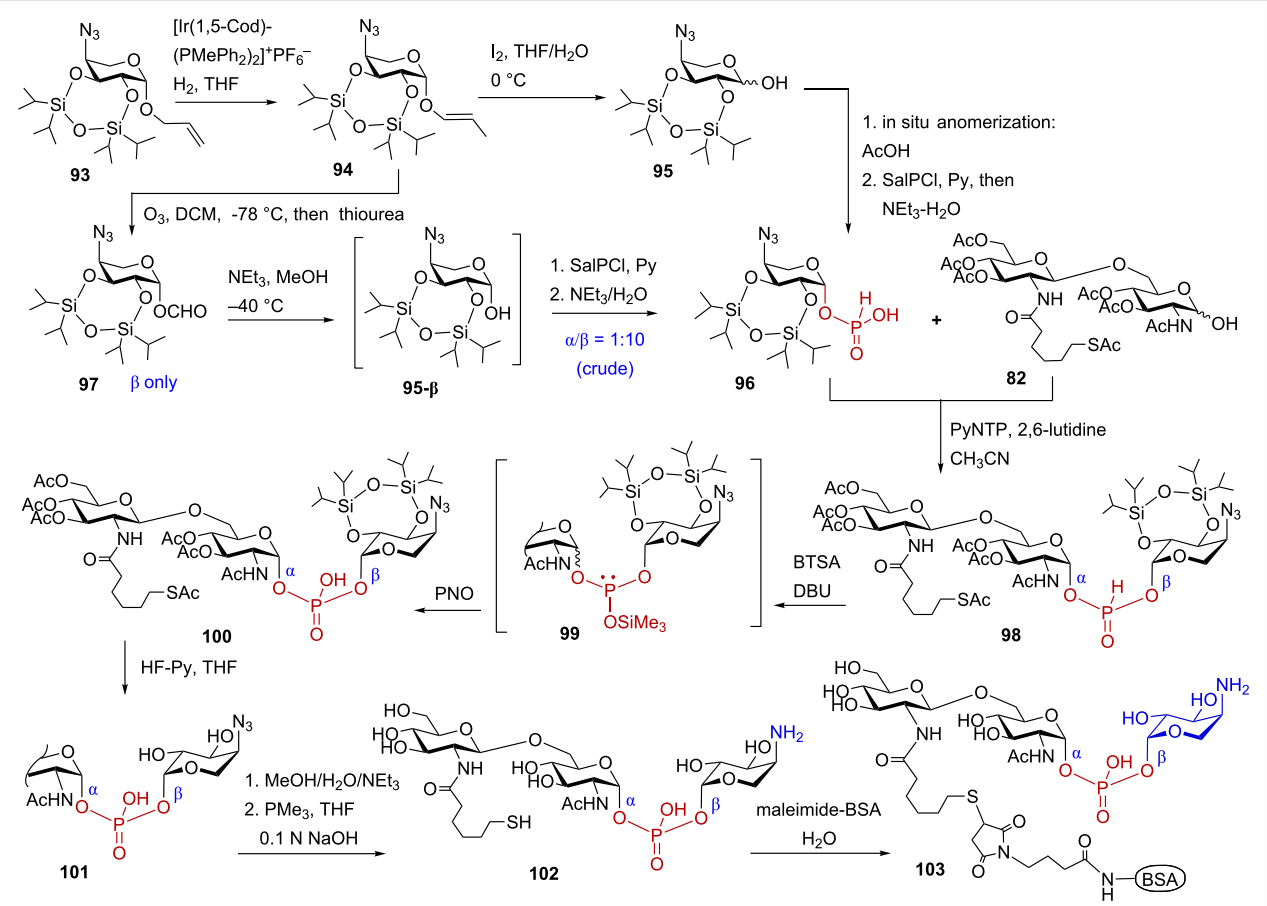
3.3. Synthesis of double glycosyl phosphodiester comprising 4-amino-4-deoxy-β-L-arabinose (β-L-Ara4N) – a partial structure of *Burkholderia* LPS

The *B. cepacia* complex (BCC) is a group of opportunistic bacterial species that can cause lethal pneumonia and septicaemia in patients with cystic fibrosis (CF) and immunocompromised patients resulting in exceptionally high mortality („the cepacia syndrome“) [146]. *Burkholderia* express an unusual lipid A structure which is modified by esterification of the phosphate groups of lipid A by 4-amino-4-deoxy-β-L-arabinose (β-L-Ara4N). A covalent attachment of β-L-Ara4N at the anomeric 1-phosphate group or at the 4'-phosphate group of *Burkholderia* lipid A is estimated as a major pathogenic factor responsible for bacterial virulence and endurance in pulmonary airways [27]. Treatment with antibiotics inflicts selective pressure on BCC in the airways of immunocompromised patients which similarly results in the substitution of the lipid A phosphates by β-L-Ara4N. Addition of the cationic sugar β-L-Ara4N reduces the net negative charge of the bacterial membrane, which enhance bacterial resistance to CAMPs and aminoglycosides [146]. Incidences of profound resistance to polymyxin B – a first choice antibiotic for treatment of multidrug-resistant Gram-negative infections – is also attributed to the β-L-Ara4N modification of the lipid A moiety of LPS [32,147,148]. Accordingly, covalent modification of *Burkholderia* lipid A with Ara4N is crucial for bacterial persistence in the airways of infected patients and results in chronic inflammation and decreased survival [27]. Of special importance are the lipid A structures corresponding to highly pro-inflammatory *B. cenocepacia* [149] and *B. caryophilly* [150] LPS which are modified with β-L-Ara4N exclusively at the glycosidically linked 1-phosphate group of lipid A.

The Ara4N-modified LPS structures can hardly be obtained in pure form by isolation from bacterial cultures owing to intrinsic lability of the glycosyl phosphodiester functionality. The content of β-L-Ara4N in the bacterial isolated is usually reported as "non-stoichiometric" reflecting high degree of heterogeneity of the isolates in respect to substitution of the 1-phosphate group with β-L-Ara4N. To clarify the biological outcome of the Ara4N modification, a reliable synthetic approach toward β-L-Ara4N-containing LPS partial structures was developed [151]. To facilitate the assessment of an immunogenic potential of the unique β-L-Ara4N substitution at the glycosidically linked 1-phosphate group, a neoglycoconjugate **103** entailing an epitope βGlcN(1→6)-αGlcN(1→P←1)-β-L-Ara4N **102** was synthesised in a stereoselective manner [152] (Scheme 12).

For the assembly of binary glycosyl phosphodiester **102**, the synthesis of anomerically pure β-configured H-phosphonate monoester of the orthogonally protected β-L-Ara4N was initially performed (Scheme 12). To this end, the 2,3-*O*-tetraisopropyldisiloxane-1,3-diyl (TIPDS)-protected azide **93** was anomerically deprotected to furnish hemiacetal **95**. Since the stereoselectivity of the phosphitylation at the anomeric center generally relies on the anomeric ratio in the lactol precursor [153,154], the preparation of anomerically enriched hemiacetals which can be straightforwardly converted into the corresponding H-phosphonates comprised the foremost synthetic challenge. When the cleavage of the anomeric allyl group was carried out by sequential double bond isomerisation with $[Ir(1,5\text{-Cod})(PMePh_2)_2]^+PF_6^-$ to give propenyl glycoside **94**, followed by I_2-assisted prop-1-enyl cleavage, an anomeric mixture **95** (α/β = 1:1) was obtained. Lactol **95** could be enriched with the β-anomer (α/β = 1:3) by treatment with $CHCl_3$/MeOH/AcOH solution. Subsequent phosphitylation by reaction with salicylchlorophosphite (SalPCl) [139] in pyridine gave rise to the anomeric H-phosphonates (α/β = 1:3), whereas the β-anomer **96** could be isolated in a moderate 35% yield.

To achieve a better stereoselectivity, a novel procedure for traceless removal of the allyl group in β-allyl glycoside **93** without affecting the axial anomeric configuration at C-1 was elaborated. After allyl group isomerization, the anomeric prop-1-enyl ether **94** was oxidised by ozonolysis to give a stable formyl intermediate **97** under mild conditions (Scheme 12) [155-157]. The formate group was hydrolysed by methanolysis (NEt_3, MeOH, −40 °C) to furnished solely β-configured lactol **95β** and volatile methyl formate, so that the crude β-lactol could be directly subjected to phosphitylation without a need of chromatographic purification (which would result in a rapid anomerisation). A predominant formation of the β-configured H-phosphonate **96** was achieved by application of highly reactive phosphitylating reagent SalPCl, which quickly trapped the excess of axial β-lactol in **95β**, such that the initial α/β ratio was preserved and the anomerically pure β-glycosyl H-phosphonate **96** was obtained in 78% yield. Glycosyl-H-phosphonate **96** was initially coupled to the β(1→6)–linked diglucosamine lactol **82** [136] using pivaloyl chloride (PivCl) as activating agent [153,154,158] to furnish H-phosphonate glycosyl phosphodiester **98** as an anomeric mixture at GlcN moiety. Oxidation of **98** by treatment with aqueous I_2 at −40 °C afforded anomerically pure binary glycosyl phosphodiester **100**, whereas the more labile β-anomeric product was destroyed upon aqueous I_2-mediated oxidation and isolation of the phosphodiester **100**



Scheme 12: Synthesis of a binary 1,1′-glycosyl phosphodiester: a partial structure of β-L-Ara4N-modified *Burkholderia* Lipid A and a neoglycoconjugate based thereof.

by chromatography on silica gel [159]. Since the PivCl-mediated H-phosphonate coupling can be often accompanied by concomitant side-reactions (formation of P-acyl byproducts [140] resulting from an over-reaction of **96** or **98** with PivCl or formation of GlcNAc-derived oxazolines in the presence of an excess of chloroanhydride) [141], phosphonium type coupling reagents were optionally explored. Accordingly, the H-phosphonate **96** was activated by 3-nitro-1,2,4-triazol-1-yl-tris(pyrrolidin-1-yl)phosphonium hexafluorophosphate (PyNTP), which selectively reacted with the electrophilic phosphorus atom of the H-phosphonate to form a P–N activated intermediate [160,161]. The later was smoothly coupled to the nucleophilic component, the hemiacetal **82**. To circumvent possible hydrolysis of the binary glycosyl H-phosphonate diester **98** during the aqueous I_2-mediated oxidation step, the oxidation was performed in anhydrous conditions by transforming the tetra-coordinated H-phosphonate **98** into the three-coordinated silyl phosphite **99** (via treatment with *N,O*-bis(trimethylsilyl)acetamide (BTSA) in the presence of DBU) [162] followed by oxidation of **99** with 2-(phenylsulfonyl)-3-(3-nitrophenyl)oxaziridine (PNO) to furnish 1,1'-glycosyl phosphodiester **100**. The stepwise deprotection of **100** included a treatment with HF·Py to remove the TIPDS protecting group, a deacetylation of **101** (including deprotection of the 6-thioacetylhexanoyl residue) with MeOH/H_2O/NEt_3 and a final reduction of the 4-azido group by reaction with trimethylphosphine [142] in aq NaOH/THF which provided **102**. The formation of a disulfide bond was inhibited by application of reducing agent (PMe_3), so that the trisaccharide **102** could be directly coupled to a maleimide-activated BSA via a sulfhydryl-containing spacer group to furnish the neoglycoconjugate **103**. Thus, a novel efficient approach for anomeric deallylation with retention of configuration allowed for the stereoselective synthesis of anomerically pure β-L-Ara4N glycosyl H-phosphonate and β-L-Ara4N-containing antigenic LPS epitope as useful biochemical probe and potential diagnostic agent.

3.4. Synthesis of *Burkholderia* lipid A modified with glycosyl phosphodiester-linked β-L-Ara4N

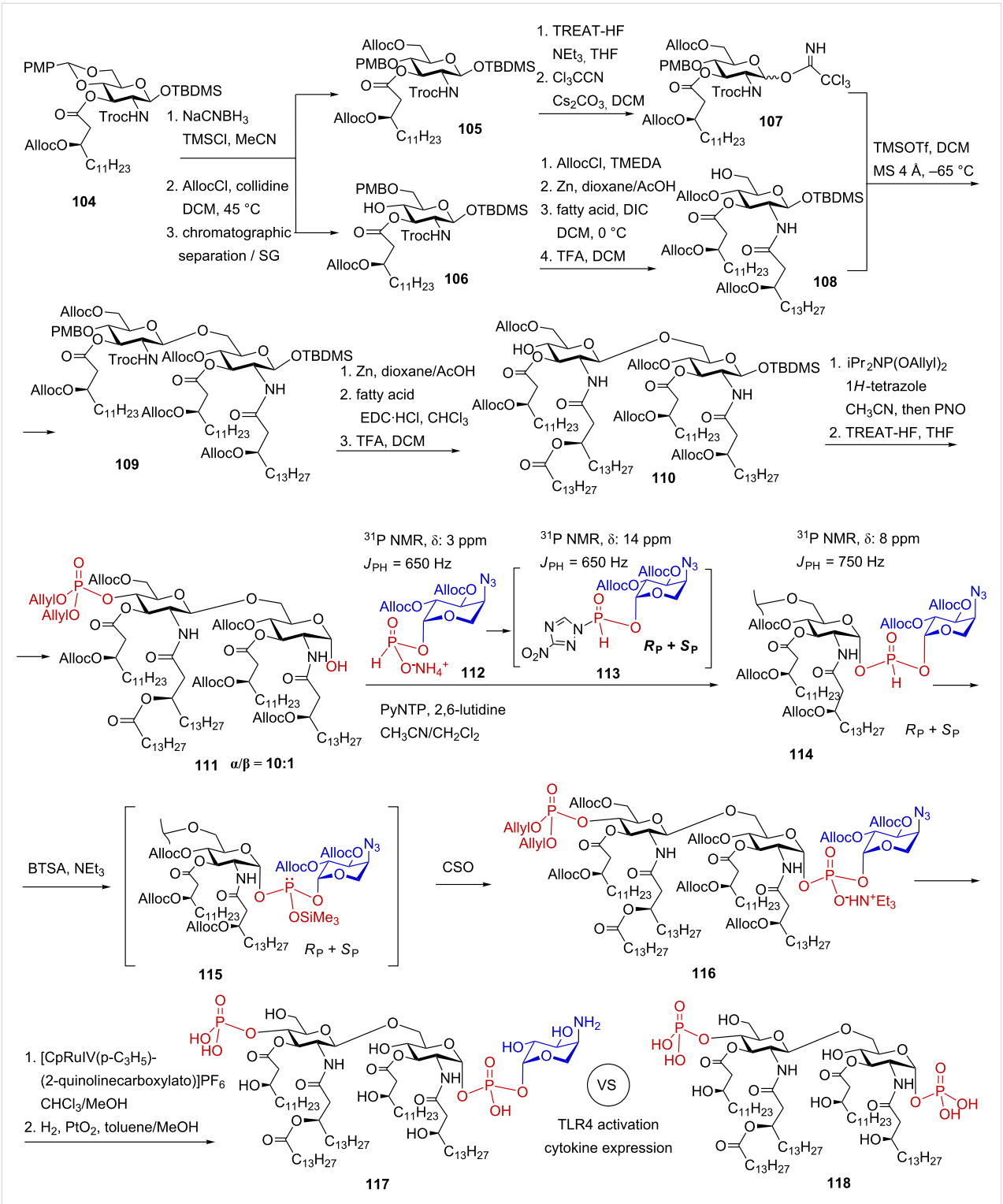
The pro-inflammatory activity of *Burkholderia* LPS isolates, which belongs to the major virulence factors of BCC species, has been extensively studied. Heterogeneous tetra- and penta-

acylated LPS/lipid A from *B. mallei* [163], *B. multivorans* [164], *B. cenocepacia* [149,165], *B. cepacia* [27] and *B. dolosa* [166] were determined as potent stimulators of the TLR4·MD-2-mediated cellular responses. Though it is generally believed that only hexaacyl lipid A (such as from *E. coli*) is capable of interacting with TLR4 complex and eliciting powerful innate immune response [18,167], underacylated β-L-Ara4N modified *Burkholderia* LPS isolates induced the expression of pro-inflammatory cytokines in vitro, and the efficiency of cytokine production was comparable with that induced by hexaacylated *E. coli* LPS [149]. The intrinsic lability of the aminosugar modification of the glycosyl phosphate group of lipid A results in a high degree of heterogeneity of lipid A preparations obtained from *Burkholderia* isolates in respect to the degree of β-L-Ara4N substitution which is commonly indicated as "non-stoichiometric". The lipid chain content in *Burkholderia* LPS also varies from species to species which makes it difficult to determine the structural characteristics of *Burkholderia* lipid A accountable for its unusual immuno-stimulating activity [168,169]. Since the 1-phosphate group of lipid A is directly involved in the formation of the dimeric MD-2·TLR4-LPS complex [42], the appendage of β-L-Ara4N might enhance the efficiency of dimerization via ionic attraction. In order to elucidate the structural determinants responsible for the unique pro-inflammatory potential of *Burkholderia* lipid A, the pentaacylated *Burkholderia* lipid A esterified by β-L-Ara4N at the anomeric phosphate **101** and its Ara4N-free counterpart **102** corresponding to native *Burkholderia* LPS were chemically synthesised [161].

The synthesis of fully orthogonally protected tetraacylated βGlcN(1→6)GlcN intermediate **109** commenced with the preparation of the GlcN-based *N*-Troc protected imidate donor **107** and the GlcN-derived bis-acylated 6-OH acceptor **108** (Scheme 13). Reductive opening of the *p*-methoxybenzylidene acetal protecting group in **104** with sodium cyanoborohydride and trimethylsilyl chloride in acetonitrile furnished a mixture of 6-OH and 4-OH (compound **106**) co-migrating regioisomers. This inseparable mixture was subjected to regioselective 6-*O*-protection with allyloxycarbonyl group by the action of allyloxycarbonyl chloride in the presence of *sym*-collidine, which transformed the 6-OH regioisomer into the 6-*O*-Alloc protected derivative **105**, whereas **106** having less reactive secondary 4-OH group did not react with AllocCl in the presence of a mild base. The resulting mixture – 6-*O*-Alloc-4-*O*-PMB protected **105** and 6-*O*-PMB protected **106** – was readily separated by conventional chromatography on silica gel. The anomeric TBDMS group in **105** was cleaved by treatment with triethylamine tris(hydrogenfluoride) (TREAT-HF) buffered by Et_3N (pH 6.5) which kept the acid labile 6-*O*-*p*-methoxybenzyl (PMB) group unaffected. The resultant hemiacetal was converted into fully protected trichloroacetimidate donor **107**.

The free secondary 4-OH group in **106** was protected by reaction with AllocCl in the presence of the stronger base *N,N,N′,N′*-tetramethylethylendiamine (TMEDA) [170]. The *N*-Troc group was subsequently reductively cleaved by treatment with Zn in acetic acid/dioxane followed by acylation of the intermediate amine by DIC-activated (*R*)-3-(allyloxycarbonyloxy)hexadecanoic acid. Succeeding acidic hydrolysis of the PMB group with trifluoroacetic acid furnished the 6-OH acceptor **108**. A TMSOTf–promoted glycosylation of **108** by the imidate donor **107** furnished a tetraacylated β(1→6)-linked disaccharide **109** (Scheme 13). Reduction of the 2′-*N*-Troc group by use of Zn in AcOH followed by *N*-acylation with (*R*)-3-acyloxyalkanoyl fatty acid in the presence of 1-ethyl-3-(3-dimethylaminopropyl) carbodiimide hydrochloride (EDC·HCl) gave fully protected pentaacylated intermediate which was treated with TFA in CH_2Cl_2 to promote hydrolysis of 4'-*O*-PMB group to furnish **110**. Compound **110** was phosphitylated at O-4' by reaction with diallyl(*N,N*-diisopropyl)phosphoramidite [171] in the presence of 1*H*-tetrazole and successive oxidation of the intermediate phosphite triester with PNO [172] to provide protected 4'-*O*-phosphate. The anomeric 1-*O*-TBDMS group in the latter was removed by treatment with TREAT-HF to give hemiacetal **111**. Since lactol **111** had to be stereoselectively coupled to the Ara4N H-phosphonate **112**, the anomeric preference of the α-configured lactol was especially important. Stabilization of the axial orientation of the 1-OH in **111** via intramolecular hydrogen bonding with the 2-NH group [154] ensured high proportion of the α-configured lactol (α/β = 10:1) and improved stereoselectivity in the next coupling step. Anomerically pure 2,3-di-*O*-Alloc protected β-L-Ara4N glycosyl H-phosphonate **112** was synthesised starting from 1-*O*-Allyl-4-azido β-L-Ara4N [173] in four steps [161].

The coupling of lactol **111** to the β-L-Ara4N glycosyl H-phosphonate **112** was promoted by 3-nitro-1,2,4-triazol-1-yl-tris(pyrrolidin-1-yl)phosphonium hexafluorophosphate (PyNTP) in the presence of 2,6-lutidine and afforded binary glycosyl H-phosphonate diester **114**. The H-phosphonate coupling reaction proceeded through formation of the tetracoordinated P(III) intermediates: H-pyrophosphonates [174] and nitrotriazol-1-yl-phosphites [175], such as β-L-Ara4N-nitrotriazol-1-yl-H-phosphonate **113** (^{31}P NMR (δ): 13 and 14 ppm, J_{PH} = 650 Hz), which instantly reacted with α-hemiacetal **111**. ^{31}P NMR spectroscopy was used to confirm the formation of a labile intermediate H-phosphonate diester **114** which displayed representative PH-coupled signals conforming with the formation of *R* and *S* diastereomers at phosphorus (^{31}P NMR (δ): 7.6 and 8.0 ppm, J_{PH} = 750 Hz). Due to exceptional lability of the



Scheme 13: Synthesis of *Burkholderia* lipid A containing binary glycosyl phosphodiester linked β-L-Ara4N.

binary glycosyl H-phosphonate diester **114**, the oxidation could not be performed under standard H-phosphonate chemistry conditions (aq. iodine) and, therefore, was accomplished in anhydrous conditions. To this end, the tetra-coordinated H-phosphonate was transformed into the three-coordinated phosphite **115** by reaction with *N,O*-bis(trimethylsilyl)acetamide [162,176] in the presence of Et_3N. The reaction was monitored by ^{31}P spectroscopy which confirmed the formation of the inter-

mediate phosphite **115**. Subsequent oxidation with (1*S*)-(+)-(10-camphorsulfonyl)oxaziridine (CSO) [177] furnished P(V) 1,1´-glycosyl phosphodiester **116**. Total cleavage of the Alloc- and Allyl- protecting groups in **116** was performed under mild neutral conditions [178] by treatment with [CpRu(IV)(π-C_3H_5)(2-quinolinecarboxylato)]PF_6 complex [179,180], so that a labile double glycosyl phosphodiester linkage was not affected. Finally, the azido group was reduced by hydrogenation on PtO_2 to give the target β-L-Ara4N-modified *Burkholderia* lipid A **117**. The availability of homogenous structurally defined synthetic β-L-Ara4N-modified *Burkholderia* lipid A provided a reliable tool for immunobiological studies. The immunomodulating potential of synthetic β-L-Ara4N-modified *Burkholderia* lipid A **117** and its non-modified synthetically prepared counterpart **118** was assessed in TLR4-transfected human embryonic kidney HEK293 cells by monitoring the activation of NF-κB signaling and in the human monocytic macrophage cell line THP-1. The β-L-Ara4N-modified lipid A **117** was considerably less efficient than *E. coli Re*-LPS in triggering the NF-κB signaling, however, it induced the expression of significantly higher levels of IL-8 compared to the non-modified pentaacyl bisphosphate lipid A **118** which was inactive at wide concentration range. Thus, the chemical synthesis of β-L-Ara4N-modified lipid A helped to reveal its immuno-modulatory potential and to demonstrate an enhancement of the pro-inflammatory activity of *Burkholderia* lipid A esterified by β-L-Ara4N at the glycosidically-linked phosphate group.

Conclusion

The synthesis of carbohydrate-based biomolecules is an area of fundamental and practical importance. Owing to immunomodulating capacities of lipid A and related glycolipids, the development of facile synthetic strategies toward these complex glycoconjugates have received particular attention. Despite huge progress achieved in the preparation of lipid A by combinatorial bioengineering of LPS and improved isolation techniques, the chemical synthesis remains the only source for sufficient amounts of structurally well-defined homogeneous materials which are completely free from any potentially pro-inflammatory biological contaminations and are suitable for biomedical or diagnostic application. Moreover, the intrinsic instability of particularly complex lipid A variants such as aminosugar-modified lipid A, renders the chemical synthesis to a single option for obtaining structurally integral compounds for biological studies. The inherent hybrid molecular structure of lipid A combining sugar-derived phosphorylated polar head group and multiple lipid moieties poses additional challenges to elaboration of efficient synthetic methodologies. Newly developed strategies allowed for divergent synthesis of LPS partial structures entailing lipid A that varies in the acylation pattern and the number of phosphate groups by the use of a single orthogonally protected disaccharide precursor. Application of advanced P(III) chemistry aided the development of stereoselective synthesis of binary glycosyl phosphodiesters comprising two aminosugars.

Acknowledgements

Financial support by Austrian Science Fund (FWF) Grants P-21276, P-22116 and P-28915 is gratefully acknowledged.

References

1. Kawai, T.; Akira, S. *Nat. Immunol.* **2010,** *11,* 373–384. doi:10.1038/ni.1863
2. Raetz, C. R. H.; Reynolds, C. M.; Trent, M. S.; Bishop, R. E. *Annu. Rev. Biochem.* **2007,** *76,* 295–329. doi:10.1146/annurev.biochem.76.010307.145803
3. Trent, M. S.; Stead, C. M.; Tran, A. X.; Hankins, J. V. *Innate Immun.* **2006,** *12,* 205–223. doi:10.1177/0968051906012004020l
4. Kobayashi, M.; Saitoh, S.-i.; Tanimura, N.; Takahashi, K.; Kawasaki, K.; Nishijima, M.; Fujimoto, Y.; Fukase, K.; Akashi-Takamura, S.; Miyake, K. *J. Immunol.* **2006,** *176,* 6211–6218. doi:10.4049/jimmunol.176.10.6211
5. Silipo, A.; Molinaro, A. Lipid A structure. In *Bacterial Lipopolysaccharides;* Knirel, Y. A.; Valvano, M. A., Eds.; Springer: Vienna, 2011; pp 1–20. doi:10.1007/978-3-7091-0733-1_1
6. Holst, O. Structure of the lipopolysaccharide core region. In *Bacterial Lipopolysaccharides;* Knirel, Y. A.; Valvano, M. A., Eds.; Springer: Vienna, 2011; pp 21–39. doi:10.1007/978-3-7091-0733-1_2
7. Knirel, Y. A. Structure of O-Antigens. In *Bacterial Lipopolysaccharides;* Knirel, Y. A.; Valvano, M. A., Eds.; Springer: Vienna, 2011; pp 41–115. doi:10.1007/978-3-7091-0733-1_3
8. Ishii, K. J.; Koyama, S.; Nakagawa, A.; Coban, C.; Akira, S. *Cell Host Microbe* **2008,** *3,* 352–363. doi:10.1016/j.chom.2008.05.003
9. Rittirsch, D.; Flierl, M. A.; Ward, P. A. *Nat. Rev. Immunol.* **2008,** *8,* 776–787. doi:10.1038/nri2402
10. Opal, S. M. *Int. J. Med. Microbiol.* **2007,** *297,* 365–377. doi:10.1016/j.ijmm.2007.03.006
11. Hunter, P. *EMBO Rep.* **2006,** *7,* 667–669. doi:10.1038/sj.embor.7400742
12. Munford, R. S. *Infect. Immun.* **2008,** *76,* 454–465. doi:10.1128/IAI.00939-07
13. Park, S. H.; Kim, N. D.; Jung, J.-K.; Lee, C.-K.; Han, S.-B.; Kim, Y. *Pharmacol. Ther.* **2012,** *133,* 291–298. doi:10.1016/j.pharmthera.2011.11.001
14. Alexander, C.; Rietschel, E. T. *J. Endotoxin Res.* **2001,** *7,* 167–202.
15. Rietschel, E. T.; Kirikae, T.; Schade, F. U.; Mamat, U.; Schmidt, G.; Loppnow, H.; Ulmer, A. J.; Zähringer, U.; Seydel, U.; Di Padova, F. *FASEB J.* **1994,** *8,* 217–225.
16. Raetz, C. R. H.; Whitfield, C. *Annu. Rev. Biochem.* **2002,** *71,* 635–700. doi:10.1146/annurev.biochem.71.110601.135414
17. Zähringer, U.; Lindner, B.; Rietschel, E. T. Chemical structure of lipid A: recent advances in structural analysis of bilogically active molecules. In *Endotoxin in health and disease;* Brade, H.; Opal, S.; Vogel, S. N.; Morrison, D. C., Eds.; Marcel Dekker, Inc.: New York - Basel, 1999; pp 93–114.
18. Needham, B. D.; Carroll, S. M.; Giles, D. K.; Georgiou, G.; Whiteley, M.; Trent, M. S. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 1464–1469. doi:10.1073/pnas.1218080110

19. Artner, D.; Oblak, A.; Ittig, S.; Garate, J. A.; Horvat, S.; Arrieumerlou, C.; Hofinger, A.; Oostenbrink, C.; Jerala, R.; Kosma, P.; Zamyatina, A. *ACS Chem. Biol.* **2013,** *8,* 2423–2432. doi:10.1021/cb4003199
20. Adanitsch, F.; Ittig, S.; Stöckl, J.; Oblak, A.; Haegman, M.; Jerala, R.; Beyaert, R.; Kosma, P.; Zamyatina, A. *J. Med. Chem.* **2014,** *57,* 8056–8071. doi:10.1021/jm500946r
21. Shimoyama, A.; Saeki, A.; Tanimura, N.; Tsutsui, H.; Miyake, K.; Suda, Y.; Fujimoto, Y.; Fukase, K. *Chem. – Eur. J.* **2011,** *17,* 14464–14474. doi:10.1002/chem.201003581
22. Christ, W. J.; McGuinness, P. D.; Asano, O.; Wang, Y.; Mullarkey, M. A.; Perez, M.; Hawkins, L. D.; Blythe, T. A.; Dubuc, G. R.; Robidoux, A. L. *J. Am. Chem. Soc.* **1994,** *116,* 3637–3638. doi:10.1021/ja00087a075
23. Oikawa, M.; Wada, A.; Yoshizaki, H.; Fukase, K.; Kusumoto, S. *Bull. Chem. Soc. Jpn.* **1997,** *70,* 1435–1440. doi:10.1246/bcsj.70.1435
24. Montminy, S. W.; Khan, N.; McGrath, S.; Walkowicz, M. J.; Sharp, F.; Conlon, J. E.; Fukase, K.; Kusumoto, S.; Sweet, C.; Miyake, K.; Akira, S.; Cotter, R. J.; Goguen, J. D.; Lien, E. *Nat. Immunol.* **2006,** *7,* 1066–1073. doi:10.1038/ni1386
25. Needham, B. D.; Trent, M. S. *Nat. Rev. Microbiol.* **2013,** *11,* 467–481. doi:10.1038/nrmicro3047
26. Li, Y.; Powell, D. A.; Shaffer, S. A.; Rasko, D. A.; Pelletier, M. R.; Leszyk, J. D.; Scott, A. J.; Masoudi, A.; Goodlett, D. R.; Wang, X.; Raetz, C. R. H.; Ernst, R. K. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109,* 8716–8721. doi:10.1073/pnas.1202908109
27. De Soyza, A.; Silipo, A.; Lanzetta, R.; Govan, J. R.; Molinaro, A. *Innate Immun.* **2008,** *14,* 127–144. doi:10.1177/1753425908093984
28. Molinaro, A.; Holst, O.; Di Lorenzo, F.; Callaghan, M.; Nurisso, A.; D'Errico, G.; Zamyatina, A.; Peri, F.; Berisio, R.; Jerala, R.; Jiménez-Barbero, J.; Silipo, A.; Martín-Santamaria, S. *Chem. – Eur. J.* **2015,** *21,* 500–519. doi:10.1002/chem.201403923
29. Breazeale, S. D.; Ribeiro, A. A.; Raetz, C. R. H. *J. Biol. Chem.* **2003,** *278,* 24731–24739. doi:10.1074/jbc.M304043200
30. Knirel, Y. A.; Dentovskaya, S. V.; Senchenkova, S. N.; Shaikhutdinova, R. Z.; Kocharova, N. A.; Anisimov, A. P. *J. Endotoxin Res.* **2006,** *12,* 3–9.
31. Marr, N.; Tirsoaga, A.; Blanot, D.; Fernandez, R.; Caroff, M. *J. Bacteriol.* **2008,** *190,* 4281–4290. doi:10.1128/JB.01875-07
32. Hamad, M. A.; Di Lorenzo, F.; Molinaro, A.; Valvano, M. A. *Mol. Microbiol.* **2012,** *85,* 962–974. doi:10.1111/j.1365-2958.2012.08154.x
33. Anisimov, A. P.; Dentovskaya, S. V.; Titareva, G. M.; Bakhteeva, I. V.; Shaikhutdinova, R. Z.; Balakhonov, S. V.; Lindner, B.; Kocharova, N. A.; Senchenkova, S. N.; Holst, O.; Pier, G. B.; Knirel, Y. A. *Infect. Immun.* **2005,** *73,* 7324–7331. doi:10.1128/IAI.73.11.7324-7331.2005
34. Knirel, Y. A.; Kocharova, N. A.; Titareva, G. M.; Bakhteeva, I. V.; Senchenkova, S. N.; Bystrova, O. V.; Dentovskaya, S. V.; Anisimov, A. P.; Pier, G. B.; Lindner, B.; Shaikhutdinova, R. Z. *Adv. Exp. Med. Biol.* **2007,** *603,* 88–96. doi:10.1007/978-0-387-72124-8_7
35. Schumann, R. R.; Leong, S. R.; Flaggs, G. W.; Gray, P. W.; Wright, S. D.; Mathison, J. C.; Tobias, P. S.; Ulevitch, R. J. *Science* **1990,** *249,* 1429–1431. doi:10.1126/science.2402637
36. Ulevitch, R. J.; Tobias, P. S. *Curr. Opin. Immunol.* **1999,** *11,* 19–22. doi:10.1016/S0952-7915(99)80004-1
37. Wright, S. D.; Ramos, R. A.; Tobias, P. S.; Ulevitch, R. J.; Mathison, J. C. *Science* **1990,** *249,* 1431–1433. doi:10.1126/science.1698311
38. Shimazu, R.; Akashi, S.; Ogata, H.; Nagai, Y.; Fukudome, K.; Miyake, K.; Kimoto, M. *J. Exp. Med.* **1999,** *189,* 1777–1782. doi:10.1084/jem.189.11.1777
39. Medzhitov, R.; Preston-Hurlburt, P.; Janeway, C. A., Jr. *Nature* **1997,** *388,* 394–397. doi:10.1038/41131
40. Miyake, K. *J. Endotoxin Res.* **2006,** *12,* 195–204.
41. Prohinar, P.; Re, F.; Widstrom, R.; Zhang, D.; Teghanemt, A.; Weiss, J. P.; Gioannini, T. L. *J. Biol. Chem.* **2007,** *282,* 1010–1017. doi:10.1074/jbc.M609400200
42. Park, B. S.; Song, D. H.; Kim, H. M.; Choi, B.-S.; Lee, H.; Lee, J.-O. *Nature* **2009,** *458,* 1191–1195. doi:10.1038/nature07830
43. Poltorak, A.; He, X.; Smirnova, I.; Liu, M.-Y.; Van Huffel, C.; Du, X.; Birdwell, D.; Alejos, E.; Silva, M.; Galanos, C.; Freudenberg, M.; Ricciardi-Castagnoli, P.; Layton, B.; Beutler, B. *Science* **1998,** *282,* 2085–2088. doi:10.1126/science.282.5396.2085
44. Chow, J. C.; Young, D. W.; Golenbock, D. T.; Christ, W. J.; Gusovsky, F. *J. Biol. Chem.* **1999,** *274,* 10689–10692. doi:10.1074/jbc.274.16.10689
45. Hold, G. L.; Bryant, C. E. The molecular basis of Lipid A and Toll-Like Receptor 4 interactions. In *Bacterial Lipopolysaccharides;* Knirel, Y. A.; Valvano, M. A., Eds.; Springer: Vienna, 2011; pp 371–387. doi:10.1007/978-3-7091-0733-1_12
46. Kumar, H.; Kawai, T.; Akira, S. *Biochem. Biophys. Res. Commun.* **2009,** *388,* 621–625. doi:10.1016/j.bbrc.2009.08.062
47. Ohto, U.; Fukase, K.; Miyake, K.; Satow, Y. *Science* **2007,** *316,* 1632–1634. doi:10.1126/science.1139111
48. Takayama, K.; Qureshi, N.; Beutler, B.; Kirkland, T. N. *Infect. Immun.* **1989,** *57,* 1336–1338.
49. Hawkins, L. D.; Christ, W. J.; Rossignol, D. P. *Curr. Top. Med. Chem.* **2004,** *4,* 1147–1171. doi:10.2174/1568026043388123
50. Shirey, K. A.; Lai, W.; Scott, A. J.; Lipsky, M.; Mistry, P.; Pletneva, L. M.; Karp, C. L.; McAlees, J.; Gioannini, T. L.; Weiss, J.; Chen, W. H.; Ernst, R. K.; Rossignol, D. P.; Gusovsky, F.; Blanco, J. C. G.; Vogel, S. N. *Nature* **2013,** *497,* 498–502. doi:10.1038/nature12118
51. Hammad, H.; Chieppa, M.; Perros, F.; Willart, M. A.; Germain, R. N.; Lambrecht, B. N. *Nat. Med.* **2009,** *15,* 410–416. doi:10.1038/nm.1946
52. Abdollahi-Roodsaz, S.; Joosten, L. A. B.; Roelofs, M. F.; Radstake, T. R. D. J.; Matera, G.; Popa, C.; van der Meer, J. W. M.; Netea, M. G.; van den Berg, W. B. *Arthritis Rheum.* **2007,** *56,* 2957–2967. doi:10.1002/art.22848
53. Abreu, M. T.; Arditi, M. *J. Pediatr.* **2004,** *144,* 421–429. doi:10.1016/j.jpeds.2004.01.057
54. Frantz, S.; Ertl, G.; Bauersachs, J. *Nat. Rev. Cardiol.* **2007,** *4,* 444–454. doi:10.1038/ncpcardio0938
55. Michaud, J.-P.; Hallé, M.; Lampron, A.; Thériault, P.; Préfontaine, P.; Filali, M.; Tribout-Jover, P.; Lanteigne, A.-M.; Jodoin, R.; Cluff, C.; Brichard, V.; Palmantier, R.; Pilorget, A.; Larocque, D.; Rivest, S. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 1941–1946. doi:10.1073/pnas.1215165110
56. Reisser, D.; Jeannin, J.-F. Lipid A in cancer therapies preclinical results. In *Lipid A in cancer therapy;* Jeannin, J.-F., Ed.; Springer, Landes Bioscience: New York, 2009; pp 101–110. doi:10.1007/978-1-4419-1603-7_9
57. Iwasaki, A.; Medzhitov, R. *Nat. Immunol.* **2004,** *5,* 987–995. doi:10.1038/ni1112
58. Medzhitov, R.; Janeway, C., Jr. *N. Engl. J. Med.* **2000,** *343,* 338–344. doi:10.1056/NEJM200008033430506
59. Casella, C. R.; Mitchell, T. C. *Cell. Mol. Life Sci.* **2008,** *65,* 3231–3240. doi:10.1007/s00018-008-8228-6

60. Ishizaka, S. T.; Hawkins, L. D. *Expert Rev. Vaccines* **2007,** *6,* 773–784. doi:10.1586/14760584.6.5.773
61. Johnson, D. A. *Curr. Top. Med. Chem.* **2008,** *8,* 64–79. doi:10.2174/156802608783378882
62. Kim, H. M.; Park, B. S.; Kim, J.-I.; Kim, S. E.; Lee, J.; Oh, S. C.; Enkhbayar, P.; Matsushima, N.; Lee, H.; Yoo, O. J.; Lee, J.-O. *Cell* **2007,** *130,* 906–917. doi:10.1016/j.cell.2007.08.002
63. Bryant, C. E.; Spring, D. R.; Gangloff, M.; Gay, N. J. *Nat. Rev. Microbiol.* **2010,** *8,* 8–14. doi:10.1038/nrmicro2266
64. Ohto, U.; Fukase, K.; Miyake, K.; Shimizu, T. *Proc. Natl. Acad. Sci. U. S. A.* **2012,** *109,* 7421–7426. doi:10.1073/pnas.1201193109
65. Meng, J.; Gong, M.; Björkbacka, H.; Golenbock, D. T. *J. Immunol.* **2011,** *187,* 3683–3693. doi:10.4049/jimmunol.1101397
66. Teghanemt, A.; Re, F.; Prohinar, P.; Widstrom, R.; Gioannini, T. L.; Weiss, J. P. *J. Biol. Chem.* **2008,** *283,* 1257–1266. doi:10.1074/jbc.M705994200
67. Resman, N.; Vašl, J.; Oblak, A.; Pristovšek, P.; Gioannini, T. L.; Weiss, J. P.; Jerala, R. *J. Biol. Chem.* **2009,** *284,* 15052–15060. doi:10.1074/jbc.M901429200
68. Kayagaki, N.; Wong, M. T.; Stowe, I. B.; Ramani, S. R.; Gonzalez, L. C.; Akashi-Takamura, S.; Miyake, K.; Zhang, J.; Lee, W. P.; Muszyński, A.; Forsberg, L. S.; Carlson, R. W.; Dixit, V. M. *Science* **2013,** *341,* 1246–1249. doi:10.1126/science.1240248
69. Hagar, J. A.; Aachoui, Y.; Miao, E. A. *Cell Res.* **2015,** *25,* 149–150. doi:10.1038/cr.2014.128
70. Hagar, J. A.; Powell, D. A.; Aachoui, Y.; Ernst, R. K.; Miao, E. A. *Science* **2013,** *341,* 1250–1253. doi:10.1126/science.1240988
71. Kagan, J. C. *Science* **2013,** *341,* 1184–1185. doi:10.1126/science.1243939
72. Shi, J.; Zhao, Y.; Wang, Y.; Gao, W.; Ding, J.; Li, P.; Hu, L.; Shao, F. *Nature* **2014,** *514,* 187–192. doi:10.1038/nature13683
73. Kajiwara, Y.; Schiff, T.; Voloudakis, G.; Gama Sosa, M. A.; Elder, G.; Bozdagi, O.; Buxbaum, J. D. *J. Immunol.* **2014,** *193,* 335–343. doi:10.4049/jimmunol.1303424
74. Zhang, Y.; Gaekwad, J.; Wolfert, M. A.; Boons, G.-J. *J. Am. Chem. Soc.* **2007,** *129,* 5200–5216. doi:10.1021/ja068922a
75. Zamyatina, A.; Sekljic, H.; Brade, H.; Kosma, P. *Tetrahedron* **2004,** *60,* 12113–12137. doi:10.1016/j.tet.2004.10.017
76. Kusumoto, S.; Fukase, K. *Chem. Rec.* **2006,** *6,* 333–343. doi:10.1002/tcr.20098
77. Yoshizaki, H.; Fukuda, N.; Sato, K.; Oikawa, M.; Fukase, K.; Suda, Y.; Kusumoto, S. *Angew. Chem., Int. Ed.* **2001,** *40,* 1475–1480. doi:10.1002/1521-3773(20010417)40:8<1475::AID-ANIE1475>3.0.CO;2-V
78. Zhang, Y.; Gaekwad, J.; Wolfert, M. A.; Boons, G.-J. *Chem. – Eur. J.* **2007,** *14,* 558–569. doi:10.1002/chem.200701165
79. Oikawa, M.; Furuta, H.; Suda, Y.; Kusumoto, S. *Tetrahedron Lett.* **1999,** *40,* 5199–5202. doi:10.1016/S0040-4039(99)00937-5
80. Parsonnet, J.; Friedman, G. D.; Vandersteen, D. P.; Chang, Y.; Vogelman, J. H.; Orentreich, N.; Sibley, R. K. *N. Engl. J. Med.* **1991,** *325,* 1127–1131. doi:10.1056/NEJM199110173251603
81. Moran, A. P. *Int. J. Med. Microbiol.* **2007,** *297,* 307–319. doi:10.1016/j.ijmm.2007.03.008
82. Moran, A. P.; Lindner, B.; Walsh, E. J. *J. Bacteriol.* **1997,** *179,* 6453–6463. doi:10.1128/jb.179.20.6453-6463.1997
83. Suda, Y.; Kim, Y.-M.; Ogawa, T.; Yasui, N.; Hasegawa, Y.; Kashihara, W.; Shimoyama, T.; Aoyama, K.; Nagata, K.; Tamura, T.; Kusumoto, S. *Innate Immun.* **2001,** *7,* 95–104. doi:10.1177/096805190100700203O1
84. Tran, A. X.; Stead, C. M.; Trent, M. S. *Innate Immun.* **2005,** *11,* 161–166. doi:10.1177/096805190501100304O1
85. Ogawa, T.; Asai, Y.; Sakai, Y.; Oikawa, M.; Fukase, K.; Suda, Y.; Kusumoto, S.; Tamura, T. *FEMS Immunol. Med. Microbiol.* **2003,** *36,* 1–7. doi:10.1016/S0928-8244(03)00093-2
86. Sakai, Y.; Oikawa, M.; Yoshizaki, H.; Ogawa, T.; Suda, Y.; Fukase, K.; Kusumoto, S. *Tetrahedron Lett.* **2000,** *41,* 6843–6847. doi:10.1016/S0040-4039(00)01158-8
87. Fujimoto, Y.; Iwata, M.; Imakita, N.; Shimoyama, A.; Suda, Y.; Kusumoto, S.; Fukase, K. *Tetrahedron Lett.* **2007,** *48,* 6577–6581. doi:10.1016/j.tetlet.2007.07.036
88. Fujimoto, Y.; Shimoyama, A.; Suda, Y.; Fukase, K. *Carbohydr. Res.* **2012,** *356,* 37–43. doi:10.1016/j.carres.2012.03.013
89. Shiina, I.; Ushiyama, H.; Yamada, Y.-k.; Kawakita, Y.-i.; Nakata, K. *Chem. – Asian J.* **2008,** *3,* 454–461. doi:10.1002/asia.200700305
90. Merzouk, A.; Guibé, F.; Loffet, A. *Tetrahedron Lett.* **1992,** *33,* 477–480. doi:10.1016/S0040-4039(00)93973-X
91. Murray, R. W.; Jeyaraman, R. *J. Org. Chem.* **1985,** *50,* 2847–2853. doi:10.1021/jo00216a007
92. Herath, T. D. K.; Darveau, R. P.; Seneviratne, C. J.; Wang, C.-Y.; Wang, Y.; Jin, L. *PLoS One* **2013,** *8,* e58496. doi:10.1371/journal.pone.0058496
93. Herath, T. D. K.; Wang, Y.; Seneviratne, C. J.; Lu, Q.; Darveau, R. P.; Wang, C.-Y.; Jin, L. *J. Clin. Periodontol.* **2011,** *38,* 694–701. doi:10.1111/j.1600-051X.2011.01741.x
94. Curtis, M. A.; Percival, R. S.; Devine, D.; Darveau, R. P.; Coats, S. R.; Rangarajan, M.; Tarelli, E.; Marsh, P. D. *Infect. Immun.* **2011,** *79,* 1187–1193. doi:10.1128/IAI.00900-10
95. Kumada, H.; Haishima, Y.; Watanabe, K.; Hasegawa, C.; Tsuchiya, T.; Tanamoto, K.; Umemoto, T. *Mol. Oral Microbiol.* **2008,** *23,* 60–69. doi:10.1111/j.1399-302X.2007.00392.x
96. Ogawa, T.; Asai, Y.; Hashimoto, M.; Takeuchi, O.; Kurita, T.; Yoshikai, Y.; Miyake, K.; Akira, S. *Int. Immunol.* **2002,** *14,* 1325–1332. doi:10.1093/intimm/dxf097
97. Reife, R. A.; Coats, S. R.; Al-Qutub, M.; Dixon, D. M.; Braham, P. A.; Billharz, R. J.; Howald, W. N.; Darveau, R. P. *Cell. Microbiol.* **2006,** *8,* 857–868. doi:10.1111/j.1462-5822.2005.00672.x
98. Zhang, Y.; Gaekwad, J.; Wolfert, M. A.; Boons, G.-J. *Org. Biomol. Chem.* **2008,** *6,* 3371–3381. doi:10.1039/b809090d
99. Fujimoto, Y.; Shimoyama, A.; Saeki, A.; Kitayama, N.; Kasamatsu, C.; Tsutsui, H.; Fukase, K. *Mol. BioSyst.* **2013,** *9,* 987–996. doi:10.1039/c3mb25477a
100. Pasare, C.; Medzhitov, R. *Microbes Infect.* **2004,** *6,* 1382–1387. doi:10.1016/j.micinf.2004.08.018
101. Kong, Q.; Six, D. A.; Roland, K. L.; Liu, Q.; Gu, L.; Reynolds, C. M.; Wang, X.; Raetz, C. R. H.; Curtiss, R., III. *J. Immunol.* **2011,** *187,* 412–423. doi:10.4049/jimmunol.1100339
102. Kong, Q.; Six, D. A.; Liu, Q.; Gu, L.; Wang, S.; Alamuri, P.; Raetz, C. R. H.; Curtiss, R. *Infect. Immun.* **2012,** *80,* 1302.
103. Casella, C. R.; Mitchell, T. C. *PLoS One* **2013,** *8,* e62622. doi:10.1371/journal.pone.0062622
104. Mata-Haro, V.; Cekic, C.; Martin, M.; Chilton, P. M.; Casella, C. R.; Mitchell, T. C. *Science* **2007,** *316,* 1628–1632. doi:10.1126/science.1138963
105. Pouliot, K.; Buglione-Corbett, R.; Marty-Roix, R.; Montminy-Paquette, S.; West, K.; Wang, S.; Lu, S.; Lien, E. *Vaccine* **2014,** *32,* 5049–5056. doi:10.1016/j.vaccine.2014.07.010
106. Maiti, K. K.; DeCastro, M.; El-Sayed, A.-B. M. A.-A.; Foote, M. I.; Wolfert, M. A.; Boons, G.-J. *Eur. J. Org. Chem.* **2010,** 80–91. doi:10.1002/ejoc.200900973

107. Jiang, Z.-H.; Budzynski, W. A.; Qiu, D.; Yalamati, D.; Koganty, R. R. *Carbohydr. Res.* **2007,** *342,* 784–796. doi:10.1016/j.carres.2007.01.012
108. Zhou, Z.; Mondal, M.; Liao, G.; Guo, Z. *Org. Biomol. Chem.* **2014,** *12,* 3238–3245. doi:10.1039/C4OB00390J
109. Wang, Q.; Xue, J.; Guo, Z. *Chem. Commun.* **2009,** 5536–5537. doi:10.1039/b907351e
110. Ingale, S.; Wolfert, M. A.; Buskas, T.; Boons, G.-J. *ChemBioChem* **2009,** *10,* 455–463. doi:10.1002/cbic.200800596
111. Bhat, U. R.; Forsberg, L. S.; Carlson, R. W. *J. Biol. Chem.* **1994,** *269,* 14402–14410.
112. Jeyaretnam, B.; Glushka, J.; Kolli, V. S. K.; Carlson, R. W. *J. Biol. Chem.* **2002,** *277,* 41802–41810. doi:10.1074/jbc.M112140200
113. Karbarz, M. J.; Kalb, S. R.; Cotter, R. J.; Raetz, C. R. H. *J. Biol. Chem.* **2003,** *278,* 39269–39279. doi:10.1074/jbc.M305830200
114. Que-Gewirth, N. L. S.; Lin, S.; Cotter, R. J.; Raetz, C. R. H. *J. Biol. Chem.* **2003,** *278,* 12109–12119. doi:10.1074/jbc.M300378200
115. Vandenplas, M. L.; Carlson, R. W.; Jeyaretnam, B. S.; McNeill, B.; Barton, M. H.; Norton, N.; Murray, T. F.; Moore, J. N. *J. Biol. Chem.* **2002,** *277,* 41811–41816. doi:10.1074/jbc.M205252200
116. Demchenko, A. V.; Wolfert, M. A.; Santhanam, B.; Moore, J. N.; Boons, G.-J. *J. Am. Chem. Soc.* **2003,** *125,* 6103–6112. doi:10.1021/ja029316s
117. Vasan, M.; Wolfert, M. A.; Boons, G.-J. *Org. Biomol. Chem.* **2007,** *5,* 2087–2097. doi:10.1039/b704427e
118. Santhanam, B.; Wolfert, M. A.; Moore, J. N.; Boons, G.-J. *Chem. – Eur. J.* **2004,** *10,* 4798–4807. doi:10.1002/chem.200400376
119. Santhanam, B.; Boons, G.-J. *Org. Lett.* **2004,** *6,* 3333–3336. doi:10.1021/ol048746f
120. Duynstee, H. I.; van Vliet, M. J.; van der Marel, G. A.; van Boom, J. H. *Eur. J. Org. Chem.* **1998,** 303–307. doi:10.1002/(SICI)1099-0690(199802)1998:2<303::AID-EJOC303>3.0.CO;2-U
121. Nikolaev, A. V.; Botvinko, I. V.; Ross, A. J. *Carbohydr. Res.* **2007,** *342,* 297–344. doi:10.1016/j.carres.2006.10.006
122. Oscarson, S. *Carbohydr. Polym.* **2001,** *44,* 305–311. doi:10.1016/S0144-8617(00)00246-0
123. Berkin, A.; Coxon, B.; Pozsgay, V. *Chem. – Eur. J.* **2002,** *8,* 4424–4433. doi:10.1002/1521-3765(20021004)8:19<4424::AID-CHEM4424>3.0.CO;2-1
124. Cowley, S. C.; Elkins, K. L. *Front. Microbiol.* **2011,** *2,* No. 26. doi:10.3389/fmicb.2011.00026
125. Conlan, J. W.; Oyston, P. C. F. *Ann. N. Y. Acad. Sci.* **2007,** *1105,* 325–350. doi:10.1196/annals.1409.012
126. Gunn, J. S.; Ernst, R. K. *Ann. N. Y. Acad. Sci.* **2007,** *1105,* 202–218. doi:10.1196/annals.1409.006
127. Hajjar, A. M.; Harvey, M. D.; Shaffer, S. A.; Goodlett, D. R.; Sjöstedt, A.; Edebro, H.; Forsman, M.; Byström, M.; Pelletier, M.; Wilson, C. B.; Miller, S. I.; Skerrett, S. J.; Ernst, R. K. *Infect. Immun.* **2006,** *74,* 6730–6738. doi:10.1128/IAI.00934-06
128. Wang, X.; Ribeiro, A. A.; Guan, Z.; Raetz, C. R. H. *Biochemistry* **2009,** *48,* 1162–1172. doi:10.1021/bi802211k
129. Phillips, N. J.; Schilling, B.; McLendon, M. K.; Apicella, M. A.; Gibson, B. W. *Infect. Immun.* **2004,** *72,* 5340–5348. doi:10.1128/IAI.72.9.5340-5348.2004
130. Wang, X.; McGrath, S. C.; Cotter, R. J.; Raetz, C. R. H. *J. Biol. Chem.* **2006,** *281,* 9321–9330. doi:10.1074/jbc.M600435200
131. Kanistanon, D.; Hajjar, A. M.; Pelletier, M. R.; Gallagher, L. A.; Kalhorn, T.; Shaffer, S. A.; Goodlett, D. R.; Rohmer, L.; Brittnacher, M. J.; Skerrett, S. J.; Ernst, R. K. *PLoS Pathog.* **2008,** *4,* e24. doi:10.1371/journal.ppat.0040024
132. Zhao, J.; Raetz, C. R. H. *Mol. Microbiol.* **2010,** *78,* 820–836. doi:10.1111/j.1365-2958.2010.07305.x
133. Chalabaev, S.; Kim, T.-H.; Ross, R.; Derian, A.; Kasper, D. L. *J. Biol. Chem.* **2010,** *285,* 34330–34336. doi:10.1074/jbc.M110.166314
134. Okan, N. A.; Chalabaev, S.; Kim, T.-H.; Fink, A.; Ross, R. A.; Kasper, D. L. *mBio* **2013,** *4,* e00638–12. doi:10.1128/mBio.00638-12
135. Wang, X.; Ribeiro, A. A.; Guan, Z.; McGrath, S. C.; Cotter, R. J.; Raetz, C. R. H. *Biochemistry* **2006,** *45,* 14427–14440. doi:10.1021/bi061767s
136. Baum, D.; Kosma, P.; Zamyatina, A. *Org. Lett.* **2014,** *16,* 3772–3775. doi:10.1021/ol501639c
137. Tärnvik, A.; Chu, M. C. *Ann. N. Y. Acad. Sci.* **2007,** *1105,* 378–404. doi:10.1196/annals.1409.017
138. Imamura, A.; Ando, H.; Ishida, H.; Kiso, M. *Curr. Org. Chem.* **2008,** *12,* 675–689. doi:10.2174/138527208784577358
139. Marugg, J. E.; Tromp, M.; Kuyl-Yeheskiely, E.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron Lett.* **1986,** *27,* 2661–2664. doi:10.1016/S0040-4039(00)84611-0
140. Froehler, B. C.; Matteucci, M. D. *Tetrahedron Lett.* **1986,** *27,* 469–472. doi:10.1016/S0040-4039(00)85507-0
141. Westerduin, P.; Veeneman, G. H.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron Lett.* **1986,** *27,* 6271–6274. doi:10.1016/S0040-4039(00)85450-7
142. Rivkin, A.; Yoshimura, F.; Gabarda, A. E.; Cho, Y. S.; Chou, T.-C.; Dong, H.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2004,** *126,* 10913–10922. doi:10.1021/ja046992g
143. Bartra, M.; Romea, P.; Urpí, F.; Vilarrasa, J. *Tetrahedron* **1990,** *46,* 587–594. doi:10.1016/S0040-4020(01)85439-9
144. Bartra, M.; Urpí, F.; Vilarrasa, J. *Tetrahedron Lett.* **1987,** *28,* 5941–5944. doi:10.1016/S0040-4039(01)81096-0
145. Burns, J. A.; Butler, J. C.; Moran, J.; Whitesides, G. M. *J. Org. Chem.* **1991,** *56,* 2648–2650. doi:10.1021/jo00008a014
146. Mahenthiralingam, E.; Urban, T. A.; Goldberg, J. B. *Nat. Rev. Microbiol.* **2005,** *3,* 144–156. doi:10.1038/nrmicro1085
147. Moskowitz, S. M.; Ernst, R. K.; Miller, S. I. *J. Bacteriol.* **2004,** *186,* 575–579. doi:10.1128/JB.186.2.575-579.2004
148. Rubin, E. J.; Herrera, C. M.; Crofts, A. A.; Trent, M. S. *Antimicrob. Agents Chemother.* **2015,** *59,* 2051–2061. doi:10.1128/AAC.05052-14
149. Silipo, A.; Molinaro, A.; Ieranò, T.; De Soyza, A.; Sturiale, L.; Garozzo, D.; Aldridge, C.; Corris, P. A.; Khan, C. M. A.; Lanzetta, R. *Chem. – Eur. J.* **2007,** *13,* 3501–3511. doi:10.1002/chem.200601406
150. Molinaro, A.; Lindner, B.; De Castro, C.; Nolting, B.; Silipo, A.; Lanzetta, R.; Parrilli, M.; Holst, O. *Chem. – Eur. J.* **2003,** *9,* 1542–1548. doi:10.1002/chem.200390177
151. Zamyatina, A.; Hollaus, R.; Blaukopf, M.; Kosma, P. *Pure Appl. Chem.* **2012,** *84,* 11–21. doi:10.1351/PAC-CON-11-08-01
152. Hollaus, R.; Kosma, P.; Zamyatina, A. *Org. Lett.* **2017,** *19,* 78–81. doi:10.1021/acs.orglett.6b03358
153. Oscarson, S.; Sehgelmeble, F. W. *Tetrahedron: Asymmetry* **2005,** *16,* 121–125. doi:10.1016/j.tetasy.2004.11.051
154. Morelli, L.; Cancogni, D.; Tontini, M.; Nilo, A.; Filippini, S.; Costantino, P.; Romano, M. R.; Berti, F.; Adamo, R.; Lay, L. *Beilstein J. Org. Chem.* **2014,** *10,* 2367–2376. doi:10.3762/bjoc.10.247

155. Kajihara, K.; Arisawa, M.; Shuto, S. *J. Org. Chem.* **2008,** *73,* 9494–9496. doi:10.1021/jo801915c
156. Smith, A. B., III; Rivero, R. A.; Hale, K. J.; Vaccaro, H. A. *J. Am. Chem. Soc.* **1991,** *113,* 2092–2112. doi:10.1021/ja00006a030
157. Marshall, J. A.; Garofalo, A. W. *J. Org. Chem.* **1993,** *58,* 3675–3680. doi:10.1021/jo00066a019
158. Slättegård, R.; Teodorovic, P.; Kinfe, H. H.; Ravenscroft, N.; Gammon, D. W.; Oscarson, S. *Org. Biomol. Chem.* **2005,** *3,* 3782–3787. doi:10.1039/b507898a
159. Zamyatina, A.; Kosma, P. Synthesis of anomeric phosphates of aldoses and 2-ulosonic acids. In *Carbohydrate Chemistry. Chemical and biological approaches;* Rauter, A. P.; Lindhorst, T. K., Eds.; RSC Publishing: Cambridge, UK, 2009; pp 71–98. doi:10.1039/b901502g
160. Saito, K.; Wada, T. *Tetrahedron Lett.* **2014,** *55,* 1991–1993. doi:10.1016/j.tetlet.2014.02.013
161. Hollaus, R.; Ittig, S.; Hofinger, A.; Haegman, M.; Beyaert, R.; Kosma, P.; Zamyatina, A. *Chem. – Eur. J.* **2015,** *21,* 4102–4114. doi:10.1002/chem.201406058
162. Wada, T.; Sato, Y.; Honda, F.; Kawahara, S.-i.; Sekine, M. *J. Am. Chem. Soc.* **1997,** *119,* 12710–12721. doi:10.1021/ja9726015
163. Brett, P. J.; Burtnick, M. N.; Snyder, D. S.; Shannon, J. G.; Azadi, P.; Gherardini, F. C. *Mol. Microbiol.* **2007,** *63,* 379–390. doi:10.1111/j.1365-2958.2006.05519.x
164. Ieranò, T.; Cescutti, P.; Leone, M. R.; Luciani, A.; Rizzo, R.; Raia, V.; Lanzetta, R.; Parrilli, M.; Maiuri, L.; Silipo, A.; Molinaro, A. *Innate Immun.* **2010,** *16,* 354–365. doi:10.1177/1753425909347400
165. De Soyza, A.; Ellis, C. D.; Khan, C. M. A.; Corris, P. A.; de Hormaeche, R. D. *Am. J. Respir. Crit. Care Med.* **2004,** *170,* 70–77. doi:10.1164/rccm.200304-592OC
166. Di Lorenzo, F.; Sturiale, L.; Palmigiano, A.; Fazio, L. L.; Paciello, I.; Coutinho, C. P.; Sá-Correia, I.; Bernardini, M.; Lanzetta, R.; Garozzo, D.; Silipo, A.; Molinaro, A. *ChemBioChem* **2013,** *14,* 1105–1115. doi:10.1002/cbic.201300062
167. Gaekwad, J.; Zhang, Y.; Zhang, W.; Reeves, J.; Wolfert, M. A.; Boons, G.-J. *J. Biol. Chem.* **2010,** *285,* 29375–29386. doi:10.1074/jbc.M110.115204
168. Silipo, A.; Molinaro, A.; Cescutti, P.; Bedini, E.; Rizzo, R.; Parrilli, M.; Lanzetta, R. *Glycobiology* **2005,** *15,* 561–570. doi:10.1093/glycob/cwi029
169. Silipo, A.; Molinaro, A.; Comegna, D.; Sturiale, L.; Cescutti, P.; Garozzo, D.; Lanzetta, R.; Parrilli, M. *Eur. J. Org. Chem.* **2006,** 4874–4883. doi:10.1002/ejoc.200600520
170. Adinolfi, M.; Barone, G.; Guariniello, L.; Iadonisi, A. *Tetrahedron Lett.* **2000,** *41,* 9305–9309. doi:10.1016/S0040-4039(00)01673-7
171. Bannwarth, W.; Küng, E. *Tetrahedron Lett.* **1989,** *30,* 4219–4222. doi:10.1016/S0040-4039(01)80694-8
172. Davis, F. A.; Chattopadhyay, S.; Towson, J. C.; Lal, S.; Reddy, T. *J. Org. Chem.* **1988,** *53,* 2087–2089. doi:10.1021/jo00244a043
173. Müller, B.; Blaukopf, M.; Hofinger, A.; Zamyatina, A.; Brade, H.; Kosma, P. *Synthesis* **2010,** 3143–3151. doi:10.1055/s-0030-1258174
174. Stawinski, J.; Kraszewski, A. *Acc. Chem. Res.* **2002,** *35,* 952–960. doi:10.1021/ar010049p
175. Oka, N.; Shimizu, M.; Saigo, K.; Wada, T. *Tetrahedron* **2006,** *62,* 3667–3673. doi:10.1016/j.tet.2006.01.084
176. de Vroom, E.; Dreef, C. E.; van den Elst, H.; van der Marel, G. A.; van Boom, J. H. *Recl. Trav. Chim. Pays-Bas* **1988,** *107,* 592–595. doi:10.1002/recl.19881071004
177. Davis, F. A.; Towson, J. C.; Weismiller, M. C.; Lal, S.; Carroll, P. J. *J. Am. Chem. Soc.* **1988,** *110,* 8477–8482. doi:10.1021/ja00233a025
178. Hayakawa, Y.; Kato, H.; Uchiyama, M.; Kajino, H.; Noyori, R. *J. Org. Chem.* **1986,** *51,* 2400–2402. doi:10.1021/jo00362a052
179. Tanaka, S.; Saburi, H.; Kitamura, M. *Adv. Synth. Catal.* **2006,** *348,* 375–378. doi:10.1002/adsc.200505401
180. Tanaka, S.; Saburi, H.; Murase, T.; Ishibashi, Y.; Kitamura, M. *J. Organomet. Chem.* **2007,** *692,* 295–298. doi:10.1016/j.jorganchem.2006.03.046

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Synthetic and semi-synthetic approaches to unprotected *N*-glycan oxazolines

Antony J. Fairbanks[1,2]

Review

Open Access

Address:
[1]Department of Chemistry, University of Canterbury, Private Bag 4800, Christchurch, 8140, New Zealand and [2]Biomolecular Interaction Centre, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand

Email:
Antony J. Fairbanks - antony.fairbanks@canterbury.ac.nz

Keywords:
DMC, ENGase; glycosyl oxazolines; *N*-glycans; oligosaccharides

Beilstein J. Org. Chem. **2018,** *14,* 416–429.
doi:10.3762/bjoc.14.30

Received: 17 November 2017
Accepted: 31 January 2018
Published: 15 February 2018

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

N-Glycan oxazolines have found widespread use as activated donor substrates for endo-β-*N*-acetylglucosaminidase (ENGase) enzymes, an important application that has correspondingly stimulated interest in their production, both by total synthesis and by semi-synthesis using oligosaccharides isolated from natural sources. Amongst the many synthetic approaches reported, the majority rely on the fabrication (either by total synthesis, or semi-synthesis from locust bean gum) of a key Manβ(1–4)GlcNAc disaccharide, which can then be elaborated at the 3- and 6-positions of the mannose unit using standard glycosylation chemistry. Early approaches subsequently relied on the Lewis acid catalysed conversion of peracetylated *N*-glycan oligosaccharides produced in this manner into their corresponding oxazolines, followed by global deprotection. However, a key breakthrough in the field has been the development by Shoda of 2-chloro-1,3-dimethylimidazolinium chloride (DMC), and related reagents, which can direct convert an oligosaccharide with a 2-acetamido sugar at the reducing terminus directly into the corresponding oxazoline in water. Therefore, oxazoline formation can now be achieved in water as the final step of any synthetic sequence, obviating the need for any further protecting group manipulations, and simplifying synthetic strategies. As an alternative to total synthesis, significant quantities of several structurally complicated *N*-glycans can be isolated from natural sources, such as egg yolks and soy bean flour. Enzymatic transformations of these materials, in concert with DMC-mediated oxazoline formation as a final step, allow access to a selection of *N*-glycan oxazoline structures both in larger quantities and in a more expedient fashion than is achievable by total synthesis.

Review

Introduction

Glycosyl oxazolines are high-energy intermediates on the hydrolytic pathway of some [1-5] (but not all) [6] of the numerous glycosidases that hydrolyse linkages between 2-acetamido sugars and other species. In particular the endo-β-*N*-acetylglucosaminidases [7] (ENGases, EC 3.2.1.96), a class of enzyme which specifically cleave between the innermost two GlcNAc residues of *N*-glycans attached to N-linked glycoproteins, all operate via a two-step mechanism involving neighbouring

group participation of the 2-acetamide group and an oxazoline as a high energy intermediate [8].

Glycosyl oxazolines first drew the attention of synthetic chemists due to their use as glycosyl donors for the synthesis of oligosaccharides that comprise 2-amino-2-dexoy sugars [9]. Though the majority of synthetic work focussed on production and reaction of *gluco*-configured oxazolines (i.e., those derived from GlcNAc), the corresponding *manno* [10,11] and *galacto*-configured [12] compounds have also been made and studied. Although the first generation of these oxazoline donors [13,14] proved to be rather unreactive, and found only limited applications [15-18], the addition of three chlorines to the methyl group did increase their potency [19-23]. However, applications were still less widespread than more conventional glucosamine-derived donors.

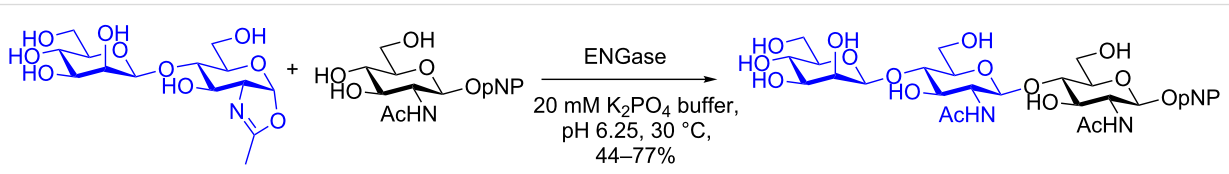
Resurgent interest in the production of glycosyl oxazolines, and in particular oxazoline derivatives of *N*-glycans, was as a direct result of their utility as activated donors species for glycosidase-catalysed synthesis [24-27]. Initially activity centred on the use of oxazolines as donors for chitinase-catalysed glycosylations [28-31]. However, a turning point occurred when, in a seminal publication in 2001 Shoda [32] and co-workers reported that a disaccharide oxazoline (Scheme 1) was an effective donor substrate for two ENGase enzymes (Endo A and Endo M), both of which were capable of using it to glycosylate two GlcNAc acceptors, to produce trisaccharide products.

Subsequently the ENGases in combination with *N*-glycan oxazolines, have become the biocatalysts of choice for the convergent production of a wide variety of biologically interesting glycopeptides and for the remodelling of glycoproteins, including mAbs [33,34]. The efficient production of *N*-glycan oxazolines as donor substrates for these enzymes has therefore become an area of significant interest over the past 15 years [35,36].

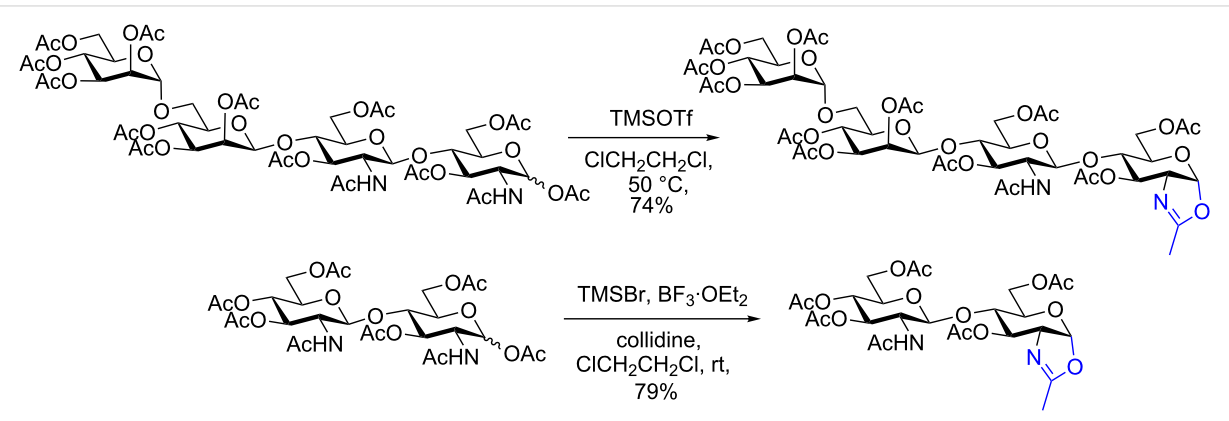
The synthesis of *N*-glycan oxazolines

Formation of glycosyl oxazolines

Glycosyl oxazolines of monosaccharides can be produced straightforwardly using strong Lewis acids (e.g., $FeCl_3$, $SnCl_4$, or TMSOTf) and a fully protected (typically peracetylated) GlcNAc or other 2-acetamido sugar [37-40]. Oxazoline formation is achieved by activation of the leaving group at the anomeric centre and neighbouring group participation by the 2-acetamide. Unfortunately application of these reaction conditions to oligosaccharide substrates leads to significant cleavage of interglycosidic linkages, and correspondingly low yields of products. However, two methods that are useful for the production of oligosaccharide oxazolines are treatment of the peracetylated sugar with either TMSOTf in dichloroethane [39], or with TMSBr, $BF_3{\cdot}Et_2O$ and 2,4,6-collidine in dichloroethane [40] (Scheme 2). Both procedures give oxazolines of *N*-glycans in moderate to good yield with no cleavage of the oligosaccharide chain; the latter method reportedly gives better yields of more structurally complex *N*-glycan oxazolines.



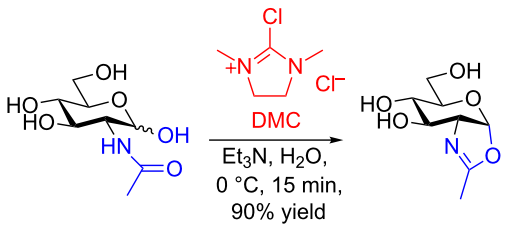
Scheme 1: The first ENGase-catalysed glycosylation of a GlcNAc acceptor using an *N*-glycan oxazoline as donor.



Scheme 2: Production of *N*-glycan oxazolines from peracetylated sugars using Lewis acids.

However, employing protected sugars as substrates presents some limitations, as any remaining protecting groups must be removed in a subsequent step. Firstly, and most importantly, glycosyl oxazolines are extremely labile to acidic hydrolysis, and so this approach precludes the use of any OH-protecting groups that require acidic conditions for their cleavage. Secondly some glycosyl oxazolines are also prone to reductive cleavage by catalytic hydrogenation [41], presenting a significant further limitation as to which OH-protecting groups may be employed. Most of the reports in the literature have therefore used a protecting group regime in which all of the sugar hydroxy groups have been protected with base-labile groups, most commonly acetate esters. Importantly glycosyl oxazolines are completely stable to the typical basic conditions used for ester removal (e.g., Zemplen deacetylation). The generally accepted approach (until 2009) was therefore to perform all protecting group manipulations/interconversions on the completed oligosaccharide to ensure that all OH groups were protected as base-labile esters, before oxazoline formation.

In 2009, Shoda published [42] a paper that was to completely change the way in which glycosyl oxazolines were made, and which would ultimately make many more readily available. In this seminal work, Shoda reported that the treatment of GlcNAc in aqueous solution with the activating agent 2-chloro-1,3-dimethylimidazolinium chloride (DMC) in the presence of triethylamine as the base, led to the formation of the glycosyl oxazoline in good yield (Scheme 3). Moreover this remarkable transformation was equally applicable to considerably larger oligosaccharide structures (vide infra). This breakthrough changed the way that all unprotected *N*-glycan oxazolines were to be made from that point in time onwards.



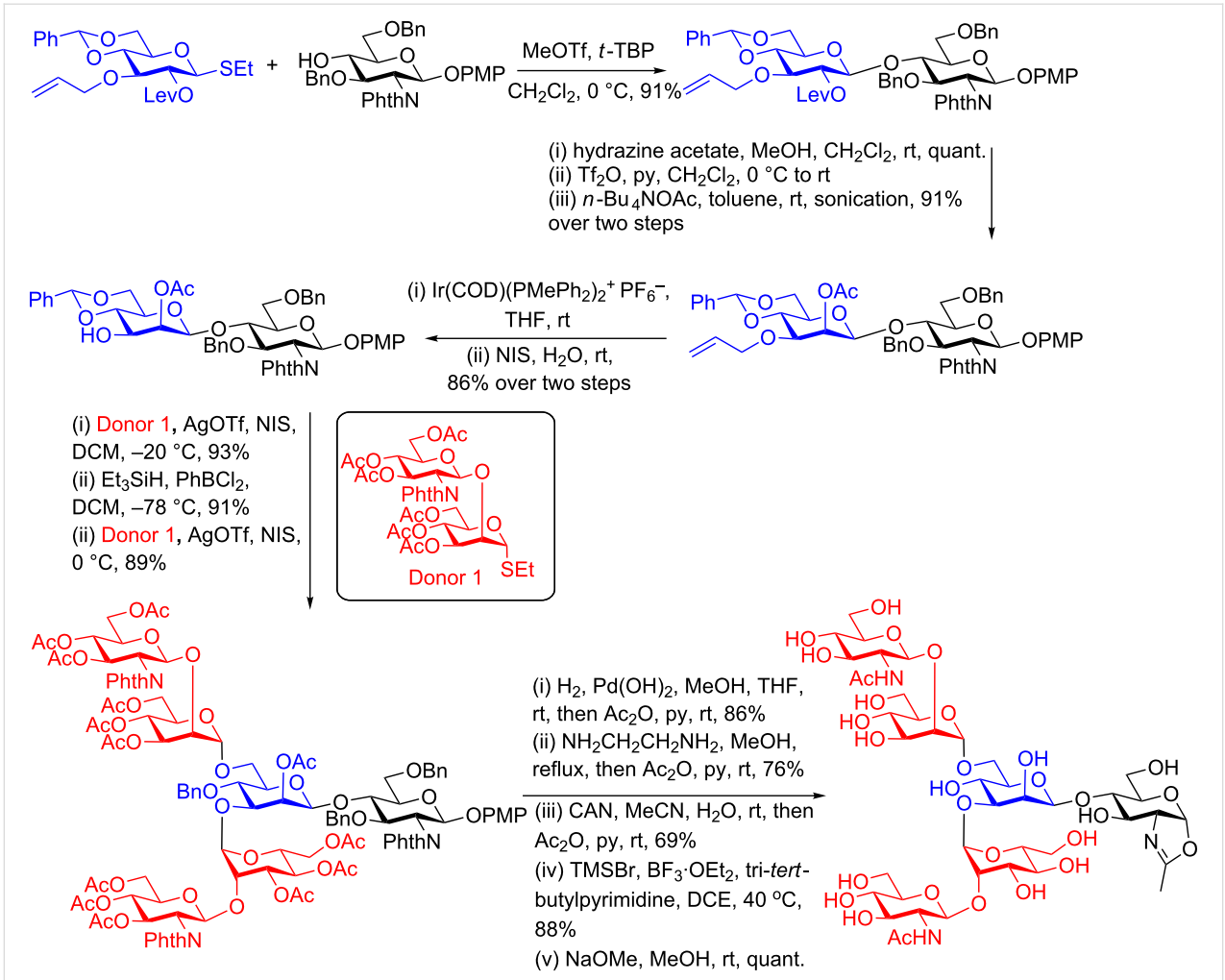
Scheme 3: Direct conversion of unprotected GlcNAc to a glycosyl oxazoline by treatment with DMC and Et_3N in water.

Although in later papers Shoda has published alternative reagents that may be used to achieve the same transformation, such as 2-chloro-1,3-dimethyl-1*H*-benzimidazol-3-ium chloride (CDMBI) [43], DMC remains the most popular reagent for glycosyl oxazoline production. DMC is remarkably tolerant of other functional groups in the oligosaccharide, for example sialic acids [44] and phosphates [45,46] are completely unaffected; the former is perhaps rather surprising since DMC was first developed as a carboxylic acid activating agent for peptide synthesis by Ishikawa [47]! One caveat to the procedure is that it is considerably less efficient for GalNAc; in this case the corresponding oxazoline is only produced in ≈50% yield. Indeed some of the other very useful DMC-mediated transformations of unprotected reducing sugars in aqueous solution that have been developed recently also work less effectively when the sugar at the reducing terminus has a *galacto* configuration [48-50].

Production of unprotected *N*-glycan oxazolines by total synthesis

The majority of the reported syntheses of *N*-glycan oxazolines have employed a key selectively protected Manβ(1–4)GlcNAc disaccharide building block which has then been extended at the 3- and 6-positions of the branching mannose unit. Amongst the possible ways to synthesise this key disaccharide [51,52] two have been used predominantly for the synthesis of *N*-glycan oxazolines. The OH-2 epimerisation approach, which uses a *gluco*-configured donor for glycosylation of the OH-4 of a selectively protected glucosamine acceptor has been used more than the other methods. Selective and orthogonal protection of OH-2 of the donor by an ester group facilitates both the stereoselective formation of the desired β-linkage, and also access to OH-2 after glycosylation for epimerisation. Amongst the many syntheses [53-58] of *N*-glycan oxazolines using this approach, the use of Lev protection on the donor, first developed by Boons [59], and then triflation and nucleophilic substitution by acetate aided by sonication, first developed by Fürstner [60,61], appear to be optimal. An example that employed these key steps was used to synthesise a truncated complex biantennary *N*-glycan oxazoline [62], as shown in Scheme 4. Following the *gluco* to *manno* epimerisation process, selective deprotection of OH-3 of the mannose unit was followed by glycosylation and extension of the 3-branched arm. Subsequent removal (or regioselective reductive ring-opening) of the 4,6-benzylidene protecting group allowed a second glycosylation at position 6. Conversion of all OH-protecting groups to acetate and the phthalamide to acetamide was followed by oxazoline formation using TMSBr, $BF_3{\cdot}Et_2O$ and 2,4,6-collidine in dichloroethane, and finally deacetylation. Modifications of this basic strategy have allowed the synthesis of a wide variety of truncated and structurally modified glycans [53-59].

Amongst other synthetic approaches that may be used to access the 'difficult' Manβ(1–4)GlcNAc linkage, including a variety of methods of intramolecular glycosylation [63-71] the most widely applied has been the Crich direct β-mannosylation [72-76]. However, one apparent limitation is that generally the reaction only works well if the GlcNAc acceptor has an azide or

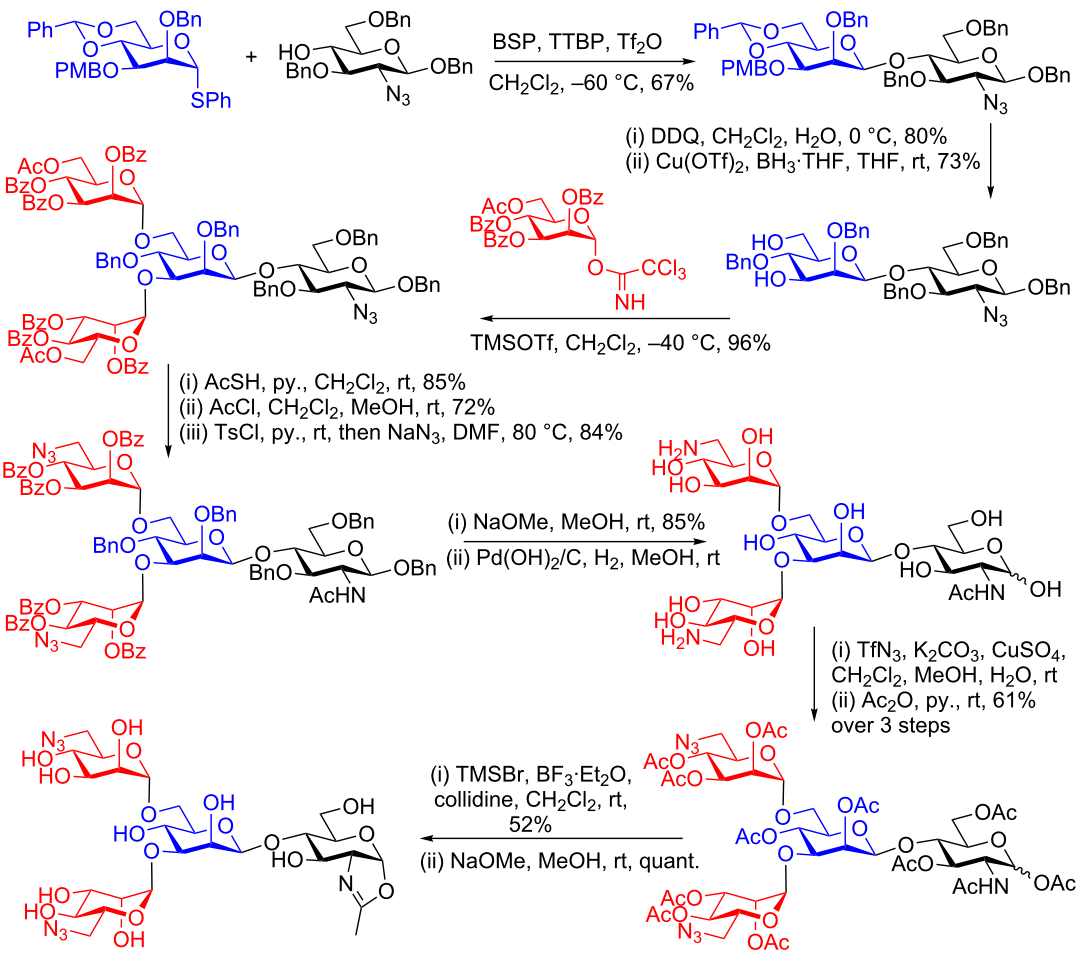


Scheme 4: Total synthesis of a truncated complex biantennary *N*-glycan oxazoline via an epimerisation approach and Lewis acid mediated oxazoline formation.

sulfonamide at position 2, rather than acetamide or *N*-phthalamide. Scheme 5 shows an example of the synthesis of a modified core *N*-glycan tetrasaccharide oxazoline from the several reported by Wang [77] using this approach. In this case following formation of the key Manβ(1–4)GlcNAc disaccharide both the 3- and 6-hydroxy groups of the mannose residue were deprotected, and the resulting diol underwent a double glycosylation with a selectively protected trichloroacetimidate donor.

An added advantaged of approaches that use total synthesis is the possibility of the incorporation of tags into the glycan structure, which allows further modifications to be made later. In this case, following conversion of the azide at position 2 of the glucosamine unit into an acetamide, azide was introduced at position 6 of the two terminal mannose residues. Protecting group interconversions, and peracetylation were followed by conversion to the oxazoline, using TMSBr, $BF_3{\cdot}Et_2O$ and collidine, and finally deacetylation. It was found that the incorporated azide was tolerated by the ENGase enzyme (Endo A), and so a modified glycoprotein (RNase) was made by enzymatic attachment of this synthetic tetrasaccharide, to which other species were then conjugated by click reactions.

In more recent examples conversion of the completely deprotected glycan to the oxazoline by treatment with DMC has become the normal (and most effective) strategy. For example the same key Manβ(1–4)GlcNAc disaccharide was used by Wang for the more extended synthesis of a dodecasaccharide oxazoline (Scheme 6) [78]. In this case selective removal of the PMB protecting group at OH-3 was followed by glycosylation with a pentasaccharide glycosyl fluoride donor, comprising one galactose, one glucose, and three mannose residues. Acid catalysed hydrolysis of the 4,6-benzylidene was followed by regioselective glycosylation of the primary 6-OH with a different pentasaccharide, this time comprised of five mannoses. Conver-



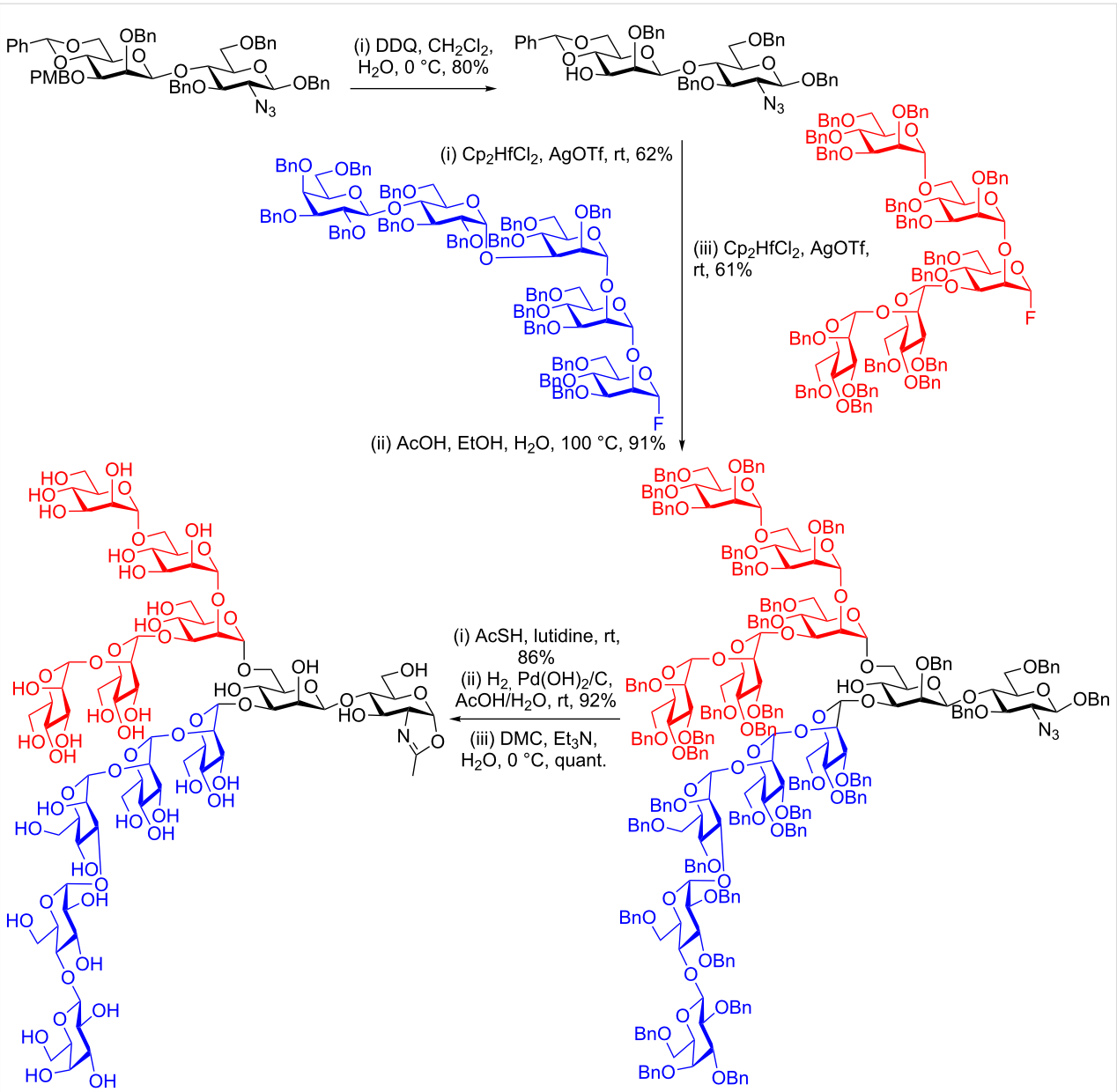
Scheme 5: Wangs's total synthesis of an *N*-glycan oxazoline incorporating click handles, employing Crich direct β-mannosylation.

sion of the azide to acetamide and removal of all benzyl groups by hydrogenolysis produced a completely deprotected dodecasaccharide, which was finally converted to the glycosyl oxazoline by treatment with DMC in quantitative yield.

Final stage conversion of the fully deprotected oligosaccharide into the oxazoline has greatly facilitated the synthesis of more complex *N*-glycan oxazolines by analogous routes [79,80], including those bearing mannose-6-phosphate residues [45,46]. For example as shown in Scheme 7 sequential glycosylation of the key Manβ(1–4)GlcNAc disaccharide at positions 3 and 6, using the same selectively protected *manno* thioglycoside donor gave a tetrasaccharide. Removal of the silyl protecting groups revealed the 6-hydroxy groups of the terminal mannose residues, which were then phosphorylated. Removal of the anomeric PMP protection was followed by global deprotection by Birch reduction to give the completely deprotected tetrasaccharide diphosphate. Finally treatment with DMC in water in the presence of Et_3N resulted in conversion to the glycosyl oxazoline in an excellent 95% yield.

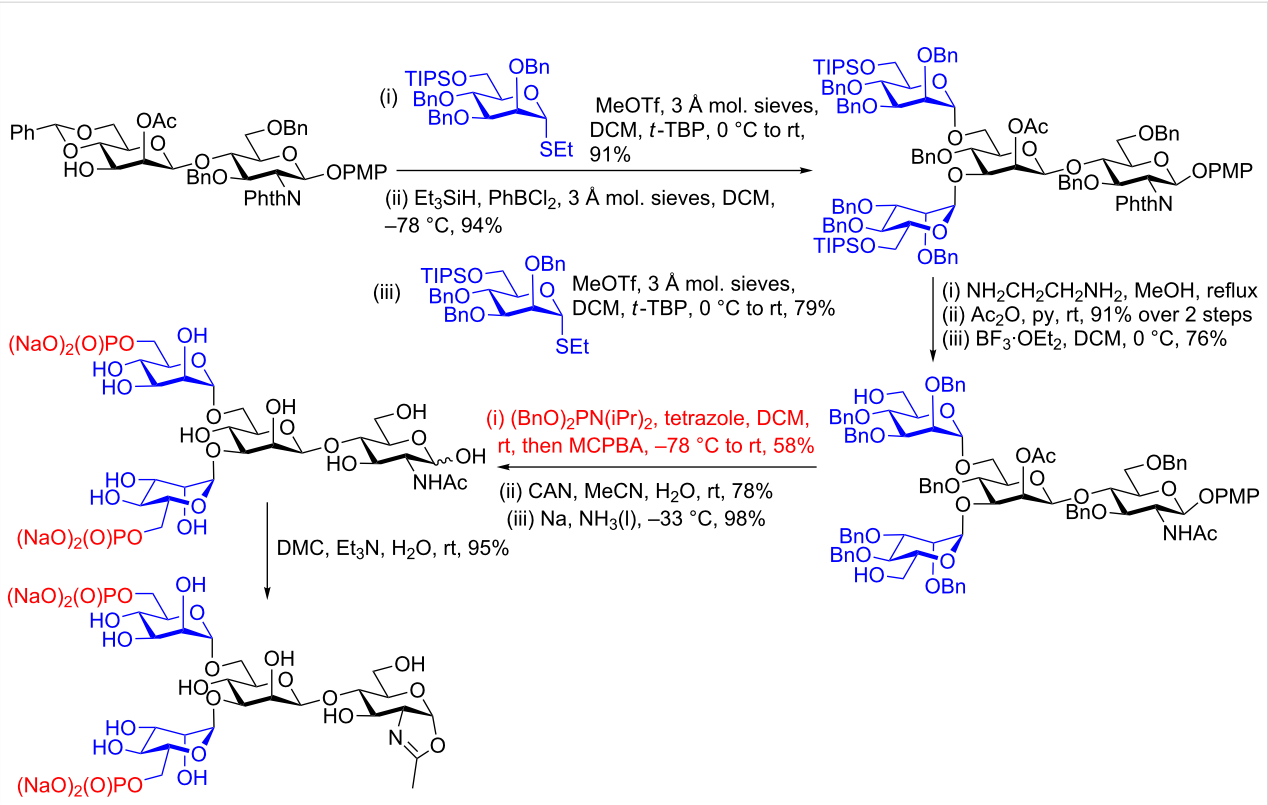
Semi-synthesis: the locust bean gum approach

The naturally occurring polysaccharide locust bean gum contains a repeating Manβ(1–4)Man disaccharide unit, which is also decorated with branching α-galactose residues attached to OH-6 of some of the mannoses. Nishimura and co-workers [81] realised the potential utility of this Manβ(1–4)Man disaccharide in an expedient route to the part of *N*-glycans that is most difficult to synthesise; namely the Manβ(1–4)GlcNAc linkage. Treatment of locust bean gum with pectinase from *Aspergillus aculeatus*, which has both mannosidase and galactosidase activity, at 50 °C for 48 h in a 50 mM acetate buffer (pH 5.0) resulted in the production of a mixture of compounds, from which the Manβ(1–4)Man disaccharide was readily purified by acetylation (typically in ≈30% overall yield, Scheme 8). In the key transformation, the mannose residue at the reducing terminus was then converted into a glucosamine derivative (in fact possessing an azide at C2) first by conversion to the glycal and then an azido nitration reaction. This innovative method is considerably shorter than other approaches to the Manβ(1–4)GlcNAc (or equivalent) disaccharide. Elegant



Scheme 6: Wangs's total synthesis of an *N*-glycan dodecasaccharide oxazoline employing final step oxazoline formation with DMC.

protecting group manipulations, involving the formation of a dibenzylidene derivative on the mannose ring, benzylation of the remaining free hydroxy groups on the glucosamine ring, and then regio- and chemoselective reductive ring opening of the less stable 5-ring benzylidene with DIBAL, led to a key disaccharide intermediate in which OH-3 of the mannose unit was unprotected and in which the 4- and 6-positions were protected as a benzylidene. Extension of this core disaccharide should be straightforward by traditional synthetic methodology, and so in principle the locust bean gum approach should allow rapid access to a wide variety of more extended *N*-glycan structures. In their original publication Nishimura and co-workers first glycosylated the free OH at position 3 with 2,4-branched trisaccharide trichloroacetimidate donor 2, removed the 4,6-benzylidene, and then regioselectively glycosylated the free primary OH at position 6 with 2,6-branched trisaccharide trichloroacetimidate donor 3. Following conversion of the Troc groups into acetamides and reduction and acetylation of the azide, all of the acetates were removed. Treatment with UDP-Gal and a β(1–4)-galactosyl transferase led to the addition of galactose residues to all of the 4-hydroxy groups of the GlcNAcs. Deprotection of the remaining benzyl protecting groups and removal of the SPh at the reducing terminus by catalytic hydrogenation gave the completely deprotected dodecasaccharide. Finally conversion to



Scheme 7: Production of a phosphorylated *N*-glycan oxazoline, employing final step oxazoline formation with DMC.

the corresponding oxazoline by the use of DMC gave the tetraantennary complex *N*-glycan oxazoline in 96% yield (Scheme 8).

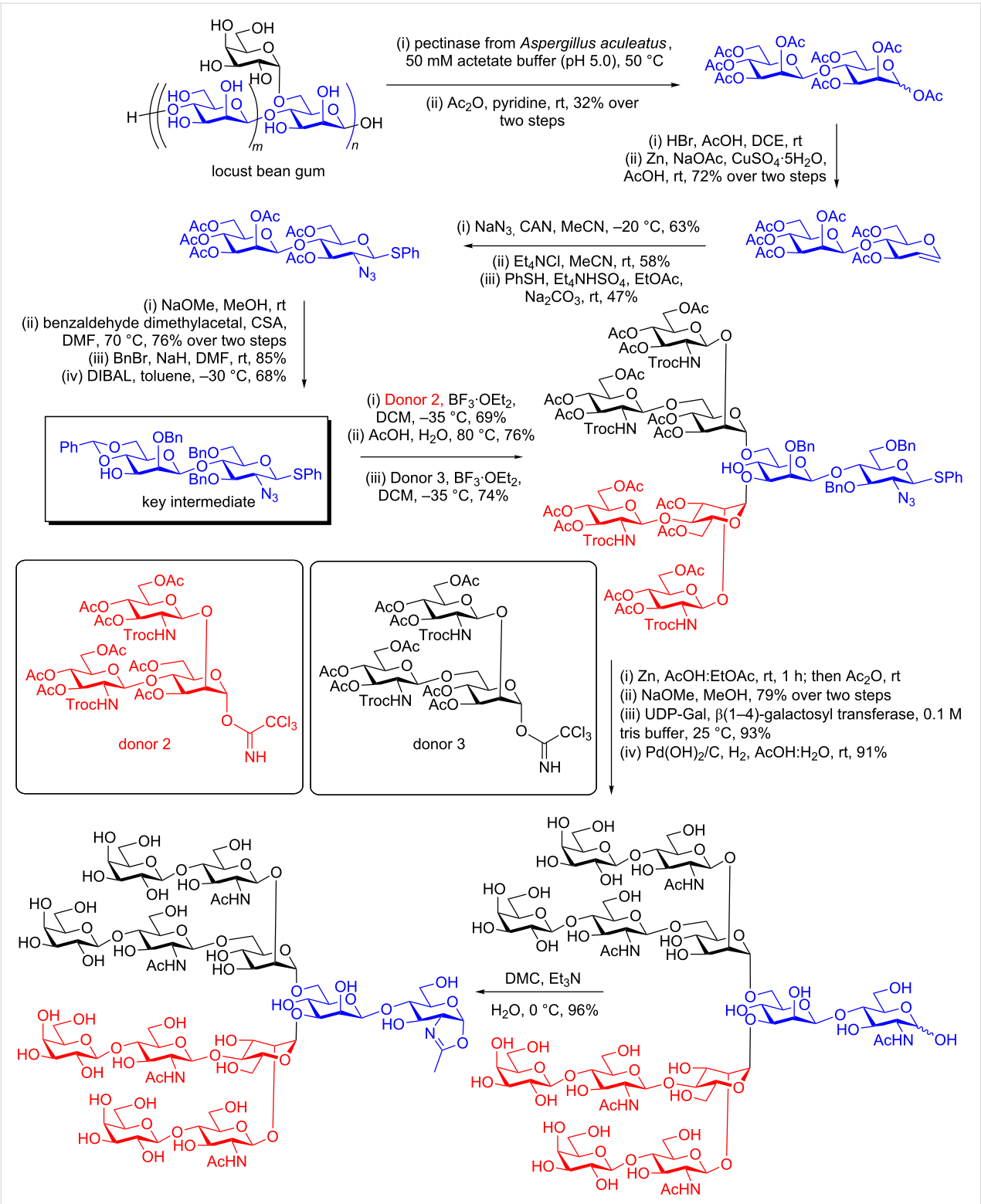
Production of *N*-glycan oxazolines using oligosaccharides isolated from natural sources

The egg yolk approach

The yolk of hens' eggs contains a glycopeptide, often termed sialylglycopeptide (SGP), which is comprised of a short peptide linked to a complex biantennary *N*-glycan. Thus egg yolks can serve as a source of this complex biantennary *N*-glycan [$(NeuAcGalGlcNAcMan)_2ManGlcNAc_2$], following isolation of SGP, and subsequent enzymatic degradation. The original procedure [82] for the isolation of SGP first involved deproteinization by treatment with 90% phenol and washing with Et_2O, and then repeated purification by size exclusion chromatography (SEC, Sephadex G-50, followed by Sephadex G-25) from which sialic acid positive fractions were collected. Further purification by anion exchange chromatography (Sephadex DEAE eluting with NaCl) removed any non-sialylated glycans, and was followed by cation exchange chromatography (Sephadex C-25). Finally desalination using SEC (Sephadex G-25) gave pure SGP.

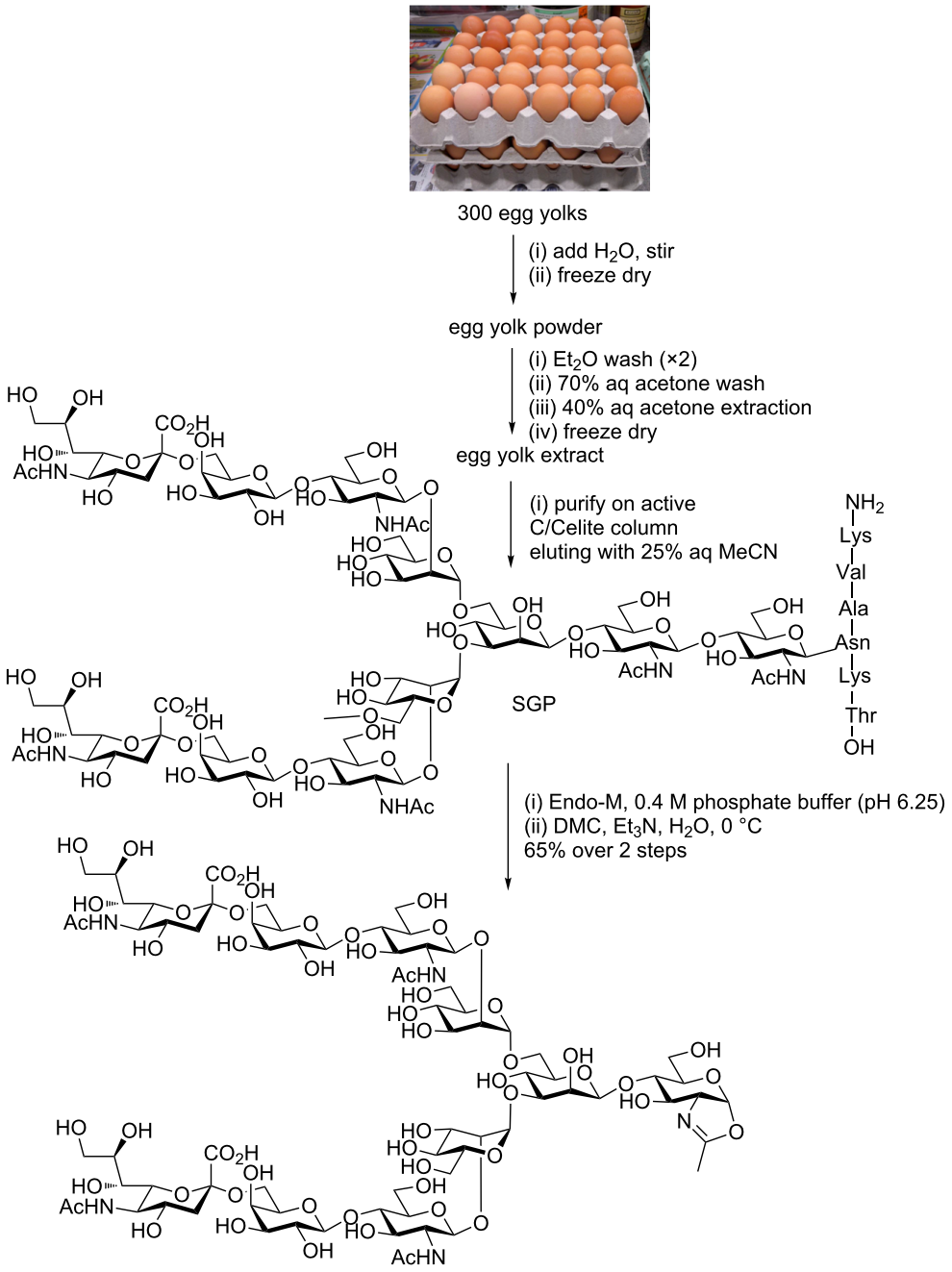
Several improvements have subsequently been published which have made the isolation process easier and improved the yield. Firstly a significantly shortened procedure [83] followed the phenol treatment with a single purification by SEC (Sephadex G-50), and then filtration through graphitized carbon cartridges. Subsequently an even better method was developed [84] which avoided the treatment with phenol and all SEC purification steps (Scheme 9). In this process the egg yolks were first stirred with water and then freeze dried to give egg yolk powder. This powder was washed successively with diethyl ether and then 70% aqueous acetone. The solid was then extracted by vigorous mixing with 40% aqueous acetone. Following filtration through Celite®, the filtrate was concentrated and freeze-dried. The powder was dissolved in water and then purified on an active carbon/Celite® (2:1) column, eluting with 25% MeCN, to give pure SGP on a gram scale; typically 1.5–2.0 g of SGP is obtained from 300 eggs.

Very recently Boons and co-workers [85] published further modifications and optimisation of this procedure, and reported that it is possible to start with commercially produced lyophilised egg yolk powder, rather than the eggs themselves. Their method, which also included purification by the use of preparative hydrophilic interaction chromatography–high per-



Scheme 8: Enzymatic degradation of locust bean gum, and chemical conversion into an *N*-glycan dodecasaccharide oxazoline.

formance liquid chromatography (HILIC–HPLC), clearly reduces time and effort by removing the need for separation of the yolks and freeze-drying. However, care has to be exercised with respect to the processing that the commercially sourced egg yolk powder has undergone; for example spray drying at >100 °C may lead to degradation of the glycans.



Scheme 9: Production of a complex biantennary *N*-glycan oxazoline from hens' eggs by semi-synthesis via isolation of SGP, enzymatic degradation, and final stage oxazoline formation.

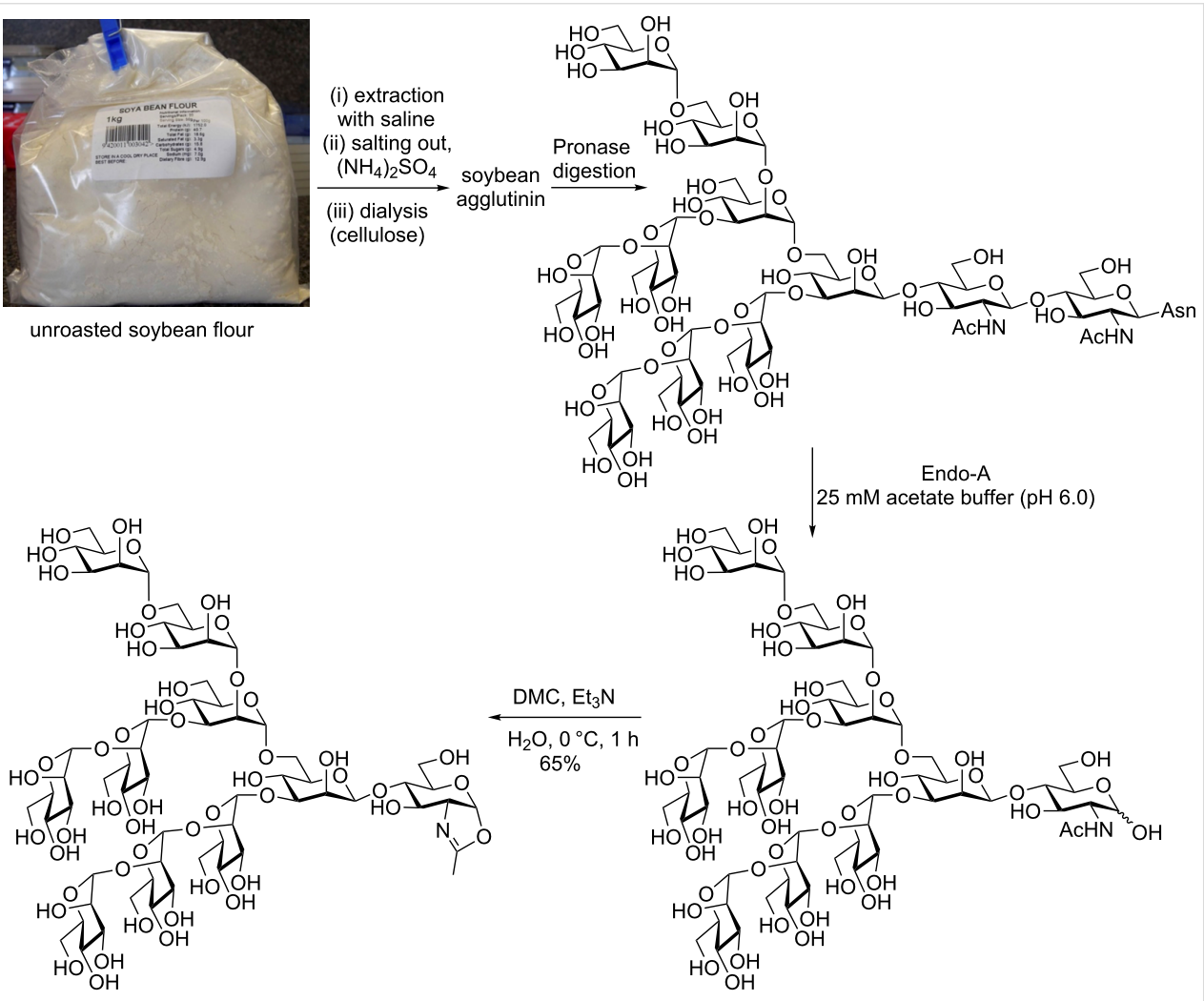
Whichever method of SGP production is used, the free oligosaccharide [(NeuAcGalGlcNAcMan)$_2$ManGlcNAc] can then be released from SGP by treatment with the ENGase Endo M [86]. Following purification by SEC (Sephadex G-25), the free glycan can be converted into the oxazoline by treatment with DMC in water (Scheme 9), as first reported by Wang [87] and Umekawa and co-workers [88], and then subsequently used by others [44,89,90]. Shoda's modified version of DMC (2-chloro-1,3-dimethyl-1*H*-benzimidazol-3-ium chloride, CDMBI) has also been reported to be efficient at this transformation [91]. Furthermore removal of the terminal sialic acid residues of the free oligosaccharide by treatment with a neuraminidase allows the production of truncated complex biantennary glycans. Originally Wang and co-workers reported [92] the synthesis of this type of oxazoline using a sequence of acetylation, treatment with TMSBr/BF_3·Et_2O/collidine and deacetylation. However,

treatment of the free reducing sugar with DMC allows the production of the truncated complex *N*-glycan oxazoline in a more efficient manner [93].

In related work, Kajihara has recently developed [94] methods that allow selective modification of the complex biantennary *N*-glycan available from egg yolks. For example, after the peptide is degraded to a single Asn residue by protease digestion (Actinase E), the sialic acids can be removed by acidic hydrolysis [95], and the amine Fmoc protected. Branch specific exo-glycosidase digestion then allows the production of a wide variety of truncated glycans. Alternatively, by forming 4,6-benzylidenes of the mannose and galactose residues, acetylating all the remaining free OH groups, and then using mild acidic hydrolysis (60% aqueous acetic acid), Kajihara was able to produce a mixture of products in which either one or both of the mannose residues had been deprotected but the galactose residues remained completely protected. HPLC separation then allowed either selective chemical glycosylation or protection of the primary OH groups; the remaining secondary hydroxy group of the products of the latter process could also be glycosylated. Ultimately this methodology allows the synthesis of the considerably more complex *N*-glycans, for example tri- (and presumably in the future tetra-) antennary glycans, starting from SGP. Although the protecting group-based reactions lack complete selectivity and the sequences require several careful HPLC separations, the fact that the complex biantennary glycan is so readily available still makes these approaches attractive with respect to total synthesis. Neither Kajihara nor others have yet to employ these routes to the production of *N*-glycan oxazolines.

The soy bean approach

Soy bean agglutinin is a glycoprotein decorated with high mannose glycans [96]. Isolation of soy bean agglutinin from unroasted soy bean flour is achieved by acidification (pH 4.6),

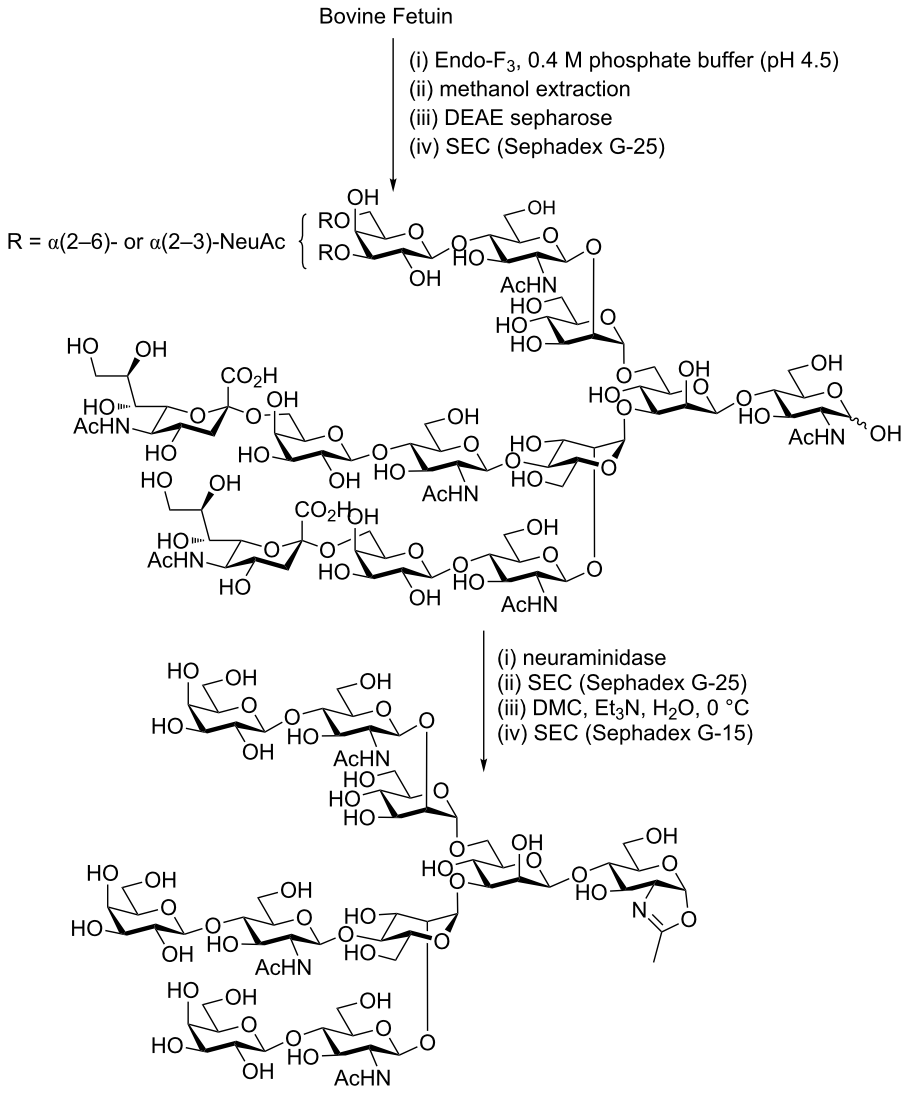


Scheme 10: Production of a high mannose (Man-9) *N*-glycan oxazoline from soy bean flour.

and salting out with ammonium sulphate [97,98]. The Asn-linked Man-9 glycan ($Man_9GlcNAc_2Asn$) can then be prepared [99] by exhaustive Pronase digestion, followed by SEC (Sephadex G-50) and further purification with HPLC on a graphitized carbon column (Scheme 10) [100]. Alternatively the glycan may also be released by hydrazinolysis [101]. Following isolation of the full-length ($Man_9GlcNAc_2$) Asn-linked glycan the truncated glycan ($Man_9GlcNAc$) can be produced by treatment with the ENGase Endo A [102], and purification by SEC (Sephadex G-15). The first production of the Man-9 oxazoline reported by Wang [103] then involved complete acetylation of this decasaccharide, treatment with $TMSBr/BF_3{\cdot}Et_2O$/collidine, and a final deacetylation step. However, this method can now be simplified by use of the Shoda DMC procedure by which the unprotected $Man_9GlcNAc$ glycan can be directly converted to the oxazoline in water [104]. This route, although lower yielding than the corresponding egg yolk procedure, is still considerably more efficient than using total synthesis to make such a highly complex decasaccharide.

Other routes to *N*-glycan oxazolines

Recently Wang and co-workers have reported [105] a semi-synthetic route to triantennary *N*-glycan oxazolines starting from bovine fetuin (Scheme 11). To enable large-scale production they first purified bovine fetuin from fetal bovine serum [106]. The *N*-glycans were then released by treatment with the ENGase Endo-F3 [107], and were partially purified by acetone precipitation and extraction with 60% methanol. The crude *N*-glycans were found to be a rather complex mixture of compounds, the four major components of which were identified as triantennary glycans with 2 or 3 sialic acids attached as regio-isomers (i.e., both α(2–6)- and α(2–3)-linked to the 6-branched



Scheme 11: Production of a triantennary *N*-glycan oxazoline from bovine feruin by semi-synthesis.

mannose arm). Isolation of these major compounds and treatment with a neuraminidase then produced essentially a single product, which was purified by SEC (Sephadex G-25), and finally converted to the corresponding triantennary oxazoline by treatment with DMC in water. Final purification was achieved by SEC (Sephadex G-15).

Commercially available chicken ovalbumin can be used as a source of high mannose *N*-glycans, from which glycopeptides can be obtained by pronase digestion [108]. A mixture of free truncated glycans ($Man_5GlcNAc$ and $Man_6GlcNAc$) can then be released [109] by the use of the ENGase Endo A, and separated by careful chromatography on a carbon-Celite® column. Although Wang and co-workers did not at that time report the conversion of these glycans into the corresponding oxazolines, the basic $Man_5GlcNAc$ structure has since been extended using a sequence of glycosyl transferases (namely a β(1–2)-GlcNAc transferase, a β(1–4)-galactosyltransferase, and an α(2–6)-sialyltransferase), and the corresponding hybrid *N*-glycan oxazoline used as a substrate for ENGases [110].

Conclusion

N-Glycan oxazolines have found widespread use as activated donor substrates for ENGase enzymes, a factor which has in turn stimulated interest in their production both by total synthesis and semi-synthesis. By far the most significant recent breakthrough in the field has been the development by Shoda of DMC (and related reagents), which can effect the direct conversion of oligosaccharides with a 2-acetamido sugar at the reducing terminus directly into the corresponding glycosyl oxazoline in water. This 'game-changer' means that nowadays no protecting group manipulations are required after oxazoline formation, which is performed as the final step; this makes production by total synthesis considerably easier. Additionally the remarkable ability of DMC and related reagents to achieve this key transformation also facilitates the use of naturally derived oligosaccharides as useful sources of *N*-glycan oxazolines. Recent work has both simplified the isolation of such *N*-glycans from natural sources, such as egg yolks, and also extended the variety of structures available by such means. It seems likely that more *N*-glycans will become available by such methods in the future. Furthermore it also appears to be only a matter of time until homogeneous glycoproteins and other glycoconjugates produced using *N*-glycans oxazolines find therapeutic and other applications; a development which will further stimulate the search for even better methods for their large scale and cost-effective production.

Acknowledgements

The author thanks the Bimolecular Interaction Centre (BIC) at the University of Canterbury for financial support.

ORCID® iDs

Antony J. Fairbanks - https://orcid.org/0000-0002-9975-3269

References

1. Terwisscha van Scheltinga, A. C.; Armand, S.; Kalk, K. H.; Isogai, A.; Henrissat, B.; Dijkstra, B. W. *Biochemistry* **1995,** *34,* 15619–15623. doi:10.1021/bi00048a003
2. Tews, I.; Terwisscha van Scheltinga, A. C.; Perrakis, A.; Wilson, K. S.; Dijkstra, B. W. *J. Am. Chem. Soc.* **1997,** *119,* 7954–7959. doi:10.1021/ja970674i
3. Drouillard, S.; Armand, S.; Davies, J. G.; Vorgias, E. C.; Henrissat, B. *Biochem. J.* **1997,** *328,* 945–949. doi:10.1042/bj3280945
4. Mark, B. L.; Vocadlo, D. J.; Knapp, S.; Triggs-Raine, B. L.; Withers, S. G.; James, M. N. G. *J. Biol. Chem.* **2001,** *276,* 10330–10337. doi:10.1074/jbc.M011067200
5. Williams, S. J.; Mark, B. L.; Vocadlo, D. J.; James, M. N. G.; Withers, S. G. *J. Biol. Chem.* **2002,** *277,* 40055–40065. doi:10.1074/jbc.M206481200
6. Sinnott, M. L. *Chem. Rev.* **1990,** *90,* 1171–1202. doi:10.1021/cr00105a006
7. Fairbanks, A. J. *Chem. Soc. Rev.* **2017,** *46,* 5128–5146. doi:10.1039/C6CS00897F
8. Abbott, D. W.; Macauley, M. S.; Vocadlo, D. J.; Boraston, A. B. *J. Biol. Chem.* **2009,** *284,* 11676–11689. doi:10.1074/jbc.M809663200
9. Banoub, J.; Boullanger, P.; Lafont, D. *Chem. Rev.* **1992,** *92,* 1167–1195. doi:10.1021/cr00014a002
10. Křenek, K.; Šimon, P.; Weignerová, L.; Fliedrová, B.; Kuzma, M.; Křen, V. *Beilstein J. Org. Chem.* **2012,** *8,* 428–432. doi:10.3762/bjoc.8.48
11. Freese, S. J.; Vann, W. F. *Carbohydr. Res.* **1996,** *281,* 313–319. doi:10.1016/0008-6215(95)00345-2
12. Kadokawa, J.-i.; Mito, M.; Takahashi, S.; Noguchi, M.; Shoda, S.-i. *Heterocycles* **2004,** *63,* 1531–1535. doi:10.3987/COM-04-10065
13. Zurabyan, S. E.; Antonenko, T. S.; Khorlin, A. Y. *Carbohydr. Res.* **1970,** *15,* 21–27. doi:10.1016/S0008-6215(00)80289-X
14. Rollin, P.; Sinaÿ, P. *J. Chem. Soc., Perkin Trans. 1* **1977,** 2513–2517. doi:10.1039/P19770002513
15. Kadokawa, J.-i.; Kasai, S.; Watanabe, Y.; Karasu, M.; Tagaya, H.; Chiba, K. *Macromolecules* **1997,** *30,* 8212–8217. doi:10.1021/ma971117j
16. Kadokawa, J.-i.; Sato, M.; Karasu, M.; Tagaya, H.; Chiba, K. *Angew. Chem., Int. Ed.* **1998,** *37,* 2373–2376. doi:10.1002/(SICI)1521-3773(19980918)37:17<2373::AID-ANIE2373>3.0.CO;2-B
17. Wittmann, V.; Lennartz, D. *Eur. J. Org. Chem.* **2002,** 1363–1367. doi:10.1002/1099-0690(200204)2002:8<1363::AID-EJOC1363>3.0.CO;2-#
18. Crasto, C. F.; Jones, G. B. *Tetrahedron Lett.* **2004,** *45,* 4891–4894. doi:10.1016/j.tetlet.2004.04.127
19. Blatter, G.; Beau, J.-M.; Jacquinet, J.-C. *Carbohydr. Res.* **1994,** *260,* 189–202. doi:10.1016/0008-6215(94)84038-5
20. Bartek, J.; Müller, K.; Kosma, P. *Carbohydr. Res.* **1998,** *308,* 259–273. doi:10.1016/S0008-6215(98)00082-2
21. Bélot, F.; Jacquinet, J.-C. *Carbohydr. Res.* **2000,** *325,* 93–106. doi:10.1016/S0008-6215(99)00322-5
22. Sherman, A. A.; Yudina, O. N.; Mironov, Y. V.; Sukhova, E. V.; Shashkov, A. S.; Menshov, V. M.; Nifantiev, N. E. *Carbohydr. Res.* **2001,** *336,* 13–46. doi:10.1016/S0008-6215(01)00213-0

23. Donohoe, T. J.; Logan, J. G.; Laffan, D. D. P. *Org. Lett.* **2003,** *5,* 4995–4998. doi:10.1021/ol0359620
24. Cobucci-Ponzano, B.; Strazzulli, A.; Rossi, M.; Moracci, M. *Adv. Synth. Catal.* **2011,** *353,* 2284–2300. doi:10.1002/adsc.201100461
25. Bojarová, P.; Křen, V. *Chimia* **2011,** *65,* 65–70. doi:10.2533/chimia.2011.65
26. Danby, P. M.; Withers, S. G. *ACS Chem. Biol.* **2016,** *11,* 1784–1794. doi:10.1021/acschembio.6b00340
27. Hayes, M. R.; Pietruszka, J. *Molecules* **2017,** *22,* No. 1434. doi:10.3390/molecules22091434
28. Kobayashi, S.; Kiyosada, T.; Shoda, S.-i. *Tetrahedron Lett.* **1997,** *38,* 2111–2112. doi:10.1016/S0040-4039(97)00319-5
29. Shoda, S.-i.; Kiyosada, T.; Mori, H.; Kobayashi, S. *Heterocycles* **2000,** *52,* 599–602. doi:10.3987/COM-99-S40
30. Sakamoto, J.; Watanabe, T.; Ariga, Y.; Kobayashi, S. *Chem. Lett.* **2001,** *30,* 1180–1181. doi:10.1246/cl.2001.1180
31. Ochiai, H.; Ohmae, M.; Kobayashi, S. *Carbohydr. Res.* **2004,** *339,* 2769–2788. doi:10.1016/j.carres.2004.08.016
32. Fujita, M.; Shoda, S.-i.; Haneda, K.; Inazu, T.; Takegawa, K.; Yamamoto, K. *Biochim. Biophys. Acta* **2001,** *1528,* 9–14. doi:10.1016/S0304-4165(01)00164-7
33. Huang, W.; Giddens, J.; Fan, S.-Q.; Toonstra, C.; Wang, L.-X. *J. Am. Chem. Soc.* **2012,** *134,* 12308–12318. doi:10.1021/ja3051266
34. Goodfellow, J. J.; Baruah, K.; Yamamoto, K.; Bonomelli, C.; Krishna, B.; Harvey, D. J.; Crispin, M.; Scanlan, C. N.; Davis, B. G. *J. Am. Chem. Soc.* **2012,** *134,* 8030–8033. doi:10.1021/ja301334b
35. For a review see: Fairbanks, A. J. *Pure Appl. Chem.* **2013,** *85,* 1847–1863. doi:10.1351/pac-con-12-09-10
36. Fairbanks, A. J. *C. R. Chim.* **2011,** *14,* 44–58. doi:10.1016/j.crci.2010.05.014
37. Matta, K. L.; Johnson, E. A.; Barlow, J. J. *Carbohydr. Res.* **1973,** *26,* 215–218. doi:10.1016/S0008-6215(00)85039-9
38. Srivastava, V. K. *Carbohydr. Res.* **1982,** *103,* 286–292. doi:10.1016/S0008-6215(00)80691-6
39. Nakabayashi, S.; Warren, C. D.; Jeanloz, R. W. *Carbohydr. Res.* **1986,** *150,* C7–C10. doi:10.1016/0008-6215(86)80028-3
40. Colon, M.; Staveski, M. M.; Davis, J. T. *Tetrahedron Lett.* **1991,** *32,* 4447–4450. doi:10.1016/0040-4039(91)80008-T
41. Cox, D. J.; Parsons, T. B.; Fairbanks, A. J. *Synlett* **2010,** 1315–1318. doi:10.1055/s-0029-1219919
42. Noguchi, M.; Tanaka, T.; Gyakushi, H.; Kobayashi, A.; Shoda, S.-i. *J. Org. Chem.* **2009,** *74,* 2210–2212. doi:10.1021/jo8024708
43. Noguchi, M.; Fujieda, T.; Huang, W. C.; Ishihara, M.; Kobayashi, A.; Shoda, S.-i. *Helv. Chim. Acta* **2012,** *95,* 1928–1936. doi:10.1002/hlca.201200414
44. Tomabechi, Y.; Krippner, G.; Rendle, P. M.; Squire, M. A.; Fairbanks, A. J. *Chem. – Eur. J.* **2013,** *19,* 15084–15088. doi:10.1002/chem.201303303
45. Priyanka, P.; Parsons, T. B.; Miller, A.; Platt, F. M.; Fairbanks, A. J. *Angew. Chem., Int. Ed.* **2016,** *55,* 5058–5061. doi:10.1002/anie.201600817
46. Yamaguchi, T.; Amin, M. N.; Toonstra, C.; Wang, L.-X. *J. Am. Chem. Soc.* **2016,** *138,* 12472–12485. doi:10.1021/jacs.6b05762
47. Isobe, T.; Ishikawa, T. *J. Org. Chem.* **1999,** *64,* 6984–6988. doi:10.1021/jo990210y
48. Alexander, S. R.; Lim, D.; Amso, Z.; Brimble, M. A.; Fairbanks, A. J. *Org. Biomol. Chem.* **2017,** *15,* 2152–2156. doi:10.1039/C7OB00112F
49. Köhling, S.; Exner, M. P.; Nojoumi, S.; Schiller, J.; Budisa, N.; Rademann, J. *Angew. Chem., Int. Ed.* **2016,** *55,* 15510–15514. doi:10.1002/anie.201607228
50. Lim, D.; Brimble, M. A.; Kowalczyk, R.; Watson, A. J. A.; Fairbanks, A. J. *Angew. Chem., Int. Ed.* **2014,** *53,* 11907–11911. doi:10.1002/anie.201406694
51. Ishiwata, A.; Ito, Y. Stereoselective Synthesis of β-*manno*-Glycosides. In *Glycoscience: Chemistry and Chemical Biology;* Fraser-Reid, B. O.; Tatsuta, K.; Thiem, J.; Coté, G. L.; Flitsch, S.; Ito, Y.; Kondo, H.; Nishimura, S.-i.; Yu, B., Eds.; Springer-Verlag: Berlin, 2008; pp 1279–1312. doi:10.1007/978-3-540-30429-6_30
52. Ennis, S. C.; Osborn, H. M. I. In *Carbohydrates;* Osborn, H. M. I., Ed.; Academic Press: Oxford, 2003; pp 239–276. doi:10.1016/B978-012312085-4/50010-6
53. Li, B.; Zeng, Y.; Hauser, S.; Song, H.; Wang, L.-X. *J. Am. Chem. Soc.* **2005,** *127,* 9692–9693. doi:10.1021/ja051715a
54. Rising, T. W. D. F.; Claridge, T. D. W.; Davies, N.; Gamblin, D. P.; Moir, J. W. B.; Fairbanks, A. J. *Carbohydr. Res.* **2006,** *341,* 1574–1596. doi:10.1016/j.carres.2006.03.007
55. Rising, T. W. D. F.; Claridge, T. D. W.; Moir, J. W. B.; Fairbanks, A. J. *ChemBioChem* **2006,** *7,* 1177–1180. doi:10.1002/cbic.200600183
56. Li, B.; Song, H.; Hauser, S.; Wang, L.-X. *Org. Lett.* **2006,** *8,* 3081–3084. doi:10.1021/ol061056m
57. Zeng, Y.; Wang, J.; Li, B.; Hauser, S.; Li, H.; Wang, L.-X. *Chem. – Eur. J.* **2006,** *12,* 3355–3364. doi:10.1002/chem.200501196
58. Rising, T. W. D. F.; Heidecke, C. D.; Moir, J. W. B.; Ling, Z.; Fairbanks, A. J. *Chem. – Eur. J.* **2008,** *14,* 6444–6464. doi:10.1002/chem.200800365
59. Watt, G. M.; Boons, G.-J. *Carbohydr. Res.* **2004,** *339,* 181–193. doi:10.1016/j.carres.2003.10.029
60. Fürstner, A.; Konetzki, I. *Tetrahedron Lett.* **1998,** *39,* 5721–5724. doi:10.1016/S0040-4039(98)01163-0
61. Fürstner, A.; Konetzki, I. *J. Org. Chem.* **1998,** *63,* 3072–3080. doi:10.1021/jo9800098
62. Parsons, T. B.; Moir, J. W. B.; Fairbanks, A. J. *Org. Biomol. Chem.* **2009,** *7,* 3128–3140. doi:10.1039/b907273j
63. Lergenmüller, M.; Nukada, T.; Kuramochi, K.; Dan, A.; Ogawa, T.; Ito, Y. *Eur. J. Org. Chem.* **1999,** 1367–1376. doi:10.1002/(SICI)1099-0690(199906)1999:6<1367::AID-EJOC1367>3.0.CO;2-N
64. Dan, A.; Lergenmüller, M.; Amano, M.; Nakahara, Y.; Ogawa, T.; Ito, Y. *Chem. – Eur. J.* **1998,** *4,* 2182–2190. doi:10.1002/(SICI)1521-3765(19981102)4:11<2182::AID-CHEM2182>3.0.CO;2-U
65. Ito, Y.; Ohnishi, Y.; Ogawa, T.; Nakahara, Y. *Synlett* **1998,** 1102–1104. doi:10.1055/s-1998-1894
66. Dan, A.; Ito, Y.; Ogawa, T. *J. Org. Chem.* **1995,** *60,* 4680–4681. doi:10.1021/jo00120a002
67. Ito, Y.; Ogawa, T. *Angew. Chem., Int. Ed. Engl.* **1994,** *33,* 1765–1767. doi:10.1002/anie.199417651
68. Chayajarus, K.; Chambers, D. J.; Chughtai, M. J.; Fairbanks, A. J. *Org. Lett.* **2004,** *6,* 3797–3800. doi:10.1021/ol048427o
69. Cumpstey, I.; Fairbanks, A. J.; Redgrave, A. J. *Org. Lett.* **2001,** *3,* 2371–2374. doi:10.1021/ol016175a
70. Seward, C. M. P.; Cumpstey, I.; Aloui, M.; Ennis, S. C.; Redgrave, A. J.; Fairbanks, A. J. *Chem. Commun.* **2000,** 1409–1410. doi:10.1039/b004522p
71. Ennis, S. C.; Fairbanks, A. J.; Slinn, C. A.; Tennant-Eyles, R. J.; Yeates, H. S. *Tetrahedron* **2001,** *57,* 4221–4230. doi:10.1016/S0040-4020(01)00308-8

72. Crich, D.; Sun, S. *J. Am. Chem. Soc.* **1997,** *119,* 11217–11223. doi:10.1021/ja971239r
73. Crich, D.; Sun, S. *J. Am. Chem. Soc.* **1998,** *120,* 435–436. doi:10.1021/ja9734814
74. Crich, D.; Sun, S. *Tetrahedron* **1998,** *54,* 8321–8348. doi:10.1016/S0040-4020(98)00426-8
75. Crich, D.; Dudkin, V. *J. Am. Chem. Soc.* **2001,** *123,* 6819–6825. doi:10.1021/ja010086b
76. Dudkin, V. Y.; Crich, D. *Tetrahedron Lett.* **2003,** *44,* 1787–1789. doi:10.1016/S0040-4039(03)00102-3
77. Ochiai, H.; Huang, W.; Wang, L.-X. *J. Am. Chem. Soc.* **2008,** *130,* 13790–13803. doi:10.1021/ja805044x
78. Amin, M. N.; Huang, W.; Mizanur, R. M.; Wang, L.-X. *J. Am. Chem. Soc.* **2011,** *133,* 14404–14417. doi:10.1021/ja204831z
79. Zou, G.; Ochiai, H.; Huang, W.; Yang, Q.; Li, C.; Wang, L.-X. *J. Am. Chem. Soc.* **2011,** *133,* 18975–18991. doi:10.1021/ja208390n
80. Priyanka, P.; Fairbanks, A. J. *Carbohydr. Res.* **2016,** *426,* 40–45. doi:10.1016/j.carres.2016.03.015
81. Ravi Kumar, H. V.; Naruchi, K.; Miyoshi, R.; Hinou, H.; Nishimura, S.-I. *Org. Lett.* **2013,** *15,* 6278–6281. doi:10.1021/ol403140h
82. Seko, A.; Koketsu, M.; Nishizono, M.; Enoki, Y.; Ibrahim, H. R.; Juneja, L. R.; Kim, M.; Yamamoto, T. *Biochim. Biophys. Acta* **1997,** *1335,* 23–32. doi:10.1016/S0304-4165(96)00118-3
83. Zou, Y.; Wu, Z.; Chen, L.; Liu, X.; Gu, G.; Xue, M.; Wang, P. G.; Chen, M. *J. Carbohydr. Chem.* **2012,** *31,* 436–446. doi:10.1080/07328303.2012.666689
84. Sun, B.; Bao, W.; Tian, X.; Li, M.; Liu, H.; Dong, J.; Huang, W. *Carbohydr. Res.* **2014,** *396,* 62–69. doi:10.1016/j.carres.2014.07.013
85. Liu, L.; Prudden, A. R.; Bosman, G. P.; Boons, G.-J. *Carbohydr. Res.* **2017,** *452,* 122–128. doi:10.1016/j.carres.2017.10.001
86. Fujita, K.; Kobayashi, K.; Iwamatsu, A.; Takeuchi, M.; Kumagai, H.; Yamamoto, K. *Arch. Biochem. Biophys.* **2004,** *432,* 41–49. doi:10.1016/j.abb.2004.09.013
87. Huang, W.; Yang, Q.; Umekawa, M.; Yamamoto, K.; Wang, L.-X. *ChemBioChem* **2010,** *11,* 1350–1355. doi:10.1002/cbic.201000242
88. Umekawa, M.; Higashiyama, T.; Koga, Y.; Tanaka, T.; Noguchi, M.; Kobayashi, A.; Shoda, S.-i.; Huang, W.; Wang, L.-X.; Ashida, H.; Yamamoto, K. *Biochim. Biophys. Acta* **2010,** *1800,* 1203–1209. doi:10.1016/j.bbagen.2010.07.003
89. Kowalczyk, R.; Brimble, M. A.; Tomabechi, Y.; Fairbanks, A. J.; Fletcher, M.; Hay, D. L. *Org. Biomol. Chem.* **2014,** *12,* 8142–8151. doi:10.1039/C4OB01208A
90. Tomabechi, Y.; Squire, M. A.; Fairbanks, A. J. *Org. Biomol. Chem.* **2014,** *12,* 942–955. doi:10.1039/C3OB42104J
91. Higuchi, Y.; Eshima, Y.; Huang, Y.; Kinoshita, T.; Sumiyoshi, W.; Nakakita, S.-i.; Takegawa, K. *Biotechnol. Lett.* **2017,** *39,* 157–162. doi:10.1007/s10529-016-2230-0
92. Huang, W.; Li, C.; Li, B.; Umekawa, M.; Yamamoto, K.; Zhang, X.; Wang, L.-X. *J. Am. Chem. Soc.* **2009,** *131,* 2214–2223. doi:10.1021/ja8074677
93. Smith, E. L.; Giddens, J. P.; Iavarone, A. T.; Godula, K.; Wang, L.-X.; Bertozzi, C. R. *Bioconjugate Chem.* **2014,** *25,* 788–795. doi:10.1021/bc500061s
94. Maki, Y.; Okamoto, R.; Izumi, M.; Murase, T.; Kajihara, Y. *J. Am. Chem. Soc.* **2016,** *138,* 3461–3468. doi:10.1021/jacs.5b13098
95. Kajihara, Y.; Suzuki, Y.; Yamamoto, N.; Sasaki, K.; Sakakibara, T.; Juneja, L. R. *Chem. – Eur. J.* **2004,** *10,* 971–985. doi:10.1002/chem.200305115
96. Lis, H.; Sharon, N.; Katchalski, E. *Biochim. Biophys. Acta* **1964,** *83,* 376–378. doi:10.1016/0926-6526(64)90024-2
97. Lis, H.; Sharon, N.; Katchalski, E. *J. Biol. Chem.* **1966,** *241,* 684–689.
98. Dorland, L.; van Halbeek, H.; Vleigenthart, J. F.; Lis, H.; Sharon, N. *J. Biol. Chem.* **1981,** *256,* 7708–7711.
99. Lis, H.; Sharon, N. *J. Biol. Chem.* **1978,** *253,* 3468–3476.
100. Fan, J.-Q.; Huynh, L. H.; Reinhold, B. B.; Reinhold, V. N.; Takegawa, K.; Iwahara, S.; Kondo, A.; Kato, I.; Lee, Y. C. *Glycoconjugate J.* **1996,** *13,* 643–652. doi:10.1007/BF00731453
101. Evers, D. L.; Hung, R. L.; Thomas, V. H.; Rice, K. G. *Anal. Biochem.* **1998,** *265,* 313–316. doi:10.1006/abio.1998.2895
102. Takegawa, K.; Yamabe, K.; Fujita, K.; Tabuchi, M.; Mita, M.; Izu, H.; Watanabe, A.; Asada, Y.; Sano, M.; Kondo, A.; Kato, I.; Iwahara, S. *Arch. Biochem. Biophys.* **1997,** *338,* 22–28. doi:10.1006/abbi.1996.9803
103. Umekawa, M.; Huang, W.; Li, B.; Fujita, K.; Ashida, H.; Wang, L.-X.; Yamamoto, K. *J. Biol. Chem.* **2008,** *283,* 4469–4479. doi:10.1074/jbc.M707137200
104. McIntosh, J. D.; Brimble, M. A.; Brooks, A. E. S.; Dunbar, P. R.; Kowalczyk, R.; Tomabechi, Y.; Fairbanks, A. *J. Chem. Sci.* **2015,** *6,* 4636–4642. doi:10.1039/C5SC00952A
105. Giddens, J. P.; Lomino, J. V.; Amin, M. N.; Wang, L.-X. *J. Biol. Chem.* **2016,** *291,* 9356–9370. doi:10.1074/jbc.M116.721597
106. Spiro, R. G. *J. Biol. Chem.* **1960,** *235,* 2860–2869.
107. Waddling, C. A.; Plummer, T. H.; Tarentino, A. L.; Van Roey, P. *Biochemistry* **2000,** *39,* 7878–7885. doi:10.1021/bi0001731
108. Huang, C.-C.; Mayer, H. E., Jr.; Montgomery, R. *Carbohydr. Res.* **1970,** *13,* 127–137. doi:10.1016/S0008-6215(00)84902-2
109. Wang, L.-X.; Ni, J.; Singh, S.; Li, H. *Chem. Biol.* **2004,** *11,* 127–134. doi:10.1016/j.chembiol.2003.12.020
110. Li, T.; Tong, X.; Yang, Q.; Giddens, J. P.; Wang, L.-X. *J. Biol. Chem.* **2016,** *291,* 16508–16518. doi:10.1074/jbc.M116.738765

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Carbohydrate inhibitors of cholera toxin

Vajinder Kumar*[1,2] and W. Bruce Turnbull*[2]

Review

Open Access

Address:
[1]Department of Chemistry, Akal University, Talwandi Sabo, Punjab, India and [2]School of Chemistry and Astbury Centre for Structural Molecular Biology, University of Leeds, LS2 9JT, UK

Email:
Vajinder Kumar* - vajinder_chm@auts.ac.in; W. Bruce Turnbull* - w.b.turnbull@leeds.ac.uk

* Corresponding author

Keywords:
carbohydrate; cholera; multivalency; toxin

Beilstein J. Org. Chem. **2018,** *14,* 484–498.
doi:10.3762/bjoc.14.34

Received: 18 August 2017
Accepted: 08 February 2018
Published: 21 February 2018

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

Cholera is a diarrheal disease caused by a protein toxin released by *Vibrio cholera* in the host's intestine. The toxin enters intestinal epithelial cells after binding to specific carbohydrates on the cell surface. Over recent years, considerable effort has been invested in developing inhibitors of toxin adhesion that mimic the carbohydrate ligand, with particular emphasis on exploiting the multivalency of the toxin to enhance activity. In this review we introduce the structural features of the toxin that have guided the design of diverse inhibitors and summarise recent developments in the field.

Introduction

Cholera, meaning a flow of bile, is caused by an acute enteric infection of the Gram-negative facultative anaerobe *Vibrio cholerae*. Not only does this disease have a disastrous effect on health, it also impacts on the socioeconomic status of societies where it is endemic. The *V. cholerae* bacterium was identified by Robert Koch in 1883, and ever since then, this scourge has grown continuously with catastrophic effects on millions of people [1]. Although appropriate water, hygiene and sanitation interventions can reduce incidence of bacterial infection, the WHO predicts that there will still continue to be millions of deaths due to diarrhoea in the developing nations of the world. While cholera is rare and seldom life threatening in developed countries, it can still pose a risk to those at the extremes of age and the immunosuppressed. However, Hispaniola Island and western African countries (Ghana, Guinea, Guinea-Bissau, Niger and Sierra Leone) are completely under the control of this epidemic. According to annual statistics of 2016 in the *Weekly Epidemiological Record* (WER) by the WHO, 172454 cases are reported in 42 endemic countries including 1304 deaths. Among 42 countries, Afghanistan, the Democratic Republic of the Congo (DRC), Haiti, Kenya, and the United Republic of Tanzania were majorly affected [2]. Recent data for the year 2017 from the GIDEON internet site (that continuously scans Medline, WHO, CDC and other peer reviewed journals), highlights the recent cholera outbreak principally affecting Somalia, DRC and Tanzania [3]. The total number of cases reported in these countries was almost 65,000 leading to 1500 deaths so far. In the Americas, the Haiti region has been fighting this

epidemic since October 2010. As of June 2017, the outbreak was still ongoing and a total of over 800,000 cases, including 10,000 deaths, had been registered [3]. This infection also prevails in the Dominican Republic and Cuba [2]. Furthermore, deaths due to cholera in Asian countries constitute 3% of the world's total [2]. However, this may be underestimated as limitations in surveillance systems in large parts of Asia, lead to millions of cholera cases not being recorded. After broad analysis, Ali et al. estimated that 2.9 million cases and 95,000 deaths happen every year worldwide [4]. Thus cholera continues to be a serious concern in many parts of the globe.

The agent responsible for causing diarrhea is an AB_5 toxin released by the bacteria. Thus, an understanding of this toxin becomes essential in finding/developing molecules that could prevent cell entry of the toxin and inhibit its activity. AB_5 toxins are an important class of bacterial toxins. They consist of a single A-subunit and a pentamer of B-subunits [5]. The catalytic activity of the toxins is due to the A-subunit, while the B-subunit enables binding of the complex to the cell surface and its delivery into the target cells, hence the complete AB_5 holotoxin is required for their toxic effects. Because of the difference in the sequence homology and catalytic activity, the classes of AB_5 toxins are subdivided into three families (Figure 3): the cholera toxin (CT) family, the shiga toxin (ST) family and the pertussis toxin (PT) family [6]. The CT family contains CT, and heat-labile toxins LT-I and LT-II [7,8]. The ST family contains the shiga toxins (SHT) themselves and the related verotoxins (also known as shiga-like toxins: SLT-I, SLT-II) [9,10] and SHT toxin comes from *Shigella dysenteriae* and verotoxin comes from enteropathogenic *E. coli* strains such as O157-H7. SHT and SLT-I are almost identical, with very little difference in the A-subunit. But the SLT-II shows more deviation in its gene sequence from the SHT and SLT-I toxins [9]. Sequence homology in the CT family is high between CTB and LTI-B (80% identical), but much lower between these proteins and the LTIIa and LTIIB toxins. PT is quite unusual in that all five of its B-subunits are different, but overall, an AB_5 architecture is still preserved [11]. A detailed knowledge of the 3D structure of these toxins is informative for the design of effective inhibitors.

Review

Structure and function of cholera toxin

Many crystallographic studies of the AB_5 toxins have been undertaken over the past 20 years [8-14]. Here, we focus solely on those describing the structure of the cholera toxin.

A-Subunit

The A-subunit of CT is the catalytic site of the AB_5 toxin, and forms a complex with the B-pentamer [15]. It is initially expressed as a single polypeptide chain which is cleaved by a protease to give two subunits, A1 and A2, remain held together by extensive non-covalent forces and a single interchain disulfide bond [16]. The A2-subunit acts as a linker between the toxic A1-subunit and CTB which is the delivery vehicle that can transport the complex into cells and direct the toxin to the endoplasmic reticulum, from where it can escape into the cytosol. The A1 chain has ADP-ribosyltransferase activity that allows the toxin to covalently modify the α-subunit of the stimulatory G protein $G_{s\alpha}$ so that it remains in its active GTP-bound state. The consequence of this change is to produce high levels of cAMP which activates protein kinase A to phosphorylate the cystic fibrosis transmembrane conductance regulator which is a chloride ion channel [15]. Transport of chloride ions to the intestine is accompanied by excessive amounts of water entering the gut and the diarrhea that is symptomatic of cholera.

The A1-subunit consists of three domains namely $A1_1$, $A1_2$ and $A1_3$ (Figure 1). While the $A1_1$ domain is responsible for catalysis, the $A1_2$ and $A1_3$ domains have been implicated in allowing the A1 subunit to escape from the endoplasmic reticulum (ER) into the cytosol. Following arrival in the ER, protein disulfide bond isomerase can reduce the disulfide bond between $A1_3$ and A2, releasing the A1 protein and causing the $A1_2$ and $A1_3$ domains to unfold [17]. The protein is then recognised by the cell as a misfolded protein and is exported into the cytolsol for degradation. However, once in the cytosol, it binds to another protein Arf6, which stabilizes the $A1_2$/$A1_3$ domains and activates the A1 enzyme.

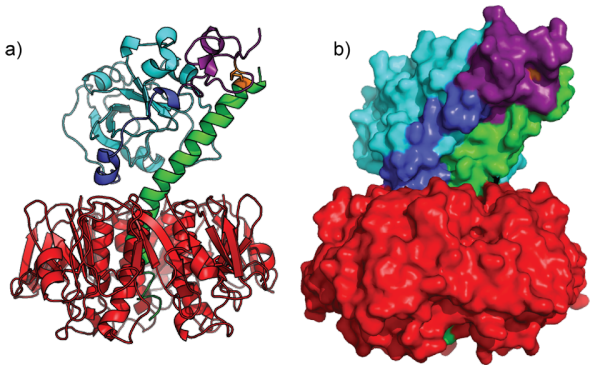


Figure 1: a) Ribbon and b) surface depictions of the cholera toxin: $A1_1$ domain in light blue; $A1_2$ domain in dark blue; $A1_3$ domain in purple; cystine disulfide in orange; A2 peptide in green and B-subunit in red. Figure prepared using the PyMOL programme from Protein Data Bank file 1XTC.pdb.

B-Subunit

The B-subunit (CTB) is a homopentamer [18,19], and crystallographic data on B subunits of the CT family showed very little deviation (less than 0.5 r.m.s.) from exact rotational symmetry.

Five long α-helices surround the central cylindrical pore through which the A2-subunit is threaded. Each subunit of a B-pentamer has a single binding site for the GM_1 oligosaccharide on the face of the pentamer distal to the A1-subunit [12,14]. GM_1 is a branched pentasaccharide [Galβ1-3GalNAcβ1-4(NeuAcα2-3)Galβ1-Glcβ1-1-ceramide] bearing a ceramide moiety at the anomeric center of the Glc moiety (Figure 2). The terminal galactose residue of GM_1 is buried most deeply inside the cavity of CTB [12,14], while the sialic acid branch sits in a wider shallow pocket. Both of these terminal sugar residues show hydrogen bonding interactions with the protein and associated water molecules. The GM_1 oligosaccharide (GM_1os) binds very tightly to CTB with a dissociation constant (K_d) of around 40 nM (measured by isothermal titration calorimetry, ITC), while simple galactosides have millimolar K_ds and little interaction can be detected for simple sialosides [20]. The distance separating the binding sites is similar for all members of the AB_5 toxin family and is believed to be instrumental in clustering the glycolipid ligands in such a way that membrane curvature is induced upon binding [21].

More recently, a second binding site has been discovered on the edge of the B-subunit sitting closer to the A-subunit face (Figure 2) [13,22-25], This secondary binding site recognises fucosylated structures including blood group oligosaccharides of the Lewis-y family. Individually, the interactions are much weaker than the CTB-GM_1os interaction (K_d ca. 1 mM measured by ITC), but even these weak binding interactions can still be functionally useful once the effect of multivalent binding enhancement has been taken into consideration. Indeed, ITC experiments have also shown the highest affinity site on the SLT-1 B-subunit has a K_d of only 1 mM [26], yet the toxin achieves sub-nanomolar affinity at a cell membrane. The purpose of the CTB blood group oligosaccharide binding site remains a topic for debate, but it may be responsible for the reported blood group dependence of the severity of cholera [13,24,27], or it could provide an independent route for cell entry through interactions with cell surface glycoproteins [28].

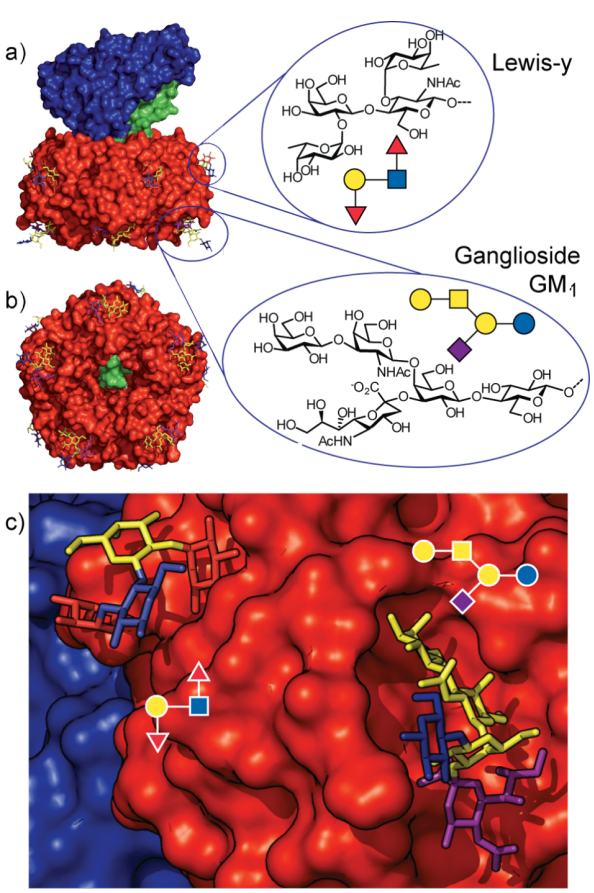


Figure 2: a) Structure of the cholera toxin showing the location of its carbohydrate binding sites and the structures of the Lewis-y and Ganglioside GM_1 ligands; A-subunit (blue), B-subunit (red) and the A2 peptide linker (green). b) Bottom face of the toxin showing the symmetry of the B-subunit and the A2 peptide linker emerging through the central channel. c) Close-up view of the two sugar binding sites. Figure prepared using the PyMOL programme from Protein Data Bank files 1XTC.pdb, 3CHB.pdb and 3EFX.pdb.

Structure-based design of inhibitors for cholera toxin

The availability of crystal structures for cholera and *E. coli* heat-labile toxins has driven opportunities for the design of potent inhibitors for these toxins. While some interest has been shown in the possibility of inhibition of cholera toxin assembly and inhibition of the enzymatic activity, most effort has been invested in seeking inhibitors of the adhesion process [29].

Designing the inhibitors for the receptor-binding process is a very compelling strategy, because the inhibitors would fight the toxin in the intestinal tract of the human host. Therefore ligands need not to cross any barrier and there is no constraint on ligand size. In the past years, several strategies have been drawn for the receptor binding to AB_5 toxins; while some target on the individual binding sites, others are intended at designing multivalent ligands against the entire toxin B pentamer [6,30,31].

Monovalent receptor-binding inhibitors

Bernardi and co-workers designed carbohydrate derivatives that mimic the natural CT receptor, ganglioside GM_1 [32]. They replaced the central 3,4-disubstituted Gal unit of GM_1 with dicarboxy cyclohexanediol (DCCHD, Figure 3). DCCHD exhibits the same absolute and relative configuration of the natural galactose residue. Taking this into account, a pseudo-tetrasaccharide **1** was made in which the recognition units, the terminal galactose and Neu5Ac, were attached onto the DCCHD scaffold. Inhibition assays of the oligosaccharide mimetic with CT and LT showed similar potency as that of natural ligands [32].

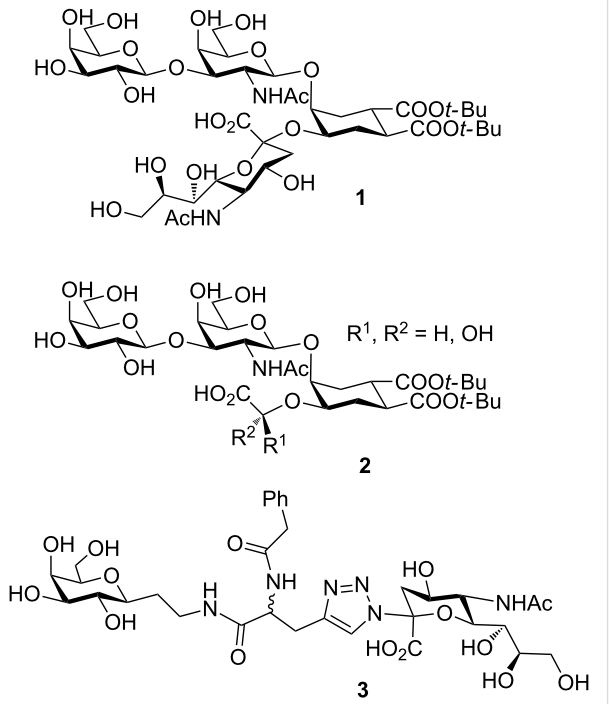


Figure 3: Bernardi and co-workers' designed oligosaccharide mimetics of GM_1.

But, the alpha-sialylation was the bottleneck step in the synthesis, so they designed second generation inhibitors by changing the synthetically challenging α-Neu5Ac with alpha-hydroxy acids **2** [33,34]. Using a combinatorial approach, a library of non-hydrolyzable, non O-glycosidic third generation inhibitors were synthesised using appropriate linkers. The CTB affinity of these inhibitors was measured using weak affinity chromatography and some molecules displayed enhancement of affinity over the individual epitome 'Galactose' [35]. One such compound **3** has found to co-crystallise with CTB in a way that the galactose and sialic acid groups bind to adjacent CTB pentamers in the crystal lattice, opening a possible route for the structure-based design of inhibitors that aggregate the toxin [36].

Hol, Verlinde and co-workers designed and synthesised a library of compounds utilizing a fragment of the toxin's natural receptor. Both CTB and LTB have specific affinity for the terminal galactose part of GM_1 [37-39]. They screened a number of galactose derivatives with substitution at O1 and C2 and found that the most potent molecule in this library was *m*-nitrophenyl α-D-galactoside (**4**) which was 100 times better than galactose for binding to CTB [38,39]. In another report, Mitchell et al. designed and synthesised twenty 3,5-substituted phenylgalactosides, e.g., **5** and when these compounds were tested on CT it was found that they have a six-fold higher affinity than *m*-nitrophenyl α-D-galactopyranoside (Figure 4) [40].

Vrasidas et al. synthesised a simple lactose-2-aminothiazoline conjugate as a CT antagonist. Its affinity for CTB was determined by monitoring the change in fluorescence of tryptophan-88, located in the GM_1 binding site, upon titration of the protein with the inhibitor. Compound **6** showed excellent binding with a K_d value of 23 μM [41]. Robina and co-workers synthesised non-hydrolyzable S-galactosides and non-carbohydrate

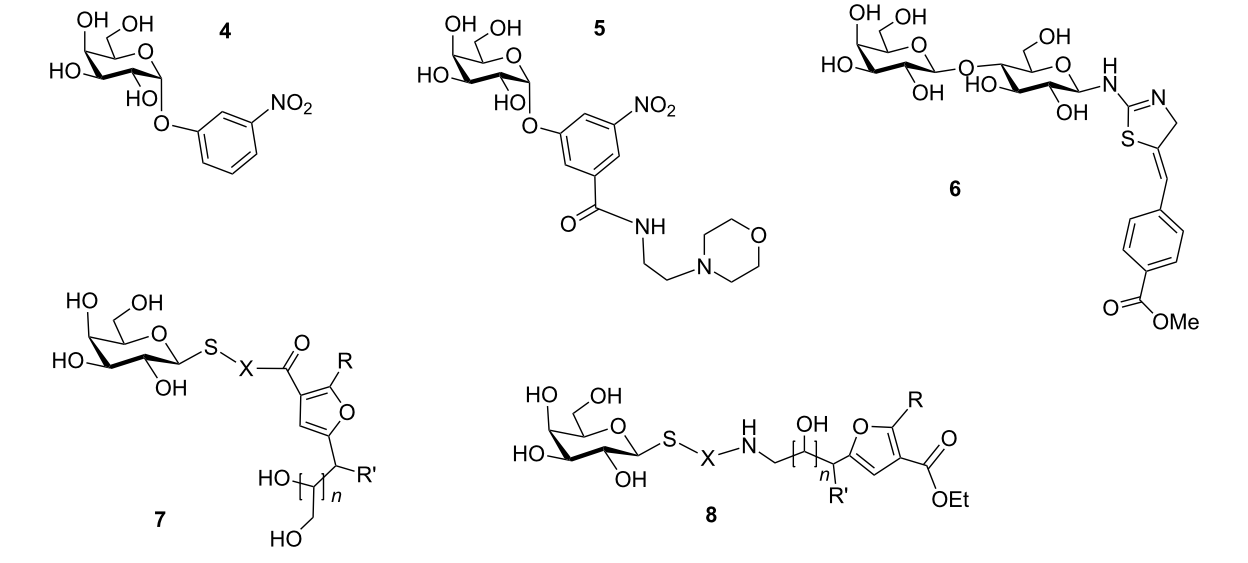


Figure 4: Structure of monomeric ligands. X = amino acid residues, aminoalkyl, 1,2,3 triazoles; *n* = 1, 2; R = H, Me, R' = OH, NHAc.

ligands based on polyhydroxyalkylfuronate moieties and measured their affinities by weak affinity chromatography (WAC) and also studied their interaction by saturation transfer difference NMR experiments [42]. Although, these compounds, **7** and **8**, did not display good inhibition, the non-glycosylated ligands offered new avenues for better CT ligand designs.

Multivalent receptor-binding inhibitors

The five-fold symmetry of AB_5 toxins provides a strong encouragement to think about multivalent inhibitor design from (even weakly binding) monovalent inhibitors [30,31]. Multivalent ligands have been long applied to a wide range of protein targets [43-45]. By having an inhibitor that may bind simultaneously with multiple binding sites, the dissociation rate of the complex is effectively reduced. Even if any individual ligand group dissociates from the protein, then the others will continue to make contact between the protein and the inhibitor, thus maintaining a high effective concentration of the dissociated ligand group in the vicinity of the binding site and increasing the probability of rebinding occurring. The gain in inhibitory potency for the multivalent ligands can be in many orders of magnitude. Here we have divided multivalent ligands and inhibitors of cholera toxin into three classes: sub-pentavalent inhibitors; pentavalent inhibitors; and inhibitors containing more than five ligands.

Sub-pentavalent inhibitors

Hol and Fan [46] designed and synthesised both spanning and non-spanning bivalent inhibitors. “Spanning” means the ligand has sufficient length of the linker to reach the two binding moieties of CT, whereas “non-spanning” means there is insufficient linker length for intra-pentamer chelation, but the second galactosyl moiety could bind to another CT molecule. They found that non-spanning bivalent inhibitors **9**–**12** as shown in Figure 5, show more binding affinity than the monovalent ones, which could also be derived from a statistical effect of a higher rebinding rate.

Bernardi, Casnati and co-workers prepared a bivalent ligand **13** for CT by attaching two copies of GM_1 mimic compound **3** to a calixarene (Figure 6) [47]. By measuring the affinity for CT by fluorescence titration, they found that the enhancement in affinity was 3800-fold as compared to the GM_1 mimic, which is consistent with a chelating mechanism.

Hughes and co-workers synthesised and evaluated bivalent 1,2,3 triazole-linked galactopyranosides **14** and **15** as shown in Figure 7 [48]. They used a piperazine core as central divalent core on to which the galactose units were attached via flexible linkers. They found that these compounds exhibit binding affinity one order higher than *m*-nitrophenyl galactopyranoside (**4**) [48]. In another recent report, low molecular weight poly(*N*-acryloylmorpholine) was used to link galactose residues to form a bivalent inhibitor, but the biological assay demonstrated only moderate inhibitory activity [49]. Liu et al. synthesised bivalent ligands **16** and **17**, for evaluation through biophysical techniques (Figure 7) [50]. They found that the enhancement in affinity and potency was due to non-specific interactions between the linker portion, nitrophenyl group and CT. The interactions increase as linker length increase. Hence, they concluded that the length, size and chemical nature of the ligand has a major effect on binding with the protein toxin.

While the ganglioside GM_1 head group is the highest affinity natural ligand for CTB, galactose and lactose (Figure 8) head groups have also been used for synthesising bivalent and

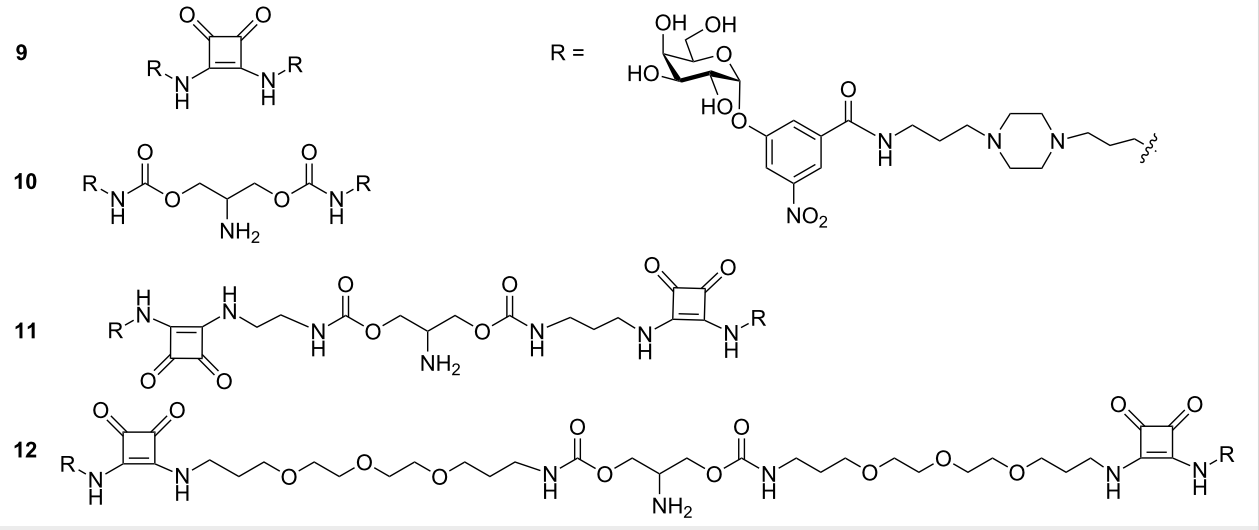
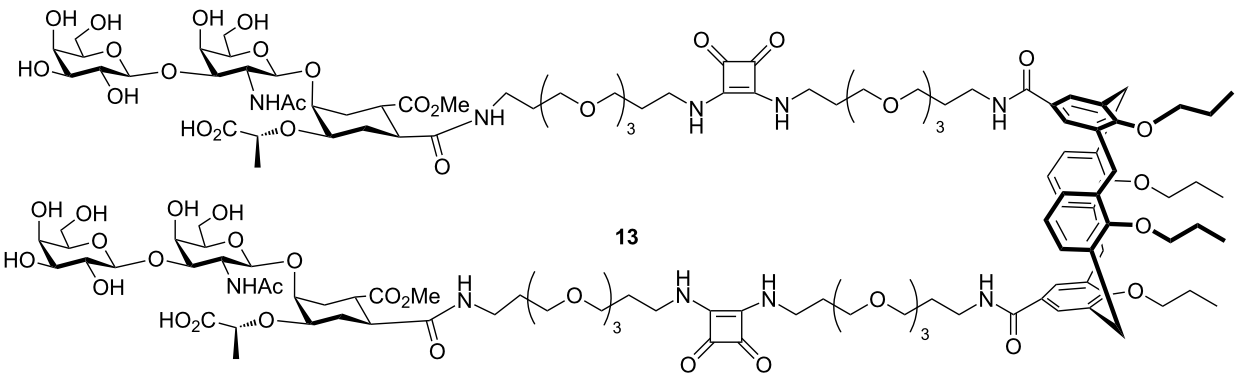
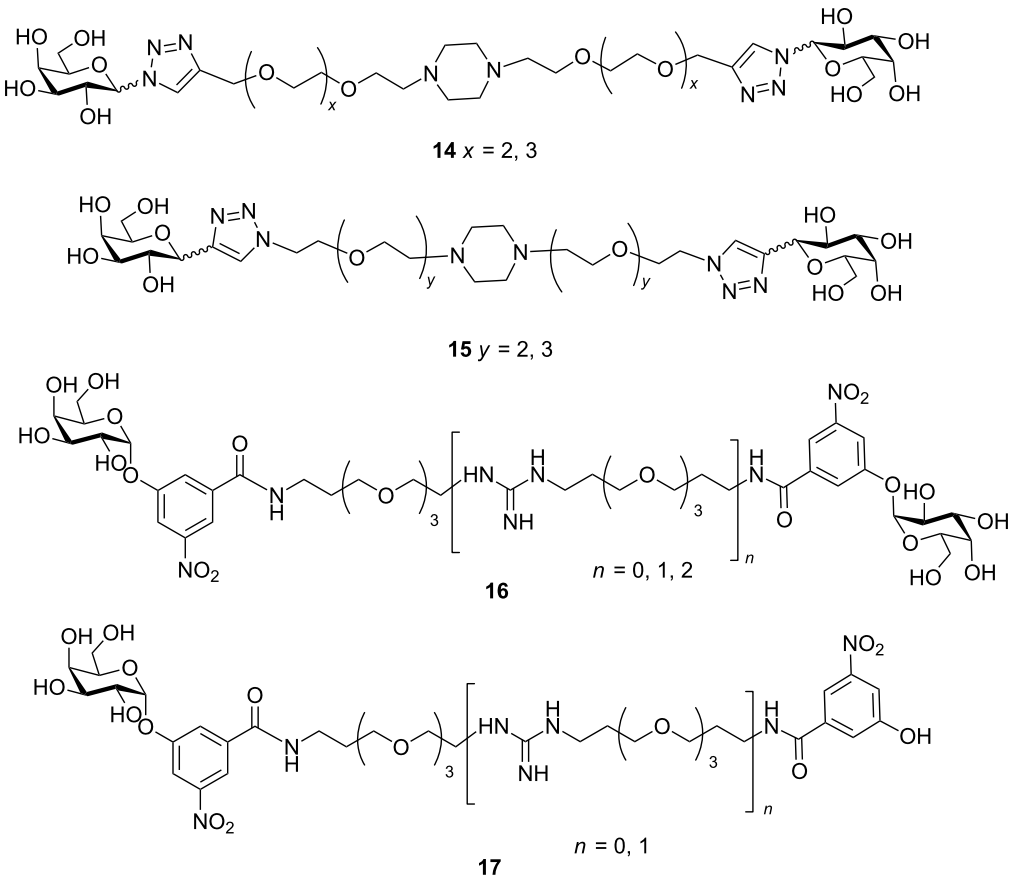


Figure 5: Bivalent inhibitor designed and synthesised by Pickens et al.



13

Figure 6: Bivalent inhibitor designed and synthesized by Arosio et al.



14 x = 2, 3

15 y = 2, 3

16 n = 0, 1, 2

17 n = 0, 1

Figure 7: Bivalent inhibitors designed and synthesised by Leaver and Liu.

tetravalent multivalent inhibitors and showed substantial gains in binding affinity in comparison to the corresponding monovalent ligands. Pieters and co-workers attached a lactose-derived monomeric ligand to the dendrimer **18**, and found that there was an affinity and potency gain from divalent and tetravalent molecules [51]. Even the galactose containing dendrimers **20** bind as strongly as that of GM_1 [52]. As an improved design of ligand for CT, the GM_1 mimic synthesised by Bernardi and co-workers was attached to the dendrimer synthesised by the Pieters group and hence compounds **19** and **21** were obtained [53]. The divalent compound **19a** and tetravalent compound **19b** exhibited IC_{50} values of 13 and 0.5 μM, respectively. In another report,

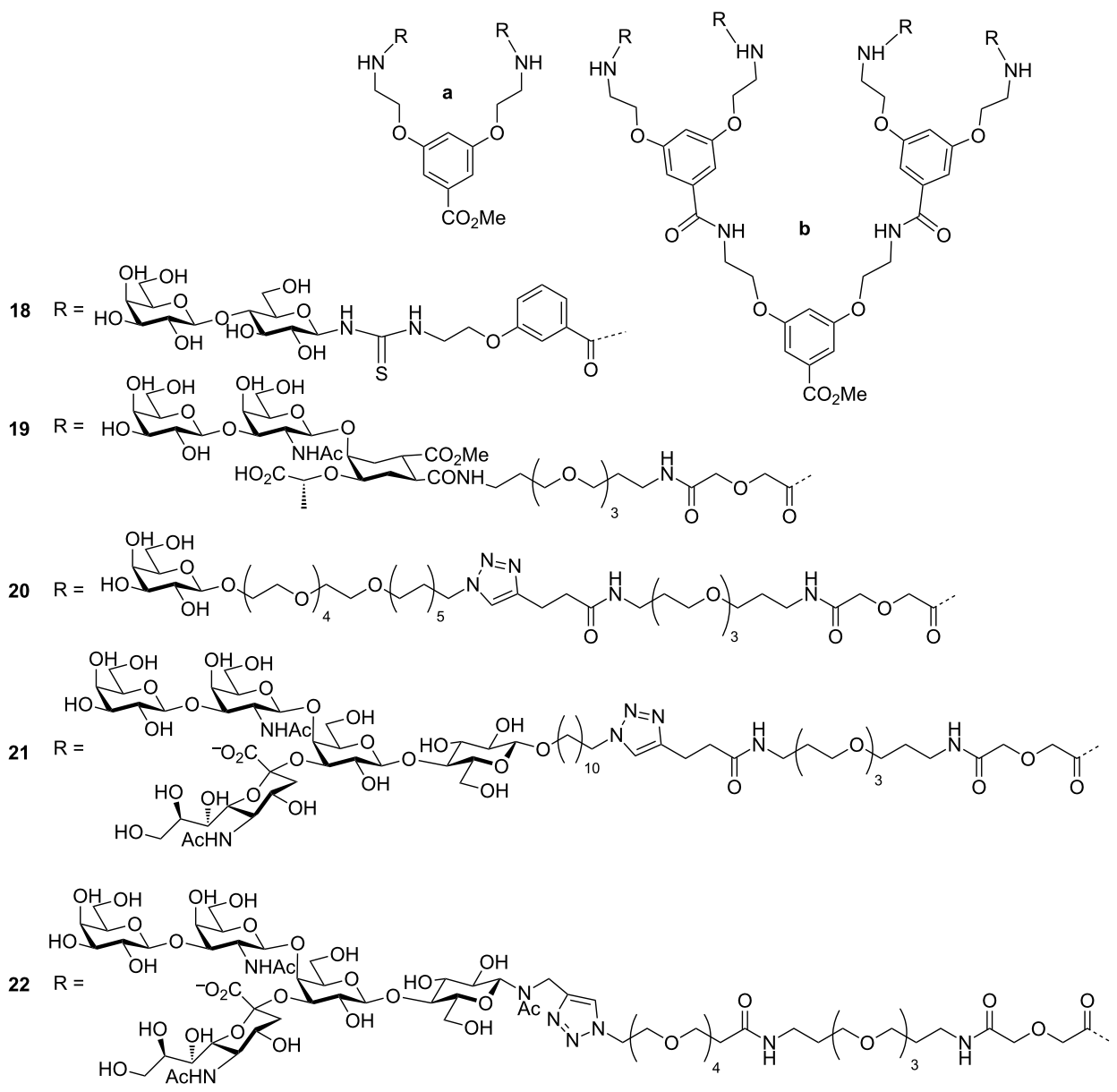


Figure 8: Bivalent and tetravalent inhibitor designed and synthesised by Pieters, and Bernardi et al.

they reported that the divalent compound **21a** and tetravalent compound **21b** displayed 9,500 and 83,000-fold enhanced potency, respectively, than monovalent GM_1 [54,55].

Fu et al. synthesised a tetravalent ligand containing highly hydrophilic spacer arms **22b**, and found that this ligand demonstrated almost the same potency with an IC_{50} value of 160 pM as that of **21b** (IC_{50} = 190 pM) [56].

To reduce the energy loss in the form of entropic penalty to be paid on binding, Kumar et al. synthesised noncyclic and cyclic neoglycopeptides and glycoamides for cholera toxin, e.g., **23**–**26** (Figure 9) [57]. They prepared divalent, trivalent, tetravalent, cyclic divalent, cyclic trivalent, cyclic tetravalent and cyclic pentavalent inhibitors with large cyclic core structures.

Pentavalent inhibitors

The pentameric structure of CTB has proved to be an enticing invitation to many scientists to develop multivalent inhibitors that are also pentavalent. Fan, Hol and co-workers were first to design and synthesise pentavalent inhibitors **27**–**29**, for the LTB/CTB [58] (although Bundle and co-workers were also working on analogous designs for shiga-like toxin [59]). They

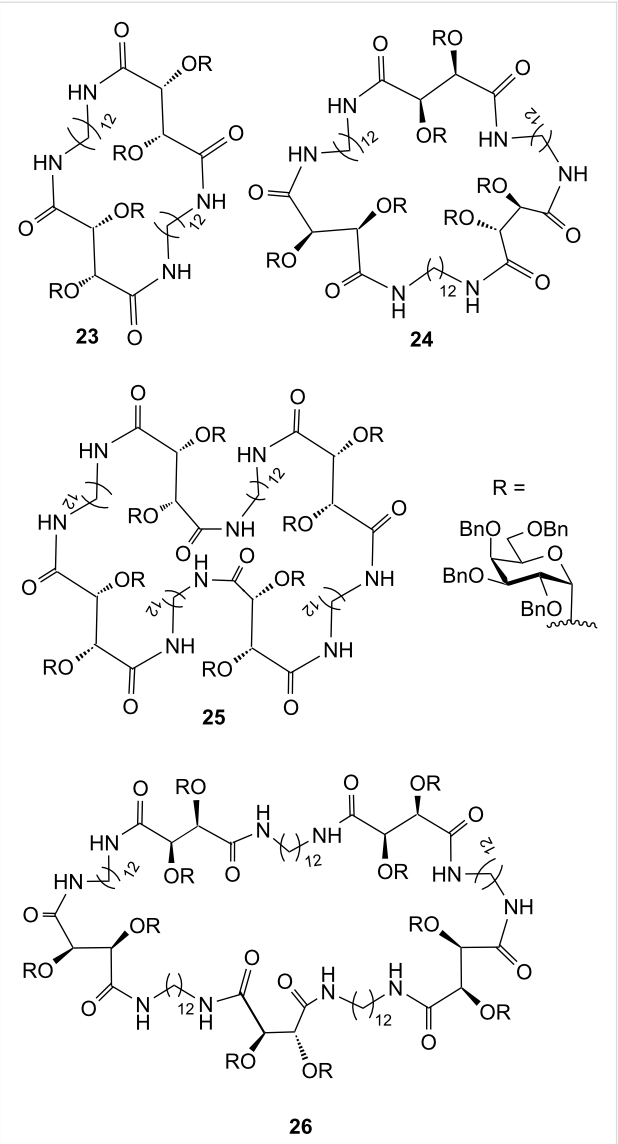


Figure 9: Cyclic inhibitors synthesised by Kumar et al. for CT.

synthesised the inhibitors on a pentacyclene core on which galactose and *m*-nitrophenyl-α-D-galactopyranoside were attached by long flexible linkers (Figure 10) [60,61]. They found million-fold increases in activity in comparison to the corresponding monovalent inhibitors with IC_{50} values of 40 nM.

Zhang et al. synthesised large cyclic decapeptides (up to 50 atoms in the ring) in a "core-linker-finger" modular setup (Figure 11) [62]. These compounds **30** showed good inhibitory results with IC_{50} values 100,000-fold more potent than monovalent galactose. This strategy facilitated a methodical study to measure the effect of linker length on the affinity of the pentavalent ligands towards the target toxin. Large affinity-gains were achieved for pentavalent ligands with short linkers on these large cyclic cores, indicating that the central cyclic peptide core probably has an expanded ring conformation.

Garcia-Hartjes et al. synthesised and evaluated the GM_1os linked calix[5]arene molecule **31** as shown in Figure 12, and found that compound **31** displayed 100,000 times more potency as compared to GM1os derivatives having an IC_{50} value of 450 pM [63].

In another report Siegel and co-workers showed that corannulene-based pentavalent glycocluster **32** (Figure 13) bearing GM_1os moieties possessed affinity for CT in low nanomolar range [64]. The IC_{50} value obtained was in the range of 5–25 nM.

Fu et al. also synthesised and evaluated a pentavalent inhibitor **33** (Figure 14), analogous to their tetravalent compound **22b** to investigate the difference between matching or mismatching the valency with that of the target CTB protein [56]. Previous biophysical studies had suggested that a mismatch in the valencies of ligand and receptor favoured an aggregation mechanism for inhibition [55] whereas matching the valency has previously been assumed to lead to the formation of 1:1 complexes [59-64]. They found that the potency exhibited by compound **33** in the usual enzyme-linked lectin assay (IC_{50} = 260 ± 20 pM) was slightly lower that for the tetravalent compound **22b** (IC_{50} = 160 ± 40 pM) [56]. Inhibition results described in this review are essentially all derived from very similar types of enzyme-linked lectin assays (ELLA) in which the inhibitors are used to prevent CTB-linked horseradish peroxidase from binding to microtitre plates coated with the ganglioside GM_1 ligand. However, it is important to note that IC_{50} values are always dependent on the experimental design and the potency of some compounds may be underestimated if the concentration of the target protein is similar to or higher than the measured IC_{50} value. Pieters and co-workers have recently reported a new type of inhibition assay based on cultured intestinal organoids [64], which when treated with the CT holotoxin swell up as fluid is transported across their epithelia. Toxin inhibition is quantified by measuring the reduction in organoid swelling. When inhibitors **22b** and **33** were re-evaluated using this new assay, they were found to be even more active than previously measured in the ELLA (IC_{50} = 34 pM for **22b** in organoid assay vs 160 pM in ELLA; IC_{50} = 15 pM for **33** in organoid assay vs 260 pM in ELLA). While enzyme-linked lectin assays will undoubtedly continue to be a popular method for easily evaluating and comparing different inhibitors, the intestinal organoid assay introduced by Pieters and co-workers is now the most sensitive and realistic in vitro assay available [65].

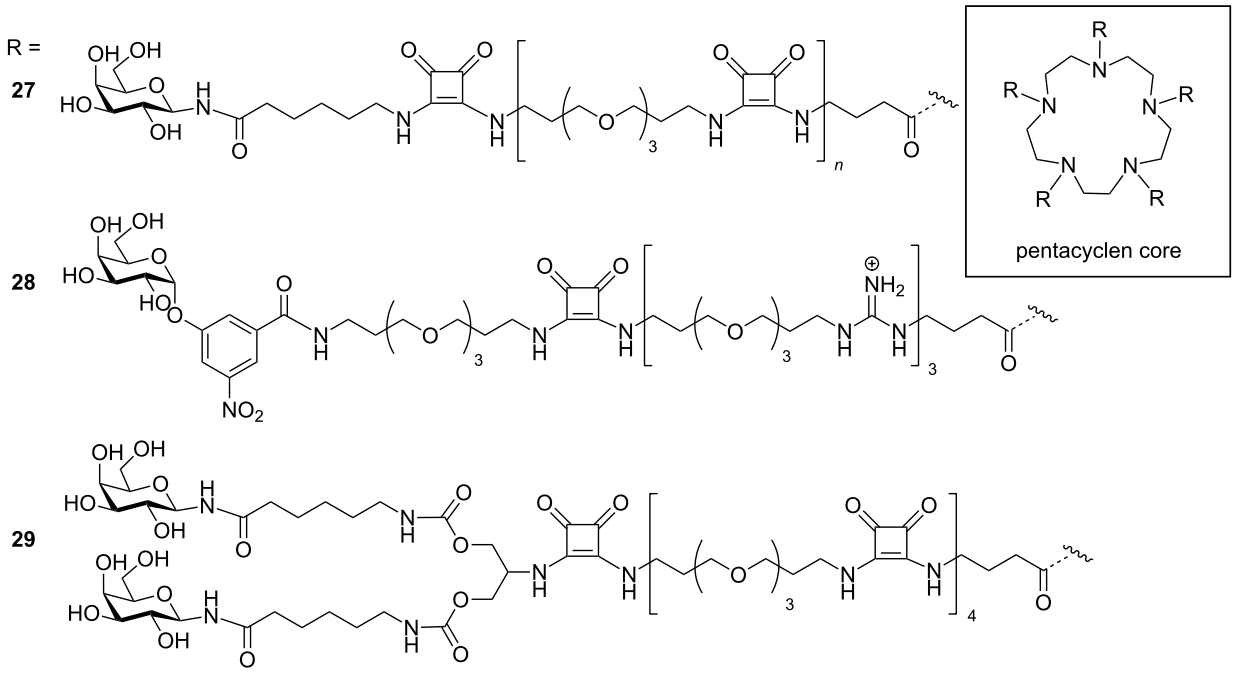


Figure 10: The star-shaped inhibitors reported by Fan, Hol and co-workers.

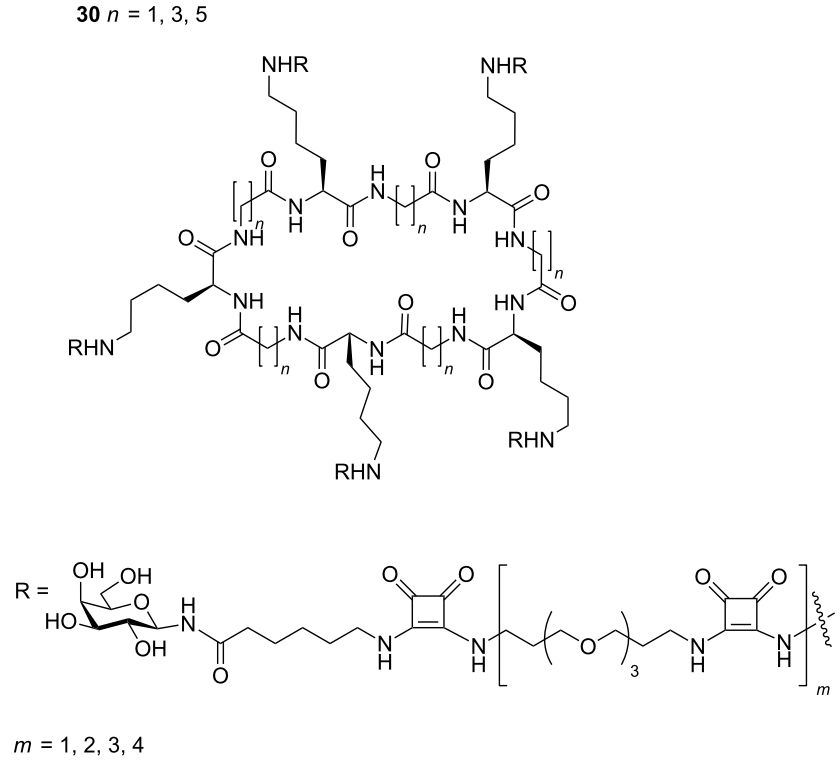


Figure 11: Differently sized cyclic decavalent peptide core designed by Zhang et al.

Branson et al. took a different approach to scaffold design in which they made a non-binding mutant of the target CTB protein [66], oxidised the N-terminal threonine residue of each subunit to an aldehyde and then chemically attached GM_1os ligands by oxime ligation (Figure 15). This neoglycoprotein was able to display the five copies of the carbohydrate ligand

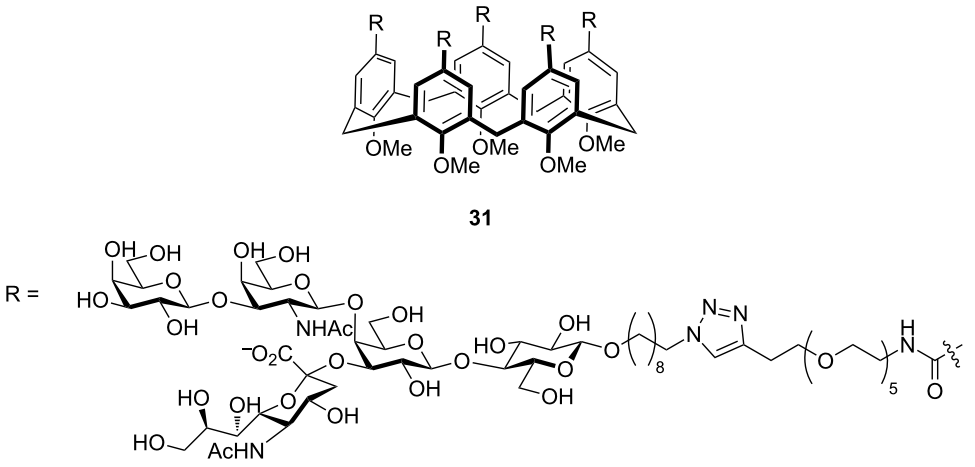


Figure 12: Calix[5]arene core-based pentavalent inhibitor designed by Garcia-Hartjes et al.

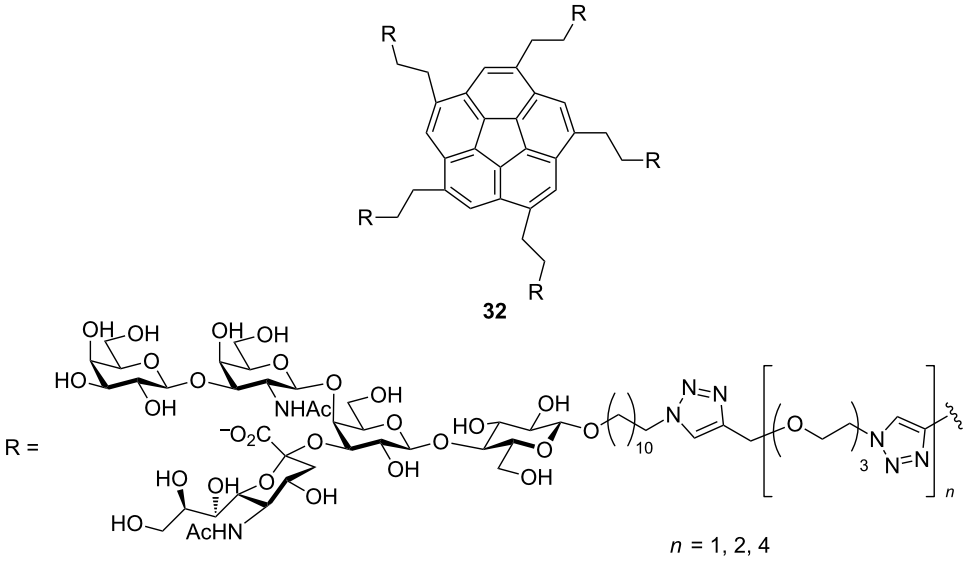


Figure 13: Corannulene core-based pentavalent inhibitor designed by Mattarella et al.

with appropriate spacing's to maximize interactions with the target protein. Dynamic light scattering and analytical ultracentrifugation demonstrated that the glycoprotein formed a 1:1 complex with the target CTB protein and was highly effective as an inhibitor with an IC_{50} value of 104 pM.

Multivalent inhibitors with more than five ligands

While pentavalent inhibitors are seductive as they match the symmetry with the CTB protein, many researchers have sought to exceed the valency of five. In many cases this is largely for convenience of preparation of polymers bearing multiple pendant groups, or to achieve inhibitors that are sufficiently long to cross-link the binding sites in a protein. Also, if multivalent molecules have not been specifically designed to match the distance between the target binding sites, then sometimes larger multivalent compounds are better. For example, Pieters and co-workers used a tryptophan fluorescence quenching assay to show that octavalent lactose-based dendrimer **34** (Figure 16) had a K_d value of 33 μM as compared to monovalent lactose derivative having a K_d value of 18,000 μM [51]. Hence, compound **34** displayed 545 fold more potency per lactose unit than monovalent lactose. In another report, they found that octavalent galactose-derived dendrimer **36** displayed excellent CT inhibition with an IC_{50} value of 12 μM and this was better than monovalent GM_1os (IC_{50} = 19 μM) [52]. From the collaboration work of Pieters and Bernardi, an ELISA assay confirmed

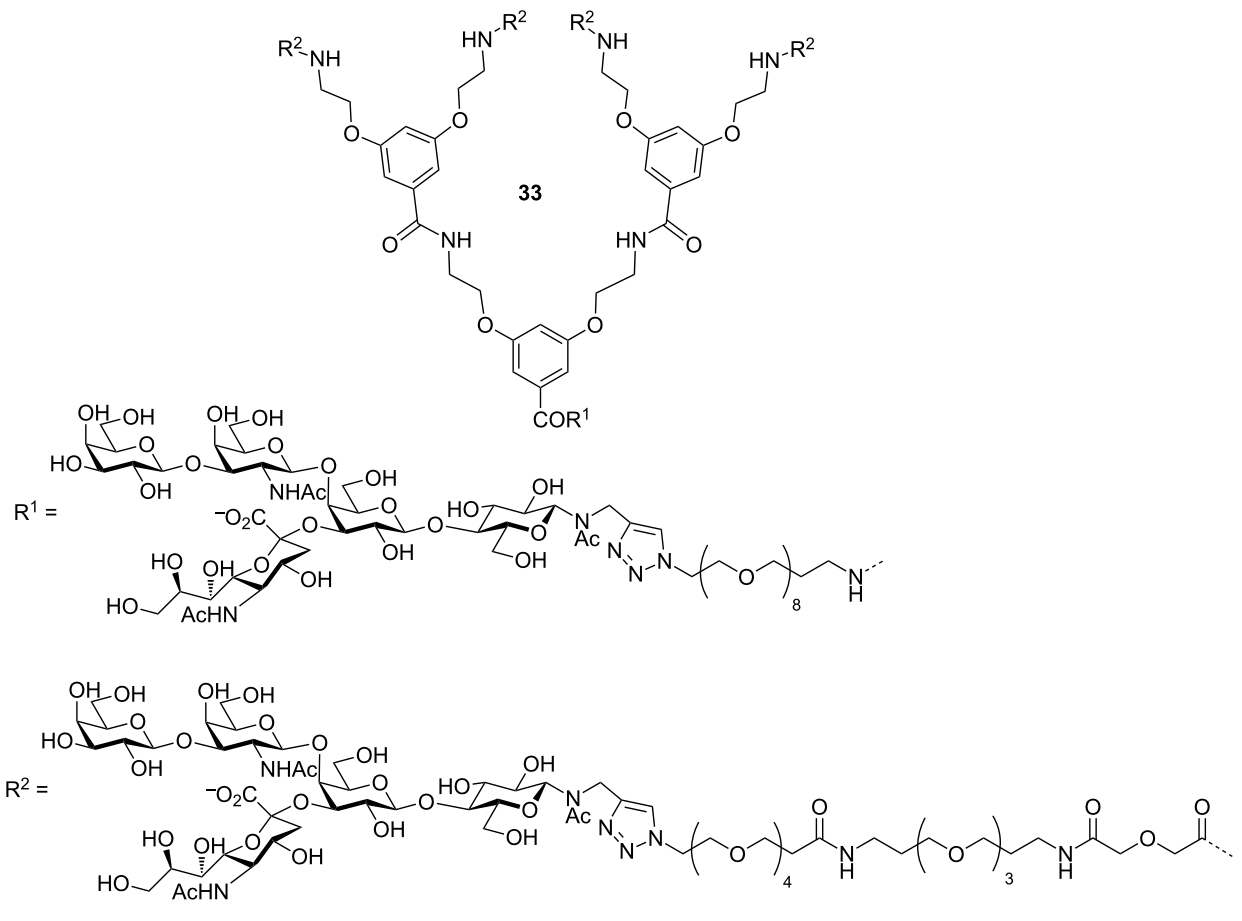


Figure 14: Pentavalent inhibitor designed by Pieters and co-workers.

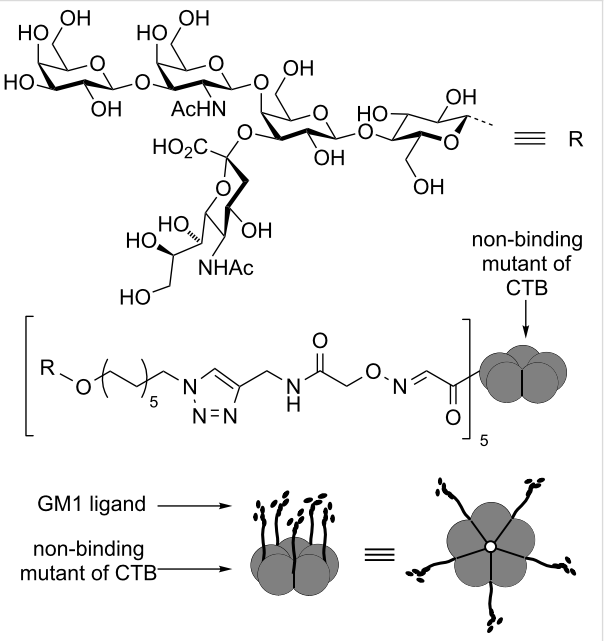


Figure 15: Neoglycoprotein inhibitor based on a non-binding mutant of CTB.

that compound **35** (Figure 16) was the most potent compound having an IC_{50} value less than 0.5 μM [53]. In another report, they reported that the octavalent GM_1os dendrimer complex **37** (Figure 16) displayed a 380,000-fold enhanced potency relative to monovalent GM_1 [54].

Polymeric scaffolds have also been used extensively over many years [67]. Some recent highlights have included using polymer backbones to identify GM1 analogues that can give enhanced multivalent interactions [68], evolving glycopolymers using exchangeable ligands [69], and tuning the way the ligands are connected to the polymer backbone for maximum interaction [70,71]. For example, using a fragment-based approach, Tran et al. synthesised and evaluated a library of polymer-based heterobifunctional ligands and found that some compounds showed low nanomolar multivalent inhibition [68]. Alpha-galactoside **38** (Figure 17) showed the highest activity when presented on the polymer scaffold with an IC_{50} value of 0.005 μM. In contrast, the IC_{50} value shown by a monomeric version of this heterobifunctional ligand **39** was in the millimolar range, similar to the compound *m*-nitrophenyl galactopyranoside (**4**).

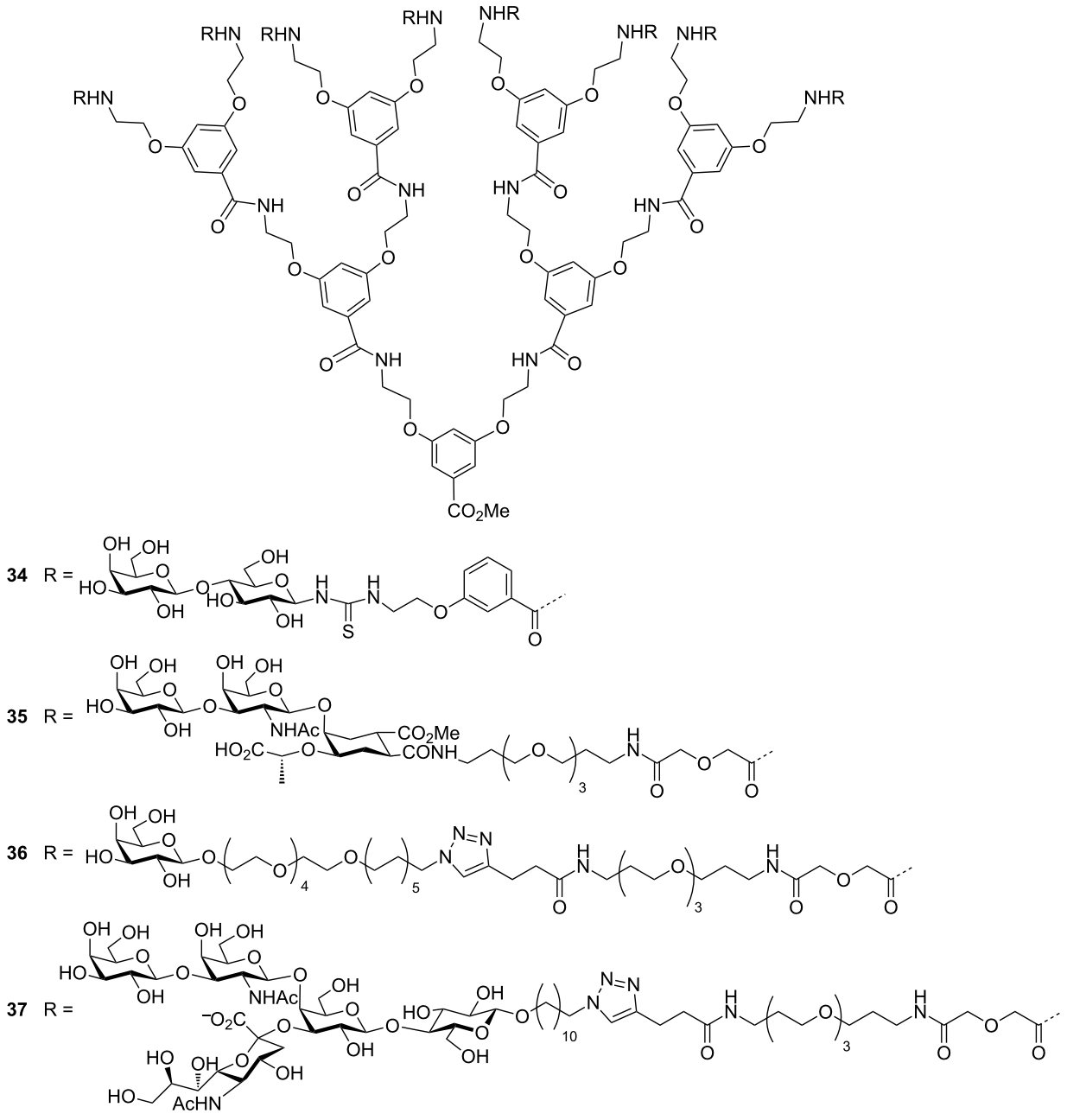


Figure 16: Octavalent inhibitor designed by Pieters, Bernardi and co-workers.

Fulton and co-workers developed a dynamic combinatorial library of glycopolymers employing exchangeable galactosyl or mannosyl hydrazide functions in conjunction with pendant benzaldehyde groups on the polymer backbone to produce exchangeable hydrazones, e.g., **40** (Figure 18) [69]. They were able to show that in the presence of LTB, the *E. coli* homologue of CTB, the polymer self-optimised its binding affinity for the protein by increasing the proportion of galactosyl residues in the backbone. In the presence of low concentrations of a dihydrazide cross-linking agent, these polymers can also be used to make crosslinked films on surfaces coated with bacterial toxin lectins [72].

Gibson and co-workers made a series of polyacrylates bearing pentafluorophenyl active ester groups which could be subsequently converted to polyacrylamides by reaction with amine linkers of varying lengths [71]. The attachment of galactosyl azides provided a series of glycopolymer inhibitors of CTB

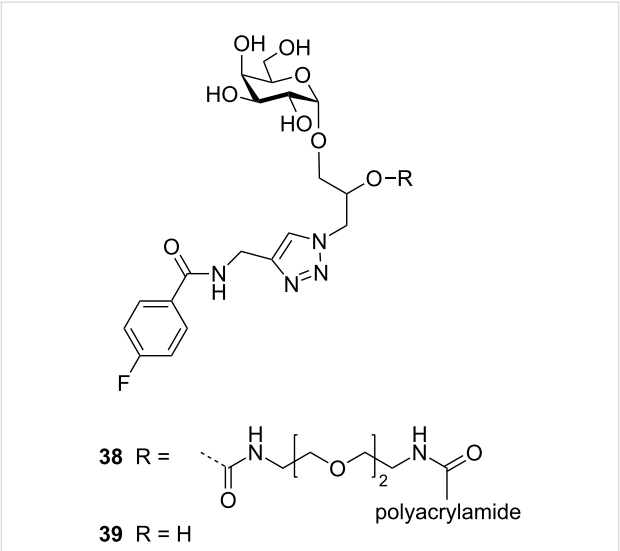


Figure 17: Hetero-bifunctional inhibitor designed by Bundle and co-workers.

(e.g., **41**). They were able to demonstrate that longer linkers between the carbohydrate and polymer backbone gave the best inhibition, probably because they were better able to reach into the binding pocket of the protein.

Multivalent scaffold bearing many galactosyl ligands need not be restricted to organic polymers. Gold nanoparticles coated in galactosyl ligands have been shown to be effective multivalent ligands for cholera toxin [73] and *E. coli* heat-labile toxin [74]. In these cases the objective of the studies was not to invoke inhibition, but rather to exploit the colour changes induced upon crosslinking the gold nanoparticles with CTB or LTB as a strategy for detecting the bacterial toxins.

Conclusion

Cholera and related diseases caused by other bacterial toxins remain a substantial threat to society. This challenge, and a molecular understanding of the basis of toxin action, has driven the development of diverse inhibitors over many years and this area of research continues to flourish with imaginative and novel strategies emerging for potential antiadhesive therapeutics. Further advances in our understanding of the structural biology of bacterial toxins, in particular the roles of secondary carbohydrate binding sites, will provide new directions for the future development of inhibitors, for example, fucosylated polymers [75], or hybrid inhibitors that can target both the blood group and the GM1 binding pockets. While other emerging, and sophisticated strategies for the use of multivalent scaffolds for displaying (dynamic) libraries of low affinity ligands may accelerate the process of finding effective mimics of the GM1 glycolipid that are simpler in structure and easier to develop into practical therapeutics. Furthermore, the introduction of diverse biophysical methods for studying inhibition mechanisms and novel inhibition assays using intestinal organoids are now providing better quality data and understanding of the action of multivalent inhibitors. The continuous innovation across this field will undoubtedly lead to many more exciting developments for years to come.

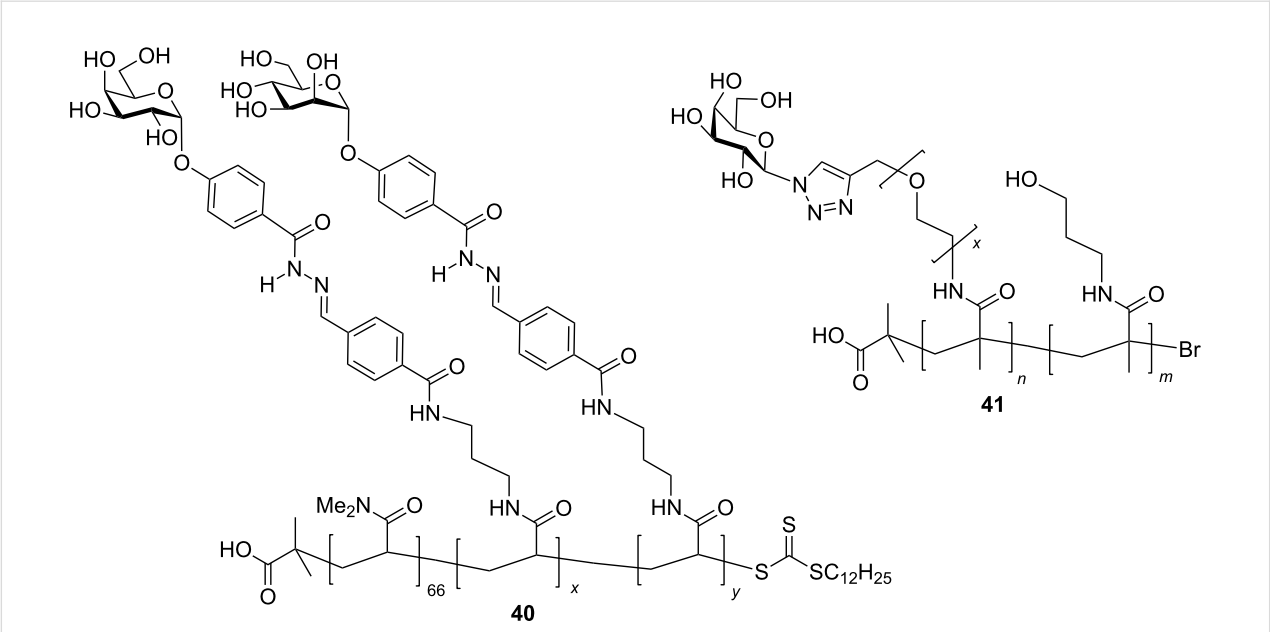


Figure 18: Glycopolymers with exchangeable sugar ligands and variable length linkers.

Acknowledgements

This review is part of a project that has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 746421.

ORCID® iDs

Vajinder Kumar - https://orcid.org/0000-0002-8825-3517
W. Bruce Turnbull - https://orcid.org/0000-0002-7352-0360

References

1. Bryce, J.; Boschi-Pinto, C.; Shibuya, K.; Black, R. E. *Lancet* **2005,** *365,* 1147–1152. doi:10.1016/S0140-6736(05)71877-8
2. WHO. *Wkly. Epidemiol. Rec.* **2016,** *38,* 433–440.
3. http://www.gideononline.com.
4. Ali, M.; Nelson, A. R.; Lopez, A. L.; Sack, D. A. *PLoS Negl. Trop. Dis.* **2015,** *9,* e0003832. doi:10.1371/journal.pntd.0003832
5. Merritt, E. A.; Hol, W. G. J. *Curr. Opin. Struct. Biol.* **1995,** *5,* 165–171. doi:10.1016/0959-440X(95)80071-9
6. Fan, E.; Merritt, E. A.; Verlinde, C. L. M. J.; Hol, W. G. J. *Curr. Opin. Struct. Biol.* **2000,** *10,* 680–686. doi:10.1016/S0959-440X(00)00152-4
7. Hol, W. G. J.; Sixma, T. K.; Merritt, E. A. *Handbook of Natural Toxins;* 1995; Vol. 8, pp 185–223.
8. van den Akker, F.; Sarfaty, S.; Twiddy, E. M.; Connell, T. D.; Holmes, R. K.; Hol, W. G. J. *Structure* **1996,** *4,* 665–678. doi:10.1016/S0969-2126(96)00073-1
9. Fraser, M. E.; Chernaia, M. M.; Kozlov, Y. V.; James, M. N. G. *Nat. Struct. Mol. Biol.* **1994,** *1,* 59–64. doi:10.1038/nsb0194-59
10. Stein, P. E.; Boodhoo, A.; Tyrrell, G. J.; Brunton, J. L.; Read, R. J. *Nature* **1992,** *355,* 748–750. doi:10.1038/355748a0
11. Stein, P. E.; Boodhoo, A.; Armstrong, G. D.; Cockle, S. A.; Klein, M. H.; Read, R. J. *Structure* **1994,** *2,* 45–57. doi:10.1016/S0969-2126(00)00007-1
12. Merritt, E. A.; Sarfaty, S.; van den Akker, F.; L'Hoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Sci.* **1994,** *3,* 166–175. doi:10.1002/pro.5560030202
13. Heggelund, J. E.; Burschowsky, D.; Bjørnestad, V. A.; Hodnik, V.; Anderluh, G.; Krengel, U. *PLoS Pathog.* **2016,** *12,* e1005567. doi:10.1371/journal.ppat.1005567
14. Merritt, E. A.; Kuhn, P.; Sarfaty, S.; Erbe, J. L.; Holmes, R. K.; Hol, W. G. J. *J. Mol. Biol.* **1998,** *282,* 1043–1059. doi:10.1006/jmbi.1998.2076
15. De Haan, L.; Hirst, T. R. *Mol. Membr. Biol.* **2004,** *21,* 77–92. doi:10.1080/09687680410001663267
16. Tomasi, M.; Battistini, A.; Araco, A.; Roda, L. G.; D'Agnolo, G. *Eur. J. Biochem.* **1979,** *93,* 621–627. doi:10.1111/j.1432-1033.1979.tb12862.x
17. Ampapathi, R. S.; Creath, A. L.; Lou, D. I.; Craft, J. W., Jr.; Blanke, S. R.; Legge, G. B. *J. Mol. Biol.* **2008,** *377,* 748–760. doi:10.1016/j.jmb.2007.12.075
18. Zhang, R.-G.; Scott, D. L.; Westbrook, M. L.; Nance, S.; Spangler, B. D.; Shipley, G. G.; Westbrook, E. M. *J. Mol. Biol.* **1995,** *251,* 563–573. doi:10.1006/jmbi.1995.0456
19. Zhang, R.-G.; Westbrook, M. L.; Westbrook, E. M.; Scott, D. L.; Otwinowski, Z.; Maulik, P. R.; Reed, R. A.; Shipley, G. G. *J. Mol. Biol.* **1995,** *251,* 550–562. doi:10.1006/jmbi.1995.0455
20. Turnbull, W. B.; Precious, B. L.; Homans, S. W. *J. Am. Chem. Soc.* **2004,** *126,* 1047–1054. doi:10.1021/ja0378207
21. Johannes, L.; Parton, R. G.; Bassereau, P.; Mayor, S. *Nat. Rev. Mol. Cell Biol.* **2015,** *16,* 311–321. doi:10.1038/nrm3968
22. Holmner, A.; Lebens, M.; Teneberg, S.; Angstrom, J.; Ökvist, M.; Krengel, U. *Structure* **2004,** *12,* 1655–1667. doi:10.1016/j.str.2004.06.022
23. Holmner, A.; Askarieh, G.; Oekvist, M.; Krengel, U. *J. Mol. Biol.* **2007,** *371,* 754–764. doi:10.1016/j.jmb.2007.05.064
24. Mandal, P. K.; Branson, T. R.; Hayes, E. D.; Ross, J. F.; Gavín, J. A.; Daranas, A. H.; Turnbull, W. B. *Angew. Chem., Int. Ed.* **2012,** *51,* 5143–5146. doi:10.1002/anie.201109068
25. Vasile, F.; Reina, J. J.; Potenza, D.; Heggelund, J. E.; Mackenzie, A.; Krengel, U.; Bernardi, A. *Glycobiology* **2014,** *24,* 766–778. doi:10.1093/glycob/cwu040
26. St. Hilaire, P. M.; Boyd, M. K.; Toone, E. J. *Biochemistry* **1994,** *33,* 14452–14463. doi:10.1021/bi00252a011
27. Holmner, A.; Mackenzie, A.; Krengel, U. *FEBS Lett.* **2010,** *584,* 2548–2555. doi:10.1016/j.febslet.2010.03.050
28. Wands, A. M.; Fujita, A.; McCombs, J. E.; Cervin, J.; Dedic, B.; Rodriguez, A. C.; Nischan, N.; Bond, M. R.; Mettlen, M.; Trudgian, D. C.; Lemoff, A.; Quiding-Järbrink, M.; Gustavsson, B.; Steentoft, C.; Clausen, H.; Mirzaei, H.; Teneberg, S.; Yrlid, U.; Kohler, J. J. *eLife* **2015,** *4,* e09545. doi:10.7554/eLife.09545
29. Verlinde, C. L. M. J.; Hol, W. G. J. *Structure* **1994,** *2,* 577–587. doi:10.1016/S0969-2126(00)00060-5
30. Branson, T. R.; Turnbull, W. B. *Chem. Soc. Rev.* **2013,** *42,* 4613–4622. doi:10.1039/C2CS35430F
31. Zuilhof, H. *Acc. Chem. Res.* **2016,** *49,* 274–285. doi:10.1021/acs.accounts.5b00480
32. Bernardi, A.; Checchia, A.; Brocca, P.; Sonnino, S.; Zuccotto, F. *J. Am. Chem. Soc.* **1999,** *121,* 2032–2036. doi:10.1021/ja983567c
33. Bernardi, A.; Arosio, D.; Sonnino, S. *Neurochem. Res.* **2002,** *27,* 539–545. doi:10.1023/A:1020251428217
34. Bernardi, A.; Carrettoni, L.; Ciponte, A. G.; Monti, D.; Sonnino, S. *Bioorg. Med. Chem. Lett.* **2000,** *10,* 2197–2200. doi:10.1016/S0960-894X(00)00428-5
35. Cheshev, P.; Morelli, L.; Marchesi, M.; Podlipnik, C.; Bergström, M.; Bernardi, A. *Chem. – Eur. J.* **2010,** *16,* 1951–1967. doi:10.1002/chem.200902469
36. Heggelund, J. E.; Mackenzie, A.; Martinsen, T.; Heim, J. B.; Cheshev, P.; Bernardi, A.; Krengel, U. *Sci. Rep.* **2017,** *7,* No. 2326. doi:10.1038/s41598-017-02179-0
37. Merritt, E. A.; Sixma, T. K.; Kalk, K. H.; van Zanten, B. A. M.; Hol, W. G. J. *Mol. Microbiol.* **1994,** *13,* 745–753. doi:10.1111/j.1365-2958.1994.tb00467.x
38. Merritt, E. A.; Sarfaty, S.; Feil, I. K.; Hol, W. G. *Structure* **1997,** *5,* 1485–1499. doi:10.1016/S0969-2126(97)00298-0
39. Minke, W. E.; Roach, C.; Hol, W. G. J.; Verlinde, C. L. M. J. *Biochemistry* **1999,** *38,* 5684–5692. doi:10.1021/bi982649a
40. Mitchell, D. D.; Pickens, J. C.; Korotkov, K.; Fan, E.; Hol, W. G. J. *Bioorg. Med. Chem.* **2004,** *12,* 907–920. doi:10.1016/j.bmc.2003.12.019
41. Vrasidas, I.; Kemmink, J.; Liskamp, R. M. J.; Pieters, R. J. *Org. Lett.* **2002,** *4,* 1807–1808. doi:10.1021/ol025909w
42. Ramos-Soriano, J.; Niss, U.; Angulo, J.; Angulo, M.; Moreno-Vargas, A. J.; Carmona, A. T.; Ohlson, S.; Robina, I. *Chem. – Eur. J.* **2013,** *19,* 17989–18003. doi:10.1002/chem.201302786

43. Mammen, M.; Chio, S.-K.; Whitesides, G. M. *Angew. Chem., Int. Ed.* **1998,** *37,* 2754–2794. doi:10.1002/(SICI)1521-3773(19981102)37:20<2754::AID-ANIE2754>3.0.CO;2-3
44. Badjic, J. D.; Nelson, A.; Cantrill, S. J.; Turnbull, W. B.; Stoddart, J. F. *Acc. Chem. Res.* **2005,** *38,* 723–732. doi:10.1021/ar040223k
45. Cecioni, S.; Imberty, A.; Vidal, S. *Chem. Rev.* **2015,** *115,* 525–561. doi:10.1021/cr500303t
46. Pickens, J. C.; Mitchell, D. D.; Liu, J.; Tan, X.; Zhang, Z.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *Chem. Biol.* **2004,** *11,* 1205–1215. doi:10.1016/j.chembiol.2004.06.008
47. Arosio, D.; Fontanella, M.; Baldini, L.; Mauri, L.; Bernardi, A.; Casnati, A.; Sansone, F.; Ungaro, R. *J. Am. Chem. Soc.* **2005,** *127,* 3660–3661. doi:10.1021/ja0444029
48. Leaver, D. J.; Dawson, R. M.; White, J. M.; Polyzos, A.; Hughes, A. B. *Org. Biomol. Chem.* **2011,** *9,* 8465–8474. doi:10.1039/c1ob06317k
49. Leaver, D. J.; Hughes, A. B.; Dawson, R. M.; Postma, A.; Malic, N.; Polyzos, A. *RSC Adv.* **2014,** *4,* 14868–14871. doi:10.1039/C3RA47500J
50. Liu, J.; Begley, D.; Mitchell, D. D.; Verlinde, C. L. M. J.; Varani, G.; Fan, E. *Chem. Biol. Drug Des.* **2008,** *71,* 408–419. doi:10.1111/j.1747-0285.2008.00648.x
51. Vrasidas, I.; de Mol, N. J.; Liskamp, R. M. J.; Pieters, R. J. *Eur. J. Org. Chem.* **2001,** 4685–4692. doi:10.1002/1099-0690(200112)2001:24<4685::AID-EJOC4685>3.0.CO;2-9
52. Branderhorst, H. M.; Liskamp, R. M. J.; Visser, G. M.; Pieters, R. J. *Chem. Commun.* **2007,** 5043–5045. doi:10.1039/b711070g
53. Arosio, D.; Vrasidas, I.; Valentini, P.; Liskamp, R. M. J.; Pieters, R. J.; Bernardi, A. *Org. Biomol. Chem.* **2004,** *2,* 2113–2124. doi:10.1039/b405344c
54. Pukin, A. V.; Branderhorst, H. M.; Sisu, C.; Weijers, C. A. G. M.; Gilbert, M.; Liskamp, R. M. J.; Visser, G. M.; Zuilhof, H.; Pieters, R. J. *ChemBioChem* **2007,** *8,* 1500–1503. doi:10.1002/cbic.200700266
55. Sisu, C.; Baron, A. J.; Branderhorst, H. M.; Connell, S. D.; Weijers, C. A. G. M.; de, V. R.; Hayes, E. D.; Pukin, A. V.; Gilbert, M.; Pieters, R. J.; Zuilhof, H.; Visser, G. M.; Turnbull, W. B. *ChemBioChem* **2009,** *10,* 329–337. doi:10.1002/cbic.200800550
56. Fu, O.; Pukin, A. V.; van Ufford, H. C. Q.; Branson, T. R.; Thies-Weesie, D. M. E.; Turnbull, W. B.; Visser, G. M.; Pieters, R. J. *ChemistryOpen* **2015,** *4,* 471–477. doi:10.1002/open.201500006
57. Kumar, V.; Yadav, N.; Kartha, K. P. R. *Carbohydr. Res.* **2016,** *431,* 47–55. doi:10.1016/j.carres.2016.05.011
58. Fan, E.; Zhang, Z.; Minke, W. E.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J. *J. Am. Chem. Soc.* **2000,** *122,* 2663–2664. doi:10.1021/ja993388a
59. Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000,** *403,* 669–672. doi:10.1038/35001095
60. Merritt, E. A.; Zhang, Z.; Pickens, J. C.; Ahn, M.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002,** *124,* 8818–8824. doi:10.1021/ja0202560
61. Zhang, Z.; Merritt, E. A.; Ahn, M.; Roach, C.; Hou, Z.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Am. Chem. Soc.* **2002,** *124,* 12991–12998. doi:10.1021/ja027584k
62. Zhang, Z.; Liu, J.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. *J. Org. Chem.* **2004,** *69,* 7737–7740. doi:10.1021/jo0489770
63. Garcia-Hartjes, J.; Bernardi, S.; Weijers, C. A. G. M.; Wennekes, T.; Gilbert, M.; Sansone, F.; Casnati, A.; Zuilhof, H. *Org. Biomol. Chem.* **2013,** *11,* 4340–4349. doi:10.1039/C3OB40515J
64. Mattarella, M.; Garcia-Hartjes, J.; Wennekes, T.; Zuilhof, H.; Siegel, J. S. *Org. Biomol. Chem.* **2013,** *11,* 4333–4339. doi:10.1039/C3OB40438B
65. Zomer-van Ommen, D. D.; Pukin, A. V.; Fu, O.; Quarles van Ufford, L. H. C.; Janssens, H. M.; Beekman, J. M.; Pieters, R. J. *J. Med. Chem.* **2016,** *59,* 6968–6972. doi:10.1021/acs.jmedchem.6b00770
66. Branson, T. R.; McAllister, T. E.; Garcia-Hartjes, J.; Fascione, M. A.; Ross, J. F.; Warriner, S. L.; Wennekes, T.; Zuilhof, H.; Turnbull, W. B. *Angew. Chem., Int. Ed.* **2014,** *53,* 8323–8327. doi:10.1002/anie.201404397
67. Schengrund, C. L.; Ringler, N. J. *J. Biol. Chem.* **1989,** *264,* 13233–13237.
68. Tran, H.-A.; Kitov, P. I.; Paszkiewicz, E.; Sadowska, J. M.; Bundle, D. R. *Org. Biomol. Chem.* **2011,** *9,* 3658–3671. doi:10.1039/c0ob01089h
69. Mahon, C. S.; Fascione, M. A.; Sakonsinsiri, C.; McAllister, T. E.; Turnbull, W. B.; Fulton, D. A. *Org. Biomol. Chem.* **2015,** *13,* 2756–2761. doi:10.1039/C4OB02587C
70. Polizzotti, B. D.; Kiick, K. L. *Biomacromolecules* **2006,** *7,* 483–490. doi:10.1021/bm050672n
71. Richards, S.-J.; Jones, M. W.; Hunaban, M.; Haddleton, D. M.; Gibson, M. I. *Angew. Chem., Int. Ed.* **2012,** *51,* 7812–7816. doi:10.1002/anie.201202945
72. Mahon, C. S.; McGurk, C. J.; Watson, S. M. D.; Fascione, M. A.; Sakonsinsiri, C.; Turnbull, W. B.; Fulton, D. A. *Angew. Chem., Int. Ed.* **2017,** *56,* 12913–12918. doi:10.1002/anie.201706379
73. Schofield, C. L.; Field, R. A.; Russell, D. A. *Anal. Chem.* **2007,** *79,* 1356–1361. doi:10.1021/ac061462j
74. Poonthiyil, V.; Golovko, V. B.; Fairbanks, A. J. *Org. Biomol. Chem.* **2015,** *13,* 5215–5223. doi:10.1039/C5OB00447K
75. Wands, A. M.; Cervin, J.; Huang, H.; Zhang, Y.; Youn, G.; Brautigam, C. A.; Matson Dzebo, M.; Bjorklund, P.; Wallenius, V.; Bright, D. K.; Bennett, C. S.; Wittung-Stafshede, P.; Sampson, N. S.; Yrlid, U.; Kohler, J. J. *ACS Infect. Dis.* **2018,** in press. doi:10.1021/acsinfecdis.7b00085

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Anomeric modification of carbohydrates using the Mitsunobu reaction

Julia Hain[1], Patrick Rollin[*2], Werner Klaffke[3] and Thisbe K. Lindhorst[*1]

Review

Open Access

Address:
[1]Christiana Albertina University of Kiel, Otto Diels Institute of Organic Chemistry, Otto-Hahn-Platz 3–4, D-24118 Kiel, Germany, Fax: +49 431 8807410, [2]Université d'Orléans et CNRS, ICOA, UMR 7311, BP 6759, 45067 Orléans, France, Fax: +33 238 417281 and [3]Haus der Technik e.V., Hollestr. 1, 45127 Essen, Germany, Fax: +49 201 1803269

Email:
Patrick Rollin[*] - Patrick.Rollin@univ-orleans.fr; Thisbe K. Lindhorst[*] - tklind@oc.uni-kiel.de

* Corresponding author

Keywords:
anomeric stereoselectivity; carbohydrates; glycoside synthesis; Mitsunobu reaction

Beilstein J. Org. Chem. **2018,** *14,* 1619–1636.
doi:10.3762/bjoc.14.138

Received: 21 March 2018
Accepted: 06 June 2018
Published: 29 June 2018

This article is part of the Thematic Series "The glycosciences". Dedicated to Professor Joachim Thiem in recognition of his constant inspiration and support as a teacher, colleague and friend.

Associate Editor: S. Flitsch

Abstract

The Mitsunobu reaction basically consists in the conversion of an alcohol into an ester under inversion of configuration, employing a carboxylic acid and a pair of two auxiliary reagents, mostly triphenylphosphine and a dialkyl azodicarboxylate. This reaction has been frequently used in carbohydrate chemistry for the modification of sugar hydroxy groups. Modification at the anomeric position, leading mainly to anomeric esters or glycosides, is of particular importance in the glycosciences. Therefore, this review focuses on the use of the Mitsunobu reaction for modifications of sugar hemiacetals. Strikingly, unprotected sugars can often be converted regioselectively at the anomeric center, whereas in other cases, the other hydroxy groups in reducing sugars have to be protected to achieve good results in the Mitsunobu procedure. We have reviewed on the one hand the literature on anomeric esterification, including glycosyl phosphates, and on the other hand glycoside synthesis, including S- and N-glycosides. The mechanistic details of the Mitsunobu reaction are discussed as well as this is important to explain and predict the stereoselectivity of anomeric modifications under Mitsunobu conditions. Though the Mitsunobu reaction is often not the first choice for the anomeric modification of carbohydrates, this review shows the high value of the reaction in many different circumstances.

Introduction

Fifty years ago, Oyo Mitsunobu reported a preparation of esters from alcohols and carboxylic acids supported by two auxiliary reagents, diethyl azodicarboxylate (DEAD) and triphenylphosphine [1]. This reaction has ever since become known as the "Mitsunobu reaction", being a frequently utilized tool in organic synthesis. In 1981, Mitsunobu published a first review about this reaction, entitled "The Use of Diethyl Azodicarboxylate and Triphenylphosphine in Synthesis and Transformation

of Natural Products" [2]. Thereafter, several further general reviews have appeared [3-6], owing to the spectacular development of diversified synthetic applications of the Mitsunobu reaction, whilst the long term debate about the mechanism of this reaction was still ongoing [7-12].

The standard Mitsunobu reaction involves coupling of an alcohol and a nucleophile in a dehydrative S_N2 process activated by a reactive combination of a triaryl- or trialkylphosphine as reducing agent and a dialkyl azodicarboxylate as oxidant. In a redox process, the phosphine species is oxidized to the respective phosphine oxide and the azo reagent is reduced to the corresponding 1,2-hydrazinodicarboxylate (Scheme 1). As we have frequently utilized this valuable reaction in carbohydrate chemistry, in this account we have compiled literature, where the Mitsunobu reaction was used for the anomeric modification of carbohydrates.

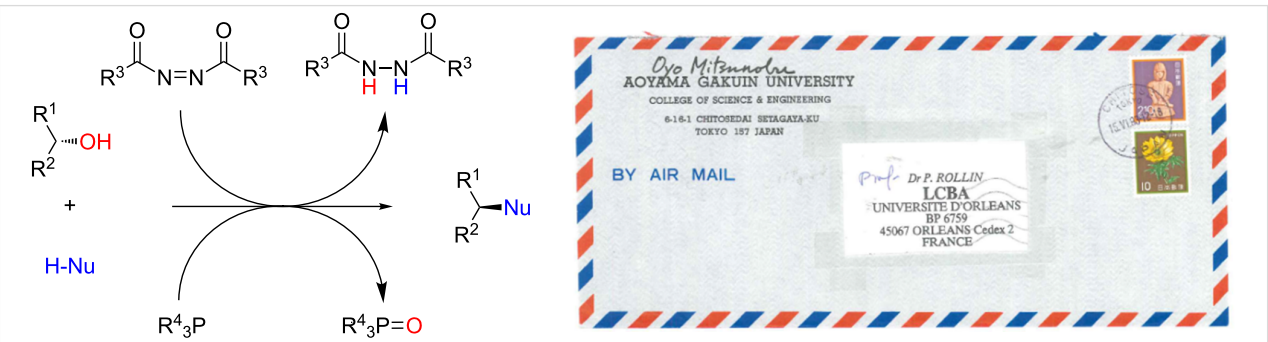
The reaction proceeds under mild, neutral conditions that are compatible with a wide range of functional groups. In the case where a stereogenic center is involved, the reaction takes place with stereochemical inversion [6]. The reaction partners are mostly primary or secondary alcohols, while the nucleophilic species needs to be acidic [13] with a $pK_a < 11$. Otherwise the azo reagent would compete with the acidic nucleophile and participate in the substitution reaction [14]. Various compounds comply with that condition: carboxylic acids, phenols, hydrazoic acid, some other NH acids, and thiols. The standard azo reagents used are diethyl- (DEAD) or diisopropyl- (DIAD) azodicarboxylate. However, alternative reagents such as azodicarboxamides [15,16] or stabilized phosphoranes were also developed to allow reaction with nucleophiles of weaker acidity. The typical phosphine reagents are triphenyl- (Ph_3P) or tributylphosphine (n-Bu_3P). In recent years, advances have been made using solid supported reagents, thus facilitating work-up conditions [17,18]. The polarity of the commonly aprotic solvents used in the Mitsunobu reaction, including toluene, tetrahydrofuran or dimethylformamide, has been shown to be influential in terms of efficacy and stereoselectivity [19].

Since its infancy, the Mitsunobu reaction has found applications in carbohydrate chemistry, as its broad scope and mild conditions are ideal for the formation of conjugates with sensitive natural products. Standard applications of the Mitsunobu reaction in glycochemistry have mostly dealt with the functionalization of the primary hydroxy group of sugars and, to a lesser extent, with modifications of the secondary alcohol array in carbohydrate rings [2-6], for example for halogenation [20]. However, the Mitsunobu reaction can also be profitably utilized for the anomeric modification of carbohydrates. Hence, we have focused this review on the utilization of the Mitsunobu reaction for manipulations of the carbohydrate hemiacetal, where reducing (anomerically unprotected) sugars react as the alcohol component to be either converted into glycosides or into other anomerically modified carbohydrate derivatives. We intend to provide a critical survey as well as a source of inspiration, even more so as glycosylation remains a challenge in carbohydrate chemistry.

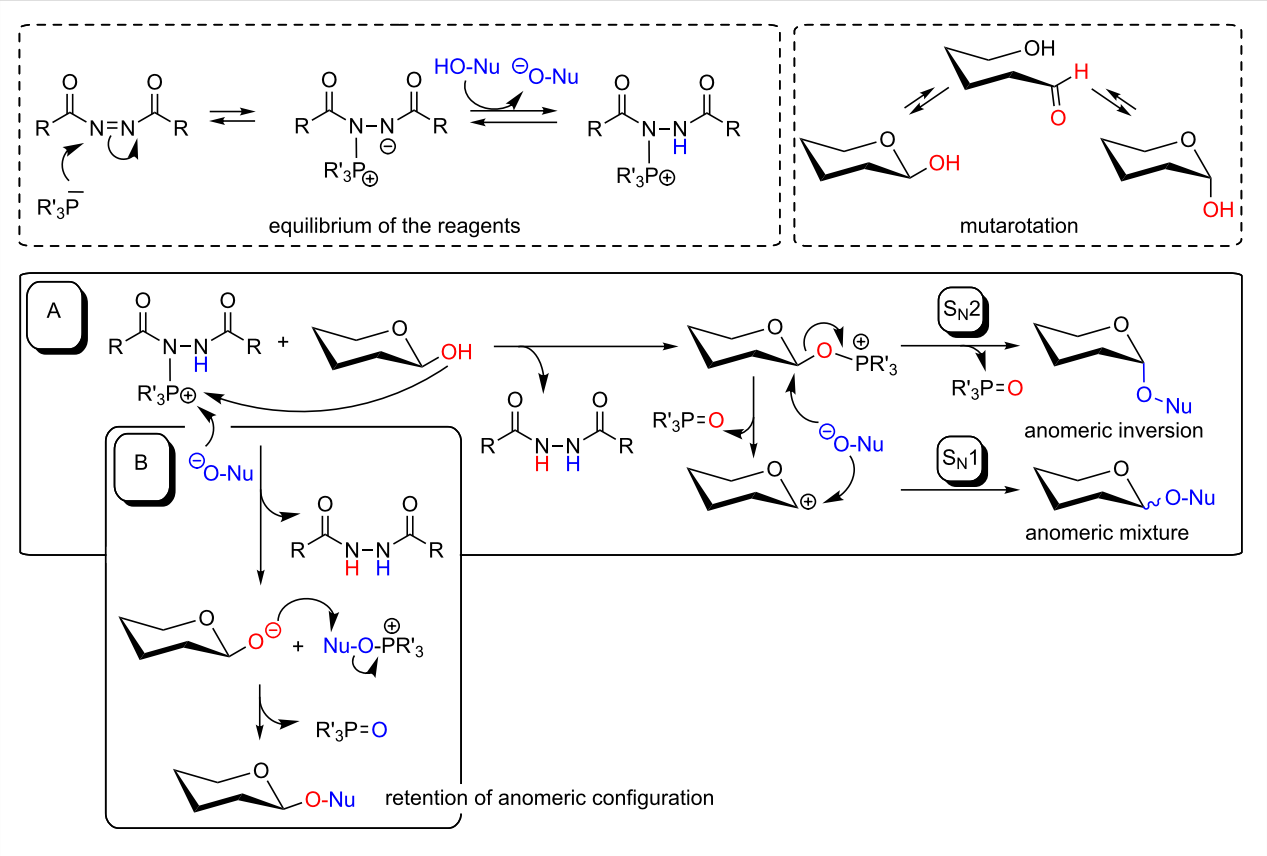
Review

Mechanistic considerations

Since Mitsunobu's postulate of a three-reaction-step mechanism in 1981 [2] many further mechanistic investigations have been performed and reported [3,7,8,19]. To rationalize the outcome of the Mitsunobu reaction with reducing sugars, special mechanistic considerations have to be taken into account. On the one hand, the equilibrium between the azodicarboxylate, the phosphine, and the acidic component, Nu-OH, is important (cf. Scheme 2, left dashed box). On the other hand, mutarotation of the sugar hemiacetal has to be discussed to predict the stereochemical outcome of the reaction. Mutarotation results in an equilibrium of both, α- and β-anomers (Scheme 2, right dashed box). However, full anomerization is often not observed as the rate and the extent of



Scheme 1: Left: The Mitsunobu reaction is essentially a nucleophilic substitution of alcohols occurring with inversion of configuration at the alcohol stereocenter. The auxiliary reagents are involved in a redox process. Right: Original correspondence with Professor Oyo Mitsunobu (letter to P.R. in 1993).



Scheme 2: Mechanistic considerations on the Mitsunobu reaction with carbohydrate hemiacetals (depicted in simplified form). Two equilibria are relevant for this reaction (top dashed boxes), (i) the formation of the "Mitsunobu reagent" from the employed azodicarboxylate, phosphine and the acidic nucleophile, and (ii) the mutarotation equilibrium of the reducing sugar in solution. These can give rise to at least two different reaction pathways, A and B, as explained in the main text. Depending on various parameters, the anomerically modified sugar, a glycoside or an anomeric ester, respectively, is obtained with full inversion of anomeric configuration or as anomeric mixture (A), or with retention of the anomeric configuration via *O*-alkylation (B). For clarity both reaction pathways are exemplified with only one sugar anomer.

mutarotation depends on various parameters such as anchimeric effects of neighboring groups and the reaction conditions. Hence it has been frequently observed in Mitsunobu reactions with carbohydrate hemiacetals, that sugar anomerization is either absent or slower than the formation of the *O*-glycosyloxyphosphonium salt, which can play the intermediate during the reaction (Scheme 2, pathway A) [21]. Another possible explanation for limited anomerization lies in the different stability of anomeric glycosyloxyphosphonium salts, where one anomer can be sterically favored over the other, thereby pushing the equilibrium to a product with the respective anomeric configuration. Regardless of the rate of mutarotation, the Mitsunobu reaction can proceed through a mechanistic pathway A or B as depicted in Scheme 2. Especially when the sugar alcohol is not sterically hindered, phosphorus transfer occurs to yield a phosphine-activated anomeric alcohol (a glycosyloxyphosphonium ion, pathway A). This in turn can be attacked by the deprotonated nucleophile resulting in an anomerically modified carbohydrate with inversion of configuration at the anomeric center, according to a S_N2 mechanism. Pathway A can also proceed through a S_N1 mechanism when the intermediate glycosyloxyphosphonium ion is less stable. Then, it can decompose into the corresponding anomeric oxocarbenium ion and phosphine oxide. The oxocarbenium ion would then react with the NuO^- anion in a S_N1 mechanism. While this would lead to racemization under normal circumstances, in most carbohydrates, participation effects of neighboring groups in the vicinity (typically at the 2-position of the sugar ring) affect the reaction outcome, favoring nucleophilic attack from a preferred face of the sugar ring [22,23]. Grynkiewicz and colleagues have discussed anchimeric assistance even when no protecting group is present at C-2, assuming a Brigl's anhydride type intermediate [24]. In the absence of a substituent at C-2, however, typically poor stereoselectivity is observed in Mitsunobu reactions with carbohydrate hemiacetals, indicating a S_N1-type pathway A of the reaction [25].

The Mitsunobu reaction can also follow a different pathway B (Scheme 2), as first suggested by Hughes [13] and later by Ahn et al. [26]. Assuming that the alcohol is sterically hindered

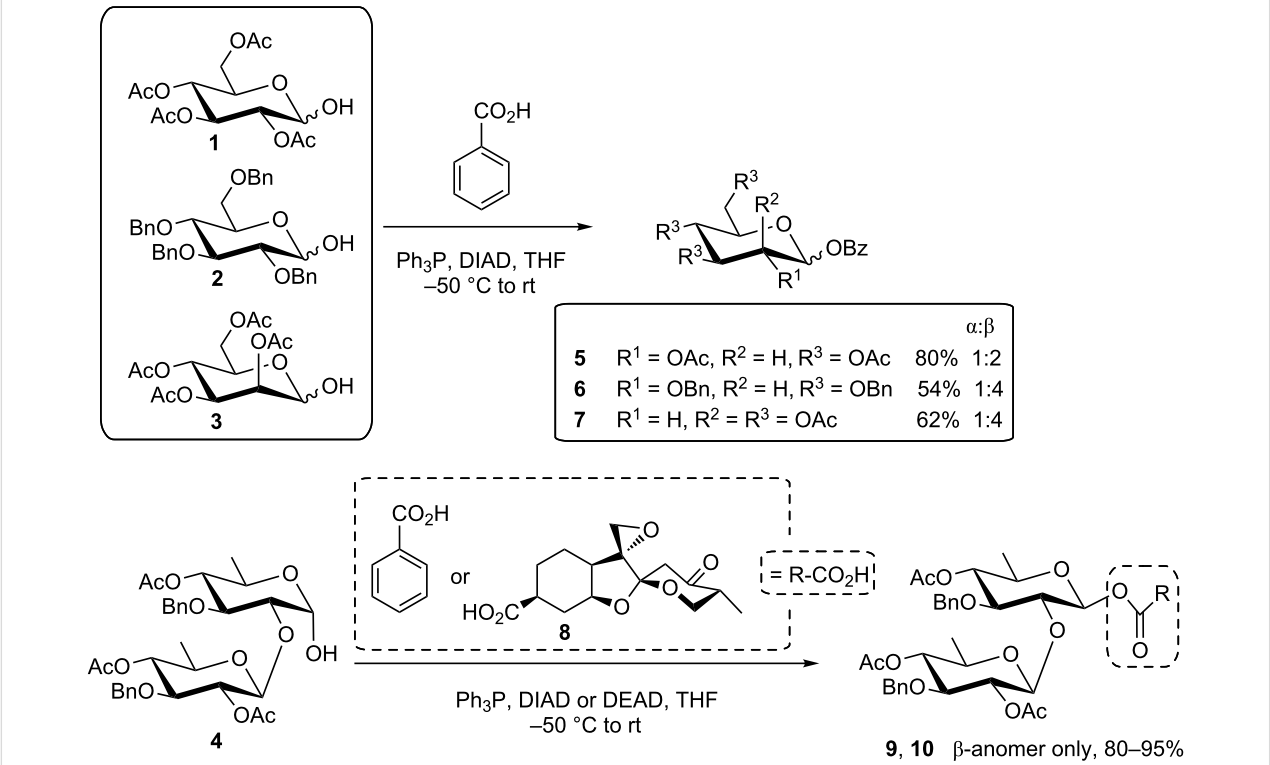
and thus represents a relatively weak nucleophile, the deprotonated acidic partner, NuO^-, can react with the phosphonium intermediate first to afford an intermediate Nu-O-PR'$_3$. In the case where a carboxylic acid is used, Nu-O-PR'$_3$ represents an acyloxyphosphonium ion. This in turn reacts with the anomeric oxyanion to furnish the anomerically modified sugar with retention of configuration via anomeric *O*-alkylation. This mechanistic proposal is in agreement with observations by Lubineau et al., who could correlate the acidity of the employed nucleophile with the anomeric outcome of the Mitsunobu reaction [27].

Both reaction pathways, A and B, have a "raison d'être" in addressing different outcomes of the Mitsunobu reaction, which vary depending on the substrates used. While these variables make the already complex Mitsunobu reaction even more demanding, they can also be manipulated to one's advantage, for example for the stereoselective formation of β-mannosides [28].

Reactions with protic acids to achieve anomeric esters

The first application of the Mitsunobu reaction involved esterification of a secondary alcohol. Although an anomeric OH group cannot be regarded as a classical secondary alcohol group but as a hemiacetal OH, it can be successfully involved in Mitsunobu reactions to achieve 1-*O*-acyl glycoses. Thus, searching for an efficient protocol for the preparation of complex, multifunctional glycosyl esters in the context of the total synthesis of phyllanthostatin antitumor agents, A. B. Smith and colleagues soundly investigated the suitability of the Mitsunobu reaction [29]. They concluded already back in 1986 that "the anomeric hydroxyl group of various pyranose hemiacetals can be esterified with inversion of configuration, conveniently, mildly and on large-scale using Ph_3P, with either DIAD or DEAD and a carboxylic acid in THF at either –50 °C or at room temperature". Hence, several protected mono- and disaccharides, such as **1**–**4** (Scheme 3) were selectively esterified with simple benzoic acid to give **5**–**7** and **9**, respectively. In addition, **4** was also converted with the phyllanthostatin aglycone **8** to give **10** with inversion of anomeric configuration. Extension of this work to other more complex antineoplastic glycosyl esters was successfully investigated by the same group [30-34].

De Mesmaeker et al. reported the stereoselective coupling of an allyl glucuronide, in which all hydroxy groups except the anomeric OH were *O*-acyl-protected, with carboxylic acids by a Mitsunobu reaction [35]. The reaction was successful even when a free phenolic function was present in the employed acid and the desired β-anomer of the 1-*O*-acyl-β-D-glucuronide products could be isolated in up to 50% yield. Similarly, regioselective esterification of unprotected allyl glucuronide **11** was



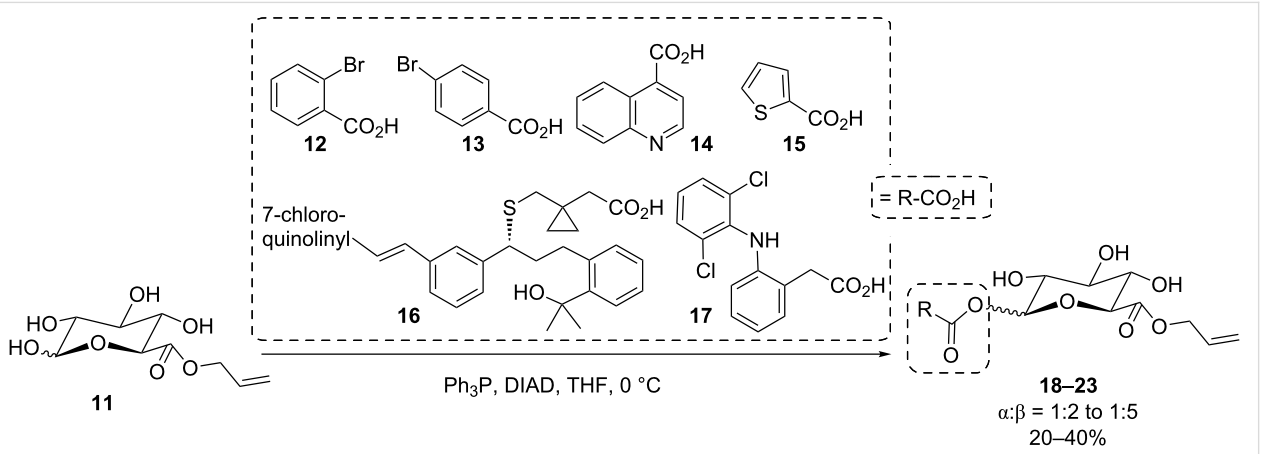
Scheme 3: Anomeric esterification using the Mitsunobu procedure [29].

performed by Juteau et al. with the acids **12**–**16** yielding anomeric mixtures of the respective 1-*O*-acyl-β-D-glucuronides **18**–**22** in quite acceptable yields even with complex acids like **16** (Scheme 4) [36]. The same approach was chosen in the Stachulski group for the anomeric modification of glucuronides with the anti-inflammatory drug diclofenac (**17**) to give the respective product **23** (Scheme 4) [37].

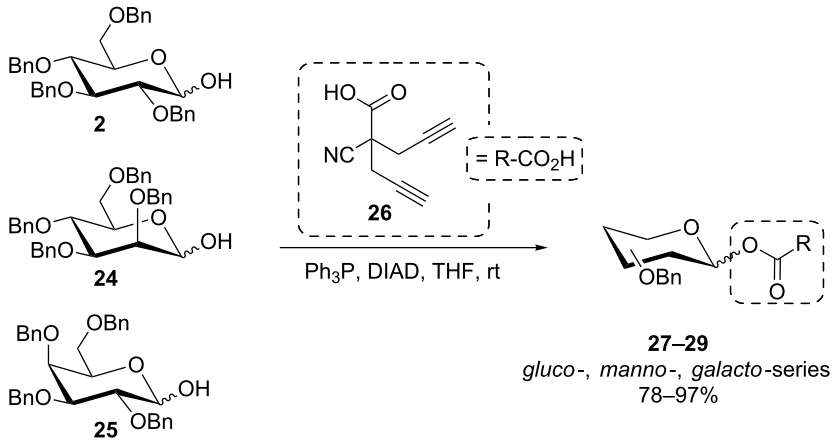
Bourhim et al. reported that the Mitsunobu reaction with native D-glucose, D-GlcNAc or D-maltose resulted in regioselective esterification of the primary OH group, leaving all other hydroxy groups including the anomeric OH unmodified [38]. On the other hand, other authors have reported that the anomeric position can be selectively modified in a Mitsunobu reaction without concomitant modification of the primary 6-OH (vide infra). Apparently, fine-tuning of reaction conditions can alter the selectivity of the Mitsunobu reaction and in addition, different regioselectivities might origin in the structure of the sugar substrate.

In the course of a synthesis of carbocyclic lignan variants related to podophyllotoxin, a pseudo-anomeric stereospecific inversion of a carbasugar was achieved in good yield in Nishimura's group [39]. More recently, the Mitsunobu procedure was applied in the context of gold-catalyzed glycosylation in order to install a reactive anomeric ester function in a series of *O*-benzylated glycoses (**2**, **24**, **25**) employing the branched carboxylic acid **26** (Scheme 5) [40]. The produced esters **27**–**29** were obtained as anomeric mixtures.

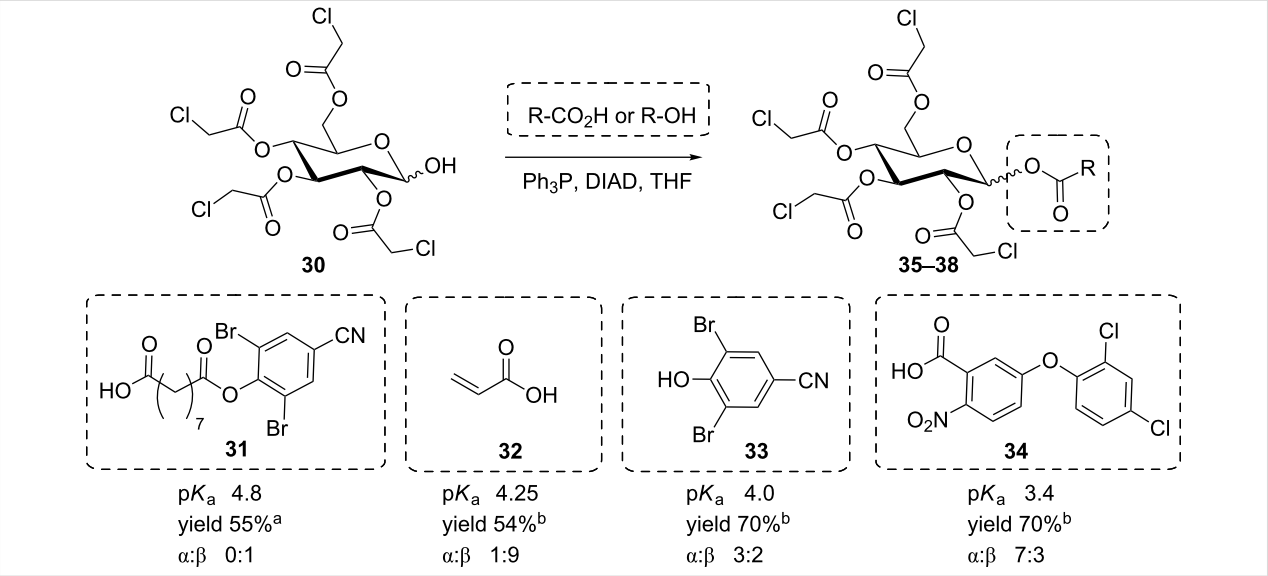
Lubineau et al. [27] investigated the stereoselectivity of the anomeric Mitsunobu coupling of 2,3,4,6-tetra-*O*-chloroacetyl-D-glucose (**30**) as well as its *galacto*-configured analogue with the carboxylic acids **31**–**34** to obtain products **35**–**38** which are related to various pesticide agents (Scheme 6). Their results supported the theory that, along with an effect of the reaction temperature, an increase of the pK_a of the employed acidic reaction partner can lead to predominant formation of the β-configured product, whereas stronger acidic reagents can favor the



Scheme 4: Conversion of allyl glucuronate into various 1-O-esterified allyl glucuronates using anomeric Mitsunobu esterification [36,37].



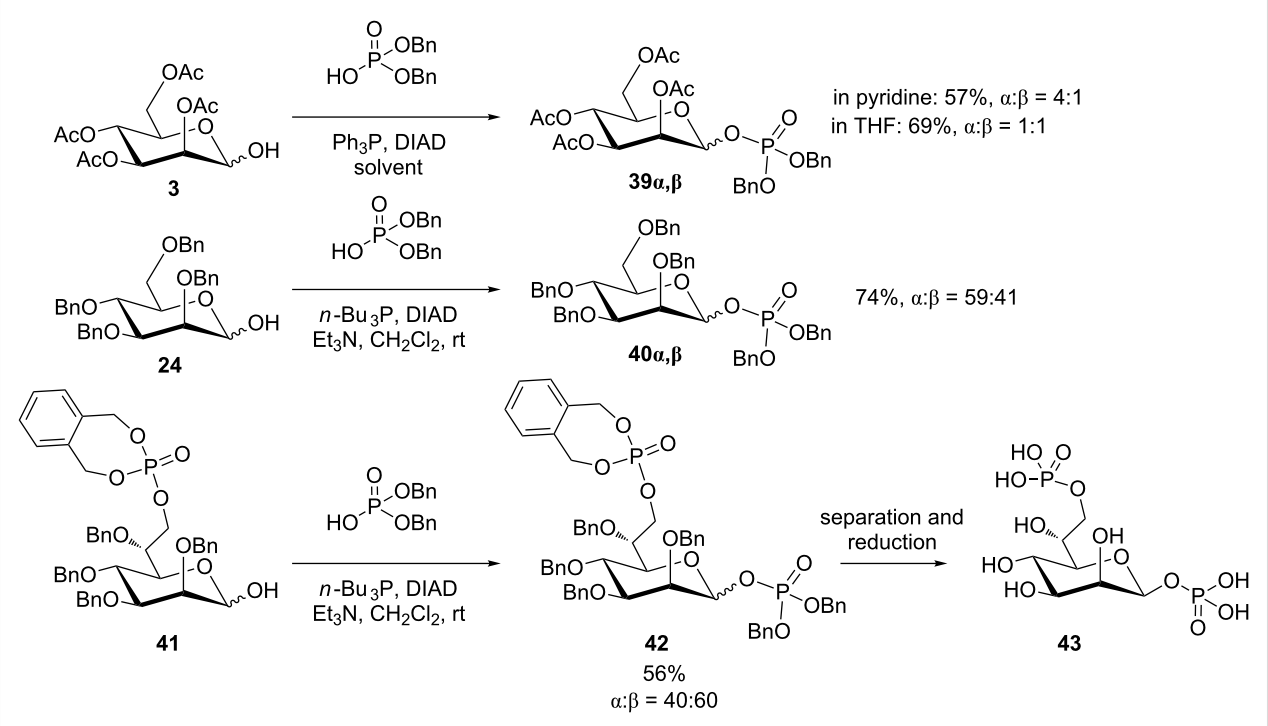
Scheme 5: Synthesis of anomeric glycosyl esters as substrates for Au-catalyzed glycosylation [40].



Scheme 6: Correlation between pK_a value of the employed acids (or alcohol) and the favoured anomeric configuration of the respective product. [a]Carried out at 0 °C; [b]carried out at rt [27].

formation of the respective α-anomers. These findings can be explained by considering the two different reaction pathways A and B as shown above in Scheme 2. The authors state that the observed pK_a effect is either due to the influence of the acidity of the employed acid on the reaction mechanism or results from the proton-catalyzed change of the anomeric ratio of the starting material **30** in solution.

Very recently, anomeric phosphorylation via a Mitsunobu approach was concomitantly undertaken by groups from Japan and Austria, respectively [41,42], aiming at the synthesis of the bacterial metabolite and potent innate immune modulator D-*glycero*-β-D-*manno*-heptose-1,7-bisphosphate (**43**, HBP, Scheme 7). The group around Zamyatina employed 2,3,4,6-tetra-*O*-acetyl-mannopyranose (**3**) as a 9:1 α,β-mixture in order



Scheme 7: Synthesis of the β-mannosyl phosphates for the synthesis of HBP **43** by anomeric phosphorylation according to Mitsunobu [41,42].

to optimize the reaction conditions for the Mitsunobu reaction with phosphoric acid dibenzyl ester. The anomeric mannosyl phosphate derivatives **39α** and **39β** were obtained in 57% total yield when pyridine was used as the solvent, as depicted in Scheme 7. In THF, the same reaction furnished a 1:1-anomeric mixture in 69% yield. The authors thus considered the Mitsunobu reaction as unsatisfactory for the synthesis of HBP. On the other hand, Inuki et al. optimized the Mitsunobu conditions with 2,3,4,6-tetra-*O*-benzyl-mannopyranose (**24**) and found that the addition of trimethylamine in dichloromethane improved the Mitsunobu process, leading to **40α** and **40β** in more than 70% yield. When such optimized conditions were applied to the mannose-6-phosphate derivative **41**, the desired bisphosphate **42** was obtained in 56% yield as a 40:60 α,β-anomeric mixture before work-up, and in a 53:47 ratio after work-up due to slight anomerization. As **42** can be easily converted into the target molecule, the authors concluded, that in spite of the poor stereoselectivity, the Mitsunobu reaction constitutes a key step in a successful access to β-mannosyl phosphates such as **43**.

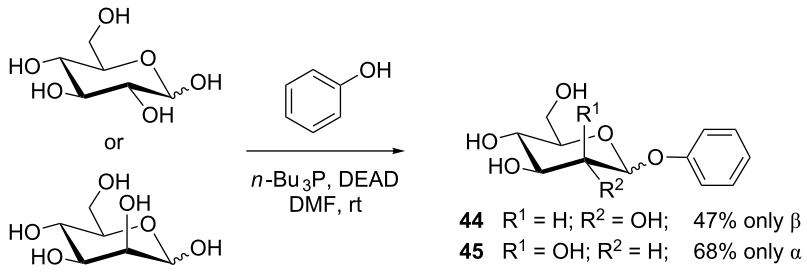
Reactions with phenols to achieve aryl glycosides

Not only anomeric esters, but also glycosides can be obtained through the Mitsunobu reaction. Dehydrative glycosylation approaches with reducing sugars were previously reviewed [43,44]. As phenols are weak acids, they are suitable reaction partners in the Mitsunobu reaction, leading to aryl glycosides with reducing sugars as the alcohol components. Grynkiewicz can be called the pioneer of Mitsunobu glycosylation, as having explored the Mitsunobu reaction for the synthesis of various aryl glycosides [24,45]. Thus, native sugars such as D-glucose and D-mannose (Scheme 8) were converted into the respective unprotected phenyl glycosides **44** and **45** with phenol in just one step in moderate to good yields.

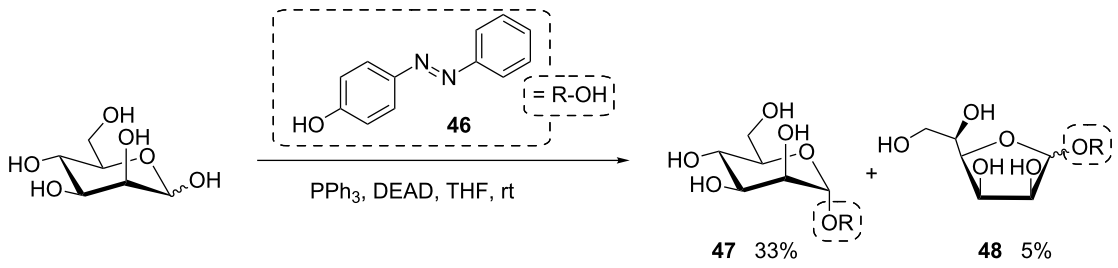
Recently, the scope of this synthetic approach was expanded by the Lindhorst group employing D-mannose and hydroxyazobenzene **46** for the synthesis of the photoswitchable azobenzene α-D-mannoside **47** (Scheme 9) [46]. Notably, in this reaction, traces of an anomeric mixture of the respective furanoside **48** were detected.

The Mitsunobu synthesis of aryl glycosides was also applied to *p*-nitrophenol [47], naphthols [48,49], or multifunctional phenols [27,50]. Such arylglycosylation was also extended for the synthesis of aureolic acid antibiotics [21,51,52]. In search of convenient methods for the synthesis of aryl sialosides, Gao et al. explored the scope of the Mitsunobu reaction with the sialic acid derivative **49**, employing a range of phenols **50**–**58** in acetonitrile to achieve sialosides **59**–**67**, albeit with modest anomeric selectivity (Scheme 10) [25].

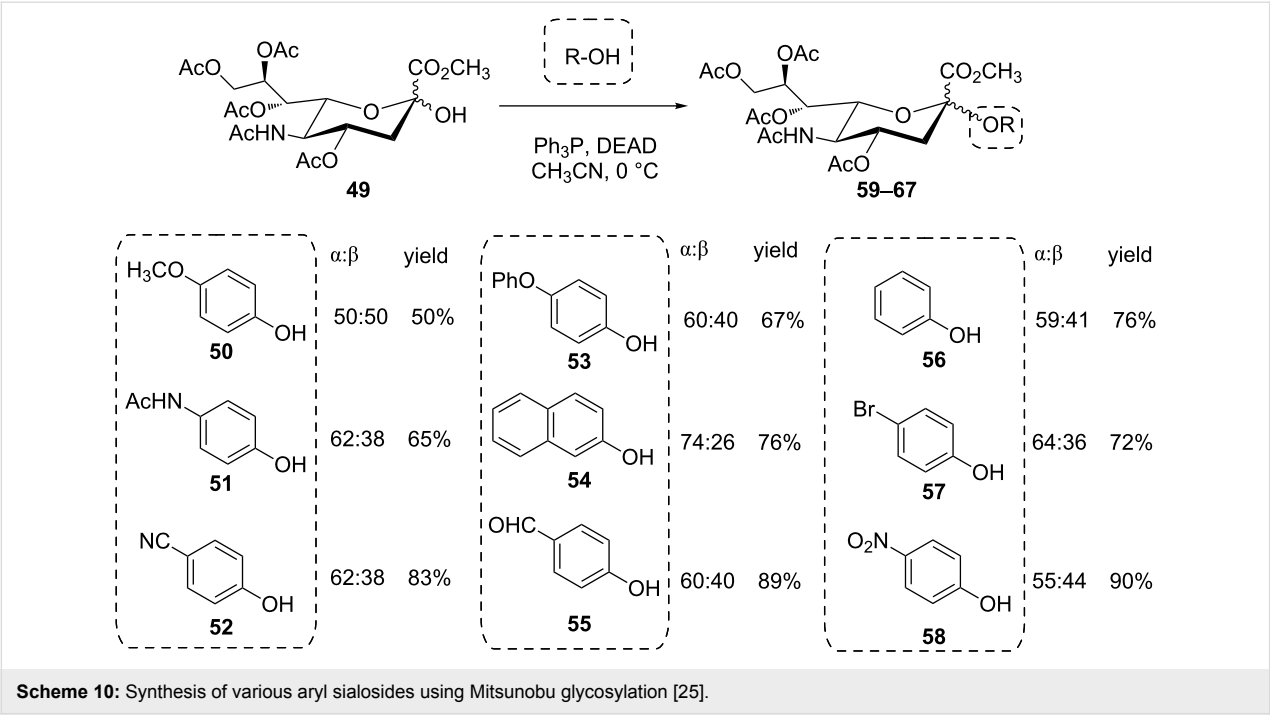
Interestingly, no correlation between the pK_a of the employed acids and the stereoselectivity of the reaction could be established in this case, since similar anomeric mixtures were obtained throughout all experiments. In this case, the absence of a neighboring group in position 3 of the sugar ring could account



Scheme 8: Synthesis of phenyl glycosides **44** and **45** from unprotected sugars [24].



Scheme 9: Synthesis of azobenzene mannosides **47** and **48** without protecting group chemistry [46].



Scheme 10: Synthesis of various aryl sialosides using Mitsunobu glycosylation [25].

for low stereoselectivity. To explain the lack of stereoselectivity, the authors considered a S_N1 reaction mechanism, involving the respective oxocarbenium ion or, alternatively, the formation of both α- and β-configured glycosyloxyphosphonium ions, which are in turn displaced by the nucleophile in the expected S_N2 fashion, resulting in a respective anomeric mixture of products (cf. Scheme 2).

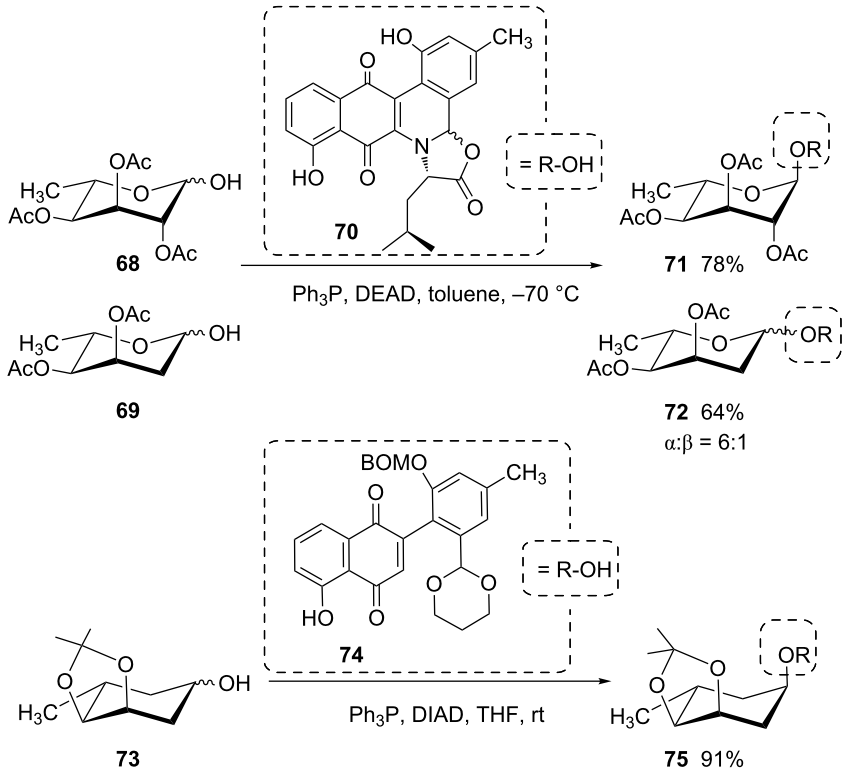
In contrast to this, the yields of the obtained aryl sialosides strongly correlated with the pK_a of the utilized phenols, with stronger acids leading to higher yields. This yield-to-pK_a correlation is in accordance with earlier findings in the synthesis of aryl glucuronides where the yields were equally affected by the pK_a of the chosen phenols, while the neighboring group effect was found to govern the stereochemical outcome of the reaction towards β-configured products [53]. In this case, phenolic chromium tricarbonyl complexes of weaker acids such as *p*-cresol were employed to improve the yield.

The challenge of glycoside synthesis using sugars devoid of a C-2 participating group is also highlighted by a total synthesis of various jadomycins [54]. Whereas the Mitsunobu glycosidation of **68** with the phenolic aglycon **70** yields the pure 1,2-*trans*-glycoside **71**, the 2-deoxy sugar **69** yields the glycoside **72** as a 6:1 α,β-anomeric mixture (Scheme 11). In contrast to this, the jadomycin B carbasugar analogue **75** was formed stereoselectively from the 2-deoxy-carbasugar **73** in a Mitsunobu reaction with the aglycon **74** [55].

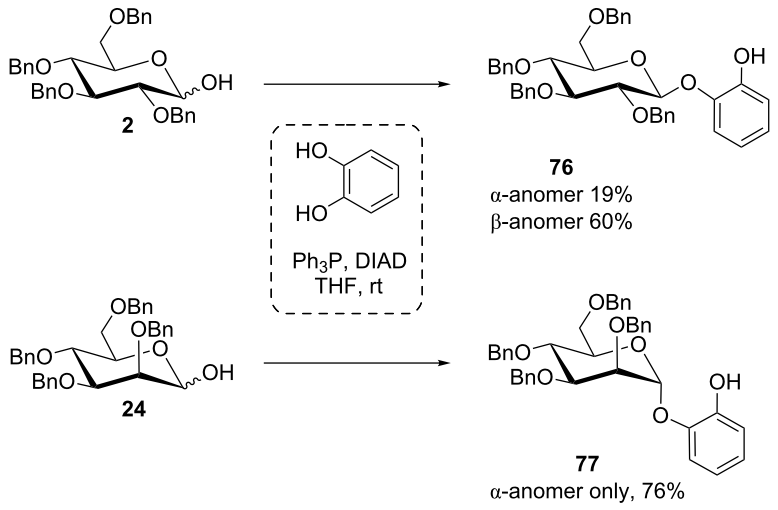
Benzyl protection, which does not exert neighboring group effects in classical glycosylations, resulted in the predominant formation of 1,2-*trans* glycosides in the Mitsunobu reaction with catechol. In fact, benzyl-protected reducing glucose derivative **2** gave the β-glucoside **76** with good stereoselectivity, and the respective mannose derivative **24** resulted in the pure α-mannoside **77** in good yield (Scheme 12) [56].

In a general approach to coumarin-derived inhibitors of gyrase B, a group working at Hoechst Marion Roussel developed a Mitsunobu process to connect noviose with a broad range of 7-hydroxycoumarins [57]. Similarly, Imamura and colleagues used 4-methylumbelliferone (**79**) as acidic reaction partner in a Mitsunobu glycosylation with a reducing galabioside **78** (Scheme 13) [22]. Advantage was taken of the bulky DTBS protecting group to enforce α-stereoselection despite of the anchimeric effect of the vicinal *N*-Troc protecting group to achieve the α-glycoside **80** in high yield. Nevertheless, this reaction needed optimization, such as an unusually high reaction temperature.

Also weakly acidic phenols were used by Vaccaro et al. [58] for Mitsunobu glycosylation in the D-glucuronic series, employing the reagent pair *n*-Bu_3P-ADDP (1,1'-(azodicarbonyl)dipiperidine) developed by Tsunoda et al. [59]. Interestingly, Davis and co-workers could employ 2,3:4,6-di-*O*-isopropylidene mannopyranose **81** in a Mitsunobu reaction with phenol to stereoselectively achieve the respective β-mannoside **82** in good yield (Scheme 14) [60].



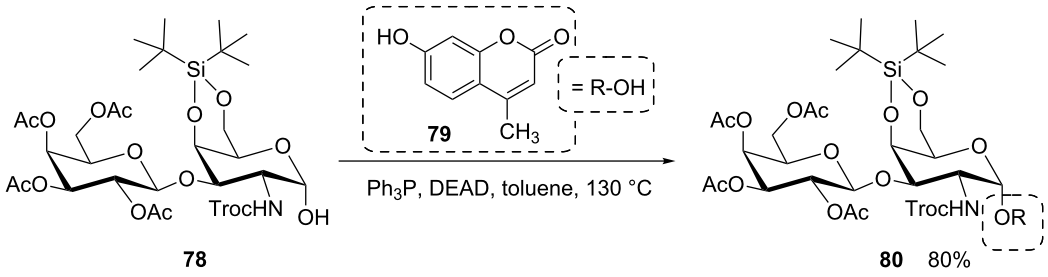
Scheme 11: Mitsunobu synthesis of different jadomycins [54,55]. BOM: benzyloxymethyl.



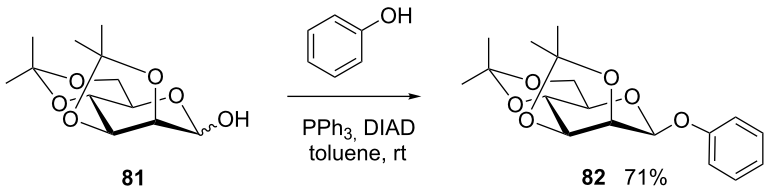
Scheme 12: Stereoselectivity in the Mitsunobu synthesis of catechol glycosides in the *gluco*- and *manno*-series [56].

This stereo-differentiating effect of isopropylidene protecting groups was also observed in other cases with D-mannopyranose [46,61]. It might be used as a key to a reliable approach to otherwise difficult to synthesize β-mannosides using the Mitsunobu procedure. This approach to 1,2-*cis*-mannosides is equally effective when cyclohexylidene protecting groups are used [28,47,62].

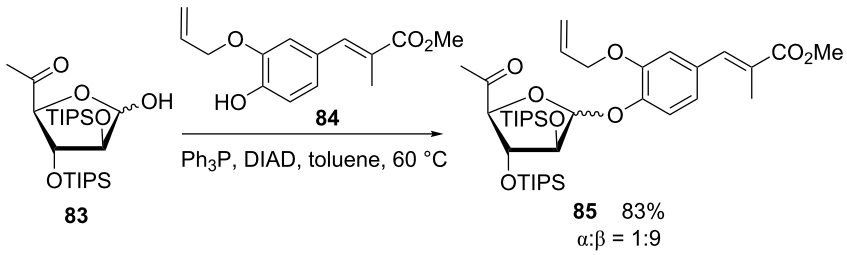
Mitsunobu glycosylation was also a successful method in total synthesis. In the course of a 17-step synthesis of hygromycin A, Donohoe et al. used a Mitsunobu glycosylation of **84** with the arabinose derivative **83**. This reaction could be tuned to deliver the required β-arabinofuranoside building block **85** with high stereoselectivity and under the assistance of triisopropylsilyl (TIPS) protecting groups (Scheme 15) [63].



Scheme 13: Formation of a 1,2-*cis* glycoside **80** assisted by steric hindrance of the β-face of the disaccharide through the DTBS protection. DTBS: di-*tert*-butylsilylene; Troc: 2,2,2-trichloroethoxycarbonyl [22].



Scheme 14: Stereoselective β-D-mannoside synthesis [60].



Scheme 15: TIPS-assisted synthesis of 1,2-*cis* arabinofuranosides [63]. TIPS: triisopropylsilyl.

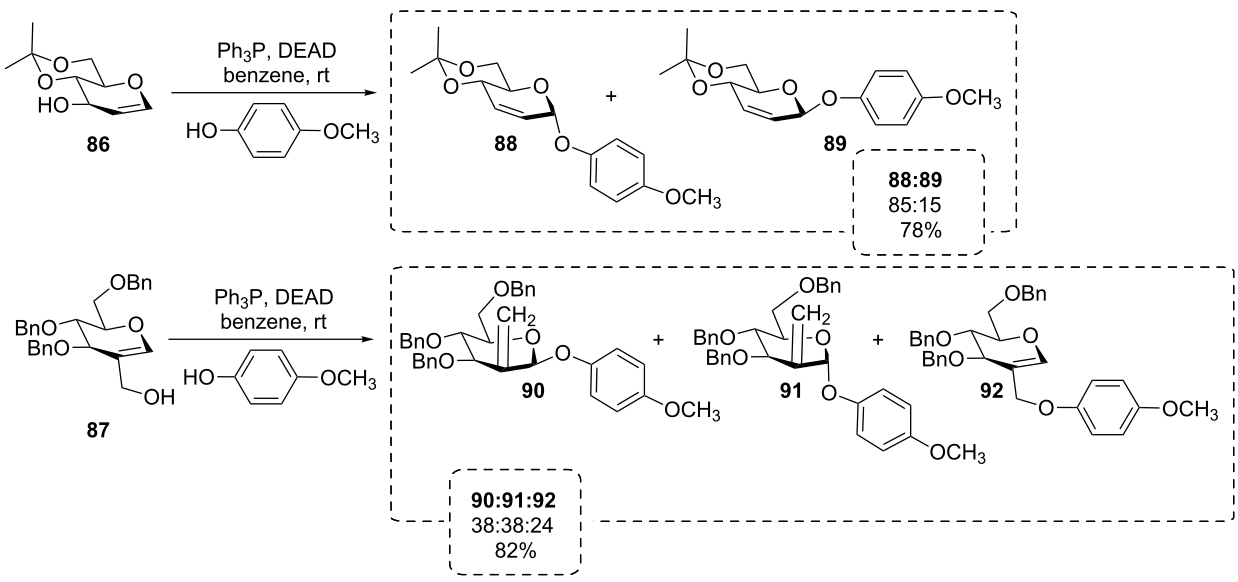
Similar reaction conditions were applied by Nie et al. in the total synthesis of the nucleoside antibiotic A201A [64]. Notably, the *n*-Bu_3P-ADDP reagent system led here to the formation of the pure α-glycoside. Likewise, a Mitsunobu glycosylation of complex phenols was successfully implemented in the preparation of novobiocin analogues [65], and formed a key step in the synthesis of new glycosidic PDE4 (phosphodiesterase type 4) inhibitors [66]. Also calix[4]arenes could be selectively mono- or diglycosylated by means of the Mitsunobu methodology [67,68].

The Mitsunobu reaction was also employed with glycals like **86** and **87** reacting with *p*-methoxyphenol as an alternative to the Ferrier rearrangement in the synthesis of 2-*C*-methylene glycosides and other rearrangement products **88**–**92**, some of which cannot be obtained in a classical Ferrier reaction (Scheme 16) [69-72]. The results outlined in Scheme 16 are consistent with early findings of Guthrie et al. exploring the Mitsunobu benzoylation of 4,6-*O*-benzylidene-D-allal [73].

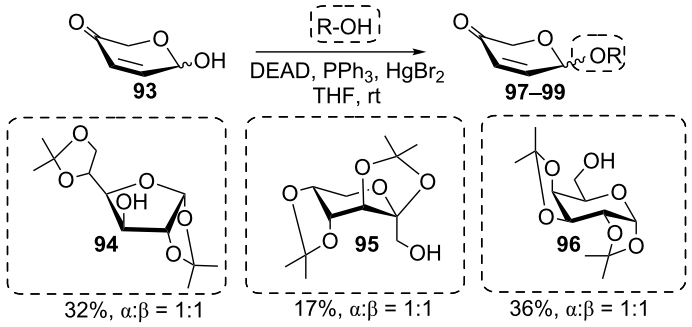
Reactions with alcohols to yield alkyl glycosides

In contrast to aryl ethers, the formation of alkyl ethers is not observed under Mitsunobu conditions. Likewise, standard alcohols are typically poor reaction partners in Mitsunobu glycosylations. Due to their high pK_a values, the formation of the transient phosphonium betaine is hampered [43]. In an effort to overcome this drawback, several decades ago, Szarek et al. tested mercuric halides to assist the betaine formation in such cases, and indeed cyclohexyl glycosides could be formed in various sugar series with decent yields [74]. Consequently, this approach was explored in a Mitsunobu-type disaccharide synthesis reacting **93** with the alcohol components **94**–**96** to give **97**–**99**, albeit with moderate success (Scheme 17) [75].

Contradictory results were reported on the Mitsunobu glycosylation of 1,3,4,6-tetra-*O*-protected fructofuranosides. In contrast to Guthrie et al. [76], Bouali and colleagues claimed an effective synthesis of alkyl fructofuranoside **101**–**103** from **100** using



Scheme 16: The Mitsunobu reaction with glycals leads to interesting rearrangement products [69].



Scheme 17: Synthesis of disaccharides using mercury(II) bromide as co-activator in the Mitsunobu reaction [75].

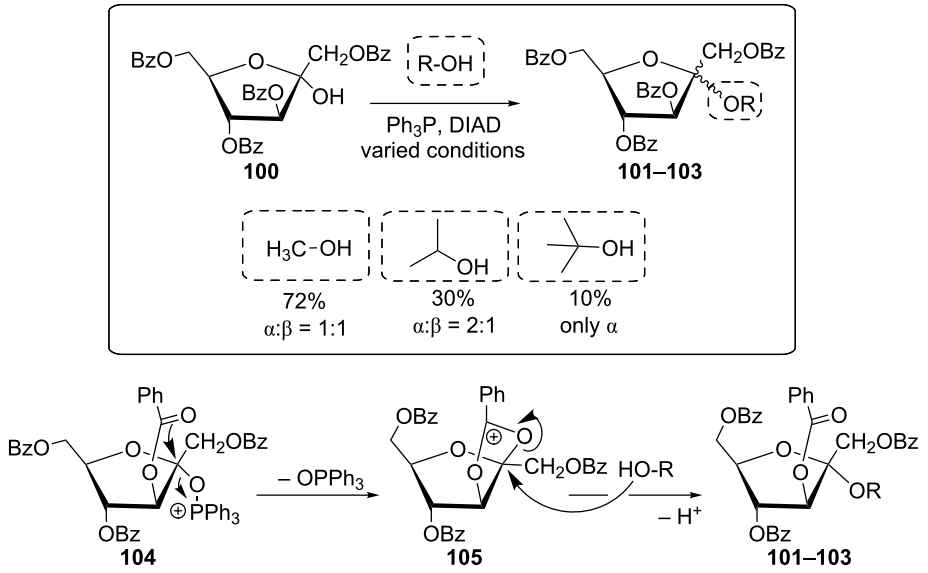
simple alcohols (Scheme 18) [23]. The reaction was rationalized by participation of the C-3 neighboring group (structure **104**) with intermediate formation of a dioxolanium derivative **105** [23].

On the other hand, the Mitsunobu reaction was advantageous for the acetalization of the antimalarial drug dihydroartemisinin **106** to give **107** with trifluoroethanol, having a pK_a of 12.4 (Scheme 19) [77]. The efficiency of the Mitsunobu glycosylation with fluorinated alcohols with pK_a values between 9 and 12 was demonstrated with several other examples [78].

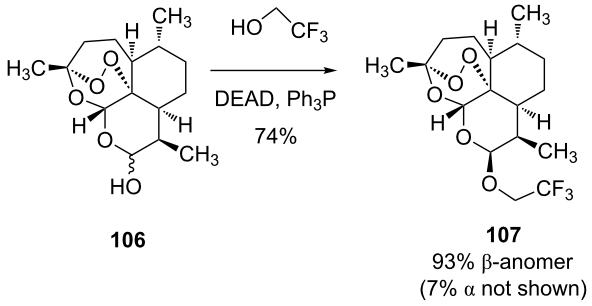
Also thiols, according to their pK_a value range between 10 and 11 should be qualified appropriate reagents for a Mitsunobu thioglycosylation. However, a competitive redox reaction with the PR_3-azodicarboxylate reagent system precludes this application [79,80]. In spite of that, thioglycosides **111**–**113** could be prepared via a Mitsunobu-type condensation of thioglycosides such as **108** and **109** with simple alcohols (Scheme 20) [81,82]. In this case, of course, the sugar thioglycoside takes the role of the nucleophile rather than of the alcohol component in the Mitsunobu reaction.

Reactions with NH acids to achieve N-glycosides

Early on, phthalimide was regarded as a good Mitsunobu reagent, owing to its NH acidity with a pK_a of 8.3, thus offering the opportunity for the synthesis of N-glycosides of the *N*-glycosylimide type [83]. However, along with the formation of *N*-glycosylphthalimides, a side-reaction takes place, producing both glycosyl carbonates and *N*-glycosyl-1,2-dialkoxycarbonylhydrazines [84]. This anomeric *N*-phthalimidation was



Scheme 18: Synthesis of various fructofuranosides according to Mitsunobu and proposed neighbouring group participation [23].

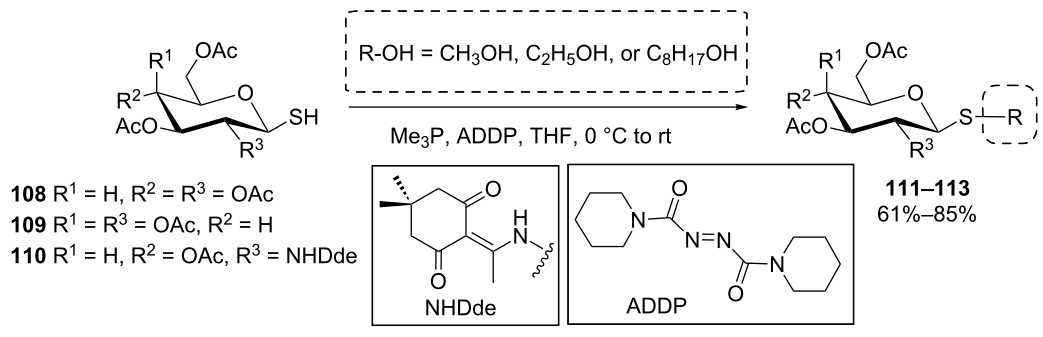


Scheme 19: The Mitsunobu reaction allows stereoslective acetalization of dihydroartemisinin [77].

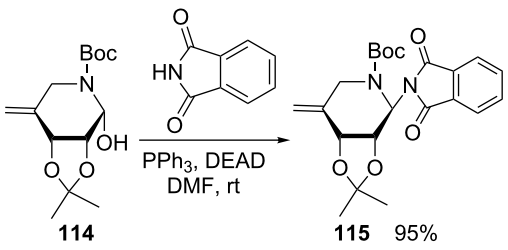
later implemented by Nishimura et al. for the iminosugar **114** with phthalimide to give **115** in a high yield, en route to a new family of α-L-fucosidase inhibitors (Scheme 21) [85].

More generally, the preparation of modified glycosylamines under Mitsunobu conditions requires a sufficiently acidic NH nucleophile. A particularly illustrative procedure was disclosed by van Boom's group, who used *N*-nosyl-activated amino-acid esters for anomeric modification of sugars in order to produce substrates for a novel route to Amadori rearrangement products [86]. The same approach was recently adopted in a total synthesis of aurantoside G, involving the Mitsunobu ligation of a D-xylopyranose derivative **116** and *N*-nosylated methyl asparaginate **117** to give **118** (Scheme 22) [87].

Compared to *N*-sulfonylation, *N*-carbamoylation can also prove effective to enhance the acidity of a NH group. Hence, the trichloroethoxycarbonyl (Troc) protection/activation of the amino group of questiomycin **119** allowed Igarashi et al. to access the *N*-glucosylated derivative **120** in good yield and complete β-stereoselectivity from hemiacetal **1** (Scheme 23) [88].



Scheme 20: Synthesis of alkyl thioglycosides by Mitsunobu reaction [81].

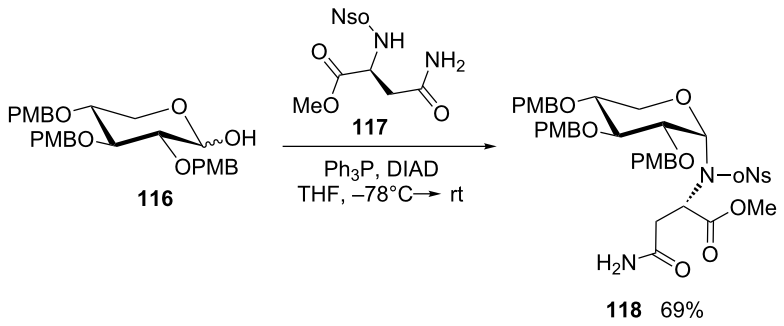


Scheme 21: Preparation of iminoglycosylphthalimide **115** from **114** [85].

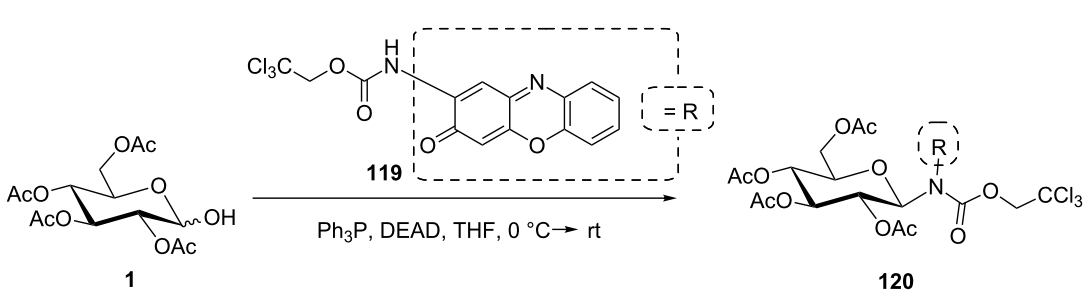
Also some aza-heterocycles bearing a free NH group possess a low enough pK_a to allow Mitsunobu coupling. In the course of the synthesis of the hexasaccharidic fragment of landomycin A, the L-rhodinose derivative **121** underwent glycosylation with 1*H*-tetrazole to give **122**, which has a pK_a that compares to carboxylic acids (Scheme 24) [89].

In spite of the fact that parent indole is too weak an acid to undergo Mitsunobu conversions, a model maleimide–indole hybrid was investigated by Ohkubo and colleagues to pave the way for the synthesis of indolo[2,3-*a*]pyrrolo[3,4-*c*]carbazole compounds with anticancer activity [90,91]. N-Glycosides of indole derivatives were also approached by Zembower et al. employing 2,3,4,6-tetra-*O*-benzyl glucopyranose in a Mitsunobu reaction [92]. In the same period, Prudhomme's group followed closely related approaches for the *N*-glycosylation of indolic structures. Various rebeccamycin analogues were efficiently synthesized from indolo[2,3-*c*]carbazole frameworks using the methodology previously developed by Voldoire et al. [93]. Further applications to 7-aza-indolic analogues of rebeccamycin [94-96], granulatimide and isogranulatimide [97-100] were also reported. In addition, using the same Mitsunobu methodology, the rebeccamycin analogue **124** was synthesized in high yield and complete β-stereoselectivity by Wang et al. from the glucose derivative **2** and **123** (Scheme 25) [101].

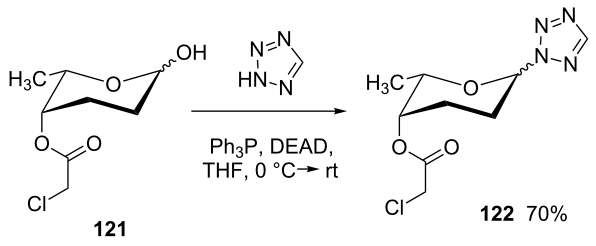
Application of the anomeric Mitsunobu coupling in nucleoside synthesis was pioneered by Szarek et al. [102], who reacted 6-chloropurine with various reducing sugars using methyldiphenylphosphine as activator. Extension to the D-*ribo* series with 6-chloro- and 2,6-dichloropurines was later reported by Hertel and co-workers [103]. In the course of an exploration of modified L-nucleosides, 6-chloropurin-9-yl derivatives were obtained in moderate yields [104]. Aiming at an improved procedure to synthesize nucleosides with glycosylation of the nucleobase, De Napoli et al. used the Bu_3P-ADDP system to connect inosine and uridine derivatives with D-ribofurano and D-glucopyrano moieties [105]. Hocek's group in 2015 published a direct one-pot synthesis of exclusively β-configured nucleosides from unprotected or 5-*O*-monoprotected



Scheme 22: Mitsunobu reaction as a key step in the total synthesis of aurantoside G [87].



Scheme 23: Utilization of an N–H acid in the Mitsunobu reaction [88].



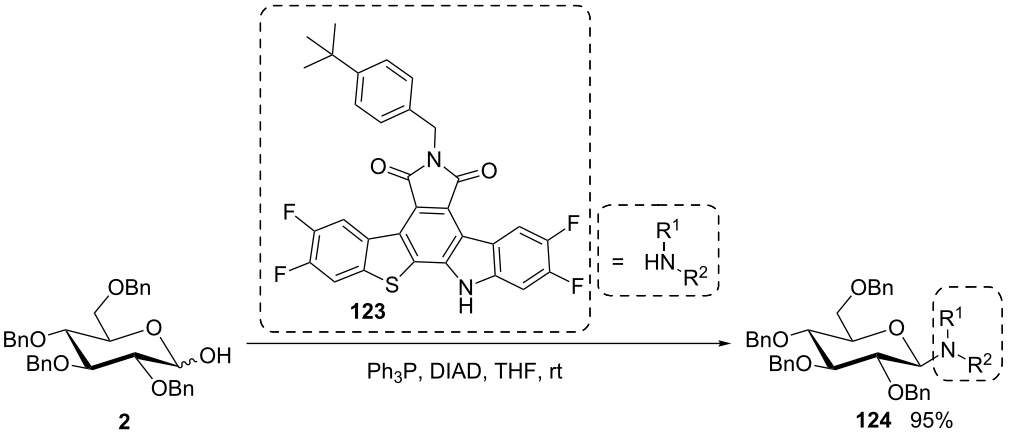
Scheme 24: Mitsunobu reaction with 1*H*-tetrazole [89].

D-ribose using optimized Mitsunobu conditions with various purine- and pyrimidine-based heterocycles. Here, DBU was applied first, followed by DIAD and P(*n*-Bu)$_3$ [106]. Two years later Seio and colleagues set out to systematically study the effect of phosphine, azodicarbonyl reagent, and solvent on the yield and α/β ratio in the synthesis of 2'-deoxynucleosides [107]. They reported that the highest yield and β-selectivity were obtained using (*n*-Bu)$_3$P and 1,1′-(azodicarbonyl)dipiperidine in DMF. In a model study directed towards the synthesis of guanofosfocin, Sugimura et al. used the Mitsunobu *N*-glycosylation to attach a glucopyranosyl donor on either 6-*N*-trityl-8-oxoadenosine or 6-*O*-benzyl-8-oxoinosine [108].

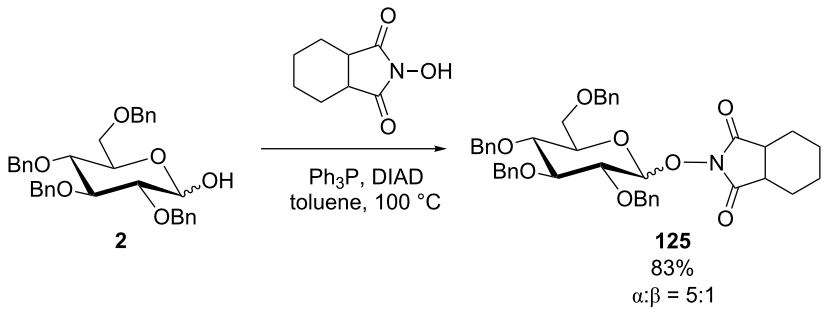
Reactions with N–OH acids to yield NO-glycosides

Because of its well-suited pK_a (6.3), *N*-hydroxyphthalimide was early considered in Mitsunobu reactions, for example by Grochowski and Jurczak to form an anomeric phthalimide–oxy bond as shown in several sugar series [109-111]. This gives access to new *O*-glycosylhydroxylamines, namely for the construction of glycosidic N–O linkages in calicheamycin oligosaccharides [112,113]. This option was applied in the synthesis of trichostatin D involving glucose derivative **2** and *N*-hydroxyhexahydrophthalimide as the glycosyl acceptor to give **125** (Scheme 26) [114].

By using diverse *N*-hydroxylated azaheterocycles in the Mitsunobu glycosylation, Grochowski explored the synthesis of new nucleoside analogues. 1-Hydroxy-benzotriazole, 1-hydroxy-2-cyanobenzimidazole, 1-hydroxyuracil, and



Scheme 25: Formation of a rebeccamycin analogue using the Mitsunobu reaction [101].



Scheme 26: Synthesis of carbohydrates with an alkoxyamine bond [114].

1-hydroxythymine were used to prepare the respective NO-furanosides in the *manno*- and *ribo*-series [115-117].

Miscellaneous

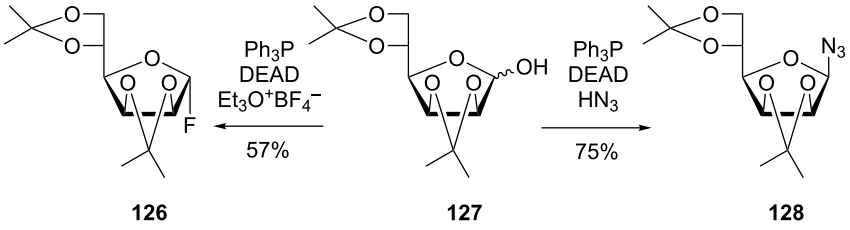
The Mitsunobu reaction was also applied for other anomeric modifications, such as fluorination, reported by Kunz et al. for the synthesis of the α-D-mannofuranosyl fluoride **126**, however, in moderate yield (Scheme 27) [118]. The advantage of this approach lies in the mild fluorine source, triethyloxonium tetrafluoroborate, which, in combination with the Ph_3P-DEAD system, leaves the acid-labile protecting groups of **127** intact, other than when HF is used. Zbiral's group on the other hand, developed the synthesis of glycosyl azides such as **128** in a Mitsunobu procedure with **127**, using hydrazoic acid as the azide source (Scheme 27) [119].

This approach was extended by Besset et al. to D-fructose and a range of unprotected mono- and disaccharides, again showing a preference of the reaction for the anomeric position instead of the primary [120]. Anomeric azidation was also investigated on diverse unprotected hexopyranoses by Larabi et al. using a modified Appel-type procedure [121].

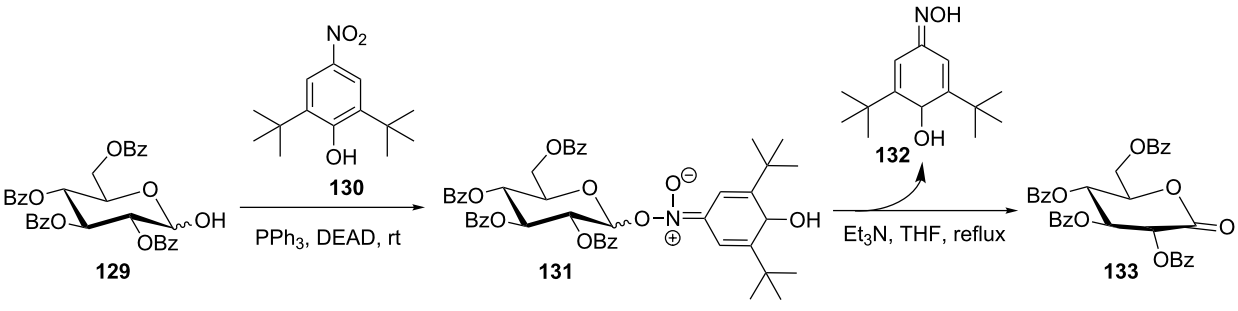
A striking oxidation reaction of alcohols to carbonyl compounds was disclosed by Mitsunobu and colleagues, involving the sterically hindered nitrophenol **130** [122]. With sugars like **129**, the Mitsunobu glycosylation is hampered, and instead an anomeric aci-nitroester **131** is formed, which is converted into the corresponding gluconolactone **133** under elimination of a quinone monoxime **132** (Scheme 28).

Conclusion

In this account, 15 years after Professor Mitsunobu has passed away, we have surveyed the literature on the Mitsunobu reaction for anomeric modifications of carbohydrates. As in classical glycosylation reactions, not all mechanistic details of the anomeric conversion of sugars in a Mitsunobu process are known and well understood. Hence until today, surprising results and unexpected side reactions are being observed in Mitsunobu type conversions of hemiacetals. In addition, the reaction conditions of a Mitsunobu process often require particular optimization efforts. Thus, the Mitsunobu reaction has not become a standard procedure in glycoside synthesis nor in anomeric esterification, but on the other hand, it was demonstrated to serve as a key step in many cases of carbohydrate modification including total synthesis of sensitive natural products. This is also due to the mild and neutral conditions under which the Mitsunobu reaction occurs. Additionally, it has a rather broad scope as many building blocks are acidic enough to react with reducing sugars representing the alcohol component of the reaction. The stereochemical outcome of a Mitsunobu glycosylation is often advantageous such as in the synthesis of β-D-mannosides, which are otherwise difficult to prepare. However, often, the stereoselectivity of the reaction is less definite than our text books claim. Unfortunately, the Mitsunobu reaction is uneasy to scale up and this is probably one of the biggest



Scheme 27: Synthesis of glycosyl fluorides and glycosyl azides according to Mitsunobu [118,119].



Scheme 28: Anomeric oxidation under Mitsunobu conditions [122].

obstacles for a broad and also technical use of this reaction. Nevertheless, this review proves that in the glycosciences, the Mitsunobu reaction must not be overlooked as it is an important method in the synthetic toolbox for anomeric modification of sugars and glycoconjugate preparation.

ORCID® iDs

Julia Hain - https://orcid.org/0000-0001-6701-283X
Patrick Rollin - https://orcid.org/0000-0001-5951-9614
Thisbe K. Lindhorst - https://orcid.org/0000-0001-6788-4224

References

1. Mitsunobu, O.; Yamada, M.; Mukaiyama, T. *Bull. Chem. Soc. Jpn.* **1967,** *40,* 935–939. doi:10.1246/bcsj.40.935
2. Mitsunobu, O. *Synthesis* **1981,** 1–28. doi:10.1055/s-1981-29317
3. Hughes, D. L. *Org. React.* **1992,** *42,* 335–656. doi:10.1002/0471264180.or042.02
4. Hughes, D. L. *Org. Prep. Proced. Int.* **1996,** *28,* 127–164. doi:10.1080/00304949609356516
5. But, T. Y. S.; Toy, P. H. *Chem. – Asian J.* **2007,** *2,* 1340–1355. doi:10.1002/asia.200700182
6. Swamy, K. C. K.; Kumar, N. N. B.; Balaraman, E.; Kumar, K. V. P. P. *Chem. Rev.* **2009,** *109,* 2551–2651. doi:10.1021/cr800278z
7. Varasi, M.; Walker, K. A. M.; Maddox, M. L. *J. Org. Chem.* **1987,** *52,* 4235–4238. doi:10.1021/jo00228a016
8. Hughes, D. L.; Reamer, R. A.; Bergan, J. J.; Grabowski, E. J. J. *J. Am. Chem. Soc.* **1988,** *110,* 6487–6491. doi:10.1021/ja00227a032
9. Camp, D.; Jenkins, I. D. *J. Org. Chem.* **1989,** *54,* 3045–3054. doi:10.1021/jo00274a016
10. Elson, K. E.; Jenkins, I. D.; Loughlin, W. A. *Org. Biomol. Chem.* **2003,** *1,* 2958–2965. doi:10.1039/B305375J
11. Schenk, S.; Weston, J.; Anders, E. *J. Am. Chem. Soc.* **2005,** *127,* 12566–12576. doi:10.1021/ja052362i
12. Camp, D.; von Itzstein, M.; Jenkins, I. D. *Tetrahedron* **2015,** *71,* 4946–4948. doi:10.1016/j.tet.2015.05.099
13. Hughes, D. L.; Reamer, R. A. *J. Org. Chem.* **1996,** *61,* 2967–2971. doi:10.1021/jo952180e
14. Lindhorst, T. K.; Thiem, J. *Liebigs Ann. Chem.* **1990,** 1237–1242. doi:10.1002/jlac.1990199001222
15. Tsunoda, T.; Otsuka, J.; Yamamiya, Y.; Itô, S. *Chem. Lett.* **1994,** *23,* 539–542. doi:10.1246/cl.1994.539
16. Itô, S.; Tsunoda, T. *Pure Appl. Chem.* **1999,** *71,* 1053–1057. doi:10.1351/pac199971061053
17. Guinó, M.; Hii, K. K. M. *Chem. Soc. Rev.* **2007,** *36,* 608–617. doi:10.1039/B603851B
18. Ma, X.; Shi, R.; Zhang, B.; Yan, B. *J. Comb. Chem.* **2009,** *11,* 438–445. doi:10.1021/cc900004m
19. Camp, D.; Harvey, P. J.; Jenkins, I. D. *Tetrahedron* **2015,** *71,* 3932–3938. doi:10.1016/j.tet.2015.04.035
20. Dar, A. R.; Aga, M. A.; Kumar, B.; Yousuf, S. K.; Taneja, S. C. *Org. Biomol. Chem.* **2013,** *11,* 6195–6207. doi:10.1039/c3ob40853a
21. Roush, W. R.; Lin, X.-F. *J. Am. Chem. Soc.* **1995,** *117,* 2236–2250. doi:10.1021/ja00113a013
22. Imamura, A.; Ando, H.; Ishida, H.; Kiso, M. *Org. Lett.* **2005,** *7,* 4415–4418. doi:10.1021/ol051592z
23. Bouali, A.; Descotes, G.; Ewing, D. F.; Grouiller, A.; Lefkidou, J.; Lespinasse, A.-D.; Mackenzie, G. *J. Carbohydr. Chem.* **1992,** *11,* 159–169. doi:10.1080/07328309208017797
24. Grynkiewicz, G. *Pol. J. Chem.* **1979,** *53,* 1571–1579.
25. Gao, G.; Schwardt, O.; Ernst, B. *Carbohydr. Res.* **2004,** *339,* 2835–2840. doi:10.1016/j.carres.2004.10.003
26. Ahn, C.; Correia, R.; DeShong, P. *J. Org. Chem.* **2002,** *67,* 1751–1753. doi:10.1021/jo001590m
27. Lubineau, A.; Meyer, E.; Place, P. *Carbohydr. Res.* **1992,** *228,* 191–203. doi:10.1016/S0008-6215(00)90559-7
28. Åkerfeldt, K.; Garegg, P. J.; Iversen, T. *Acta Chem. Scand.* **1979,** *B33,* 467–468. doi:10.3891/acta.chem.scand.33b-0467
29. Smith, A. B., III; Hale, K. J.; Rivero, R. A. *Tetrahedron Lett.* **1986,** *27,* 5813–5816. doi:10.1016/S0040-4039(00)85334-4
30. Smith, A. B., III; Rivero, R. A. *J. Am. Chem. Soc.* **1987,** *109,* 1272–1274. doi:10.1021/ja00238a061
31. Smith, A. B., III; Hale, K. J.; Vaccaro, H. A. *Tetrahedron Lett.* **1987,** *28,* 5591–5594. doi:10.1016/S0040-4039(00)96788-1
32. Vaccaro, H. A.; Rivero, R. A.; Smith, A. B., III. *Tetrahedron Lett.* **1989,** *30,* 1465–1466. doi:10.1016/S0040-4039(00)99491-7
33. Smith, A. B., III; Rivero, R. A.; Hale, K. J.; Vaccaro, H. A. *J. Am. Chem. Soc.* **1991,** *113,* 2092–2112. doi:10.1021/ja00006a030
34. Smith, A. B., III; Hale, K. J.; Vaccaro, H. A.; Rivero, R. A. *J. Am. Chem. Soc.* **1991,** *113,* 2112–2122. doi:10.1021/ja00006a031
35. De Mesmaeker, A.; Hoffmann, P.; Ernst, B. *Tetrahedron Lett.* **1989,** *30,* 3773–3776. doi:10.1016/S0040-4039(01)80651-1
36. Juteau, H.; Gareau, Y.; Labelle, M. *Tetrahedron Lett.* **1997,** *38,* 1481–1484. doi:10.1016/S0040-4039(97)00135-4
37. Kenny, J. R.; Maggs, J. L.; Meng, X.; Sinnott, D.; Clarke, S. E.; Park, B. K.; Stachulski, A. V. *J. Med. Chem.* **2004,** *47,* 2816–2825. doi:10.1021/jm030891w
38. Bourhim, A.; Czernecki, S.; Krausz, P. *J. Carbohydr. Chem.* **1993,** *12,* 853–863. doi:10.1080/07328309308020100
39. Saito, H.; Nishimura, Y.; Kondo, S.; Takeuchi, T. *Chem. Lett.* **1988,** *17,* 1235–1238. doi:10.1246/cl.1988.1235
40. Rao Koppolu, S.; Niddana, R.; Balamurugan, R. *Org. Biomol. Chem.* **2015,** *13,* 5094–5097. doi:10.1039/C5OB00248F
41. Inuki, S.; Aiba, T.; Kawakami, S.; Akiyama, T.; Inoue, J.-i.; Fujimoto, Y. *Org. Lett.* **2017,** *19,* 3079–3082. doi:10.1021/acs.orglett.7b01158
42. Borio, A.; Hofinger, A.; Kosma, P.; Zamyatina, A. *Tetrahedron Lett.* **2017,** *58,* 2826–2829. doi:10.1016/j.tetlet.2017.06.014
43. Gin, D. *J. Carbohydr. Chem.* **2002,** *21,* 645–665. doi:10.1081/CAR-120016485
44. Yang, Y.; Zhang, X.; Yu, B. *Nat. Prod. Rep.* **2015,** *32,* 1331–1355. doi:10.1039/C5NP00033E
45. Grynkiewicz, G. *Carbohydr. Res.* **1977,** *53,* C11–C12. doi:10.1016/S0008-6215(00)85467-1
46. Hain, J.; Chandrasekaran, V.; Lindhorst, T. K. *Isr. J. Chem.* **2015,** *35,* 383–386. doi:10.1002/ijch.201400211
47. Garegg, P. J.; Iversen, T.; Norberg, T. *Carbohydr. Res.* **1979,** *73,* 313–314. doi:10.1016/S0008-6215(00)85506-8
48. Kometani, T.; Kondo, H.; Fujimori, Y. *Synthesis* **1988,** 1005–1007. doi:10.1055/s-1988-27788
49. Roush, W. R.; Lin, X. F. *J. Org. Chem.* **1991,** *56,* 5740–5742. doi:10.1021/jo00020a003
50. Chida, N.; Ohtsuka, M.; Nakazawa, K.; Ogawa, S. *J. Org. Chem.* **1991,** *56,* 2976–2983. doi:10.1021/jo00009a009
51. Roush, W. R.; Lin, X.-F. *Tetrahedron Lett.* **1993,** *34,* 6829–6832. doi:10.1016/S0040-4039(00)91806-9
52. Roush, W. R.; Hartz, R. A.; Gustin, D. J. *J. Am. Chem. Soc.* **1999,** *121,* 1990–1991. doi:10.1021/ja984229e

53. Badman, G. T.; Green, D. V. S.; Voyle, M. *J. Organomet. Chem.* **1990,** *388,* 117–121. doi:10.1016/0022-328X(90)85353-Z
54. Yang, X.; Yu, B. *Chem. – Eur. J.* **2013,** *19,* 8431–8434. doi:10.1002/chem.201301297
55. Shan, M.; Sharif, E. U.; O'Doherty, G. A. *Angew. Chem.* **2010,** *122,* 9682–9685. doi:10.1002/ange.201005329
56. Luo, S.-Y.; Jang, Y.-J.; Liu, J.-Y.; Chu, C.-S.; Liao, C.-C.; Hung, S.-C. *Angew. Chem.* **2008,** *120,* 8202–8205. doi:10.1002/ange.200802693
57. Ferroud, D.; Collard, J.; Klich, M.; Dupuis-Hamelin, C.; Mauvais, P.; Lassaigne, P.; Bonnefoy, A.; Musicki, B. *Bioorg. Med. Chem. Lett.* **1999,** *9,* 2881–2886. doi:10.1016/S0960-894X(99)00493-X
58. Vaccaro, W. D.; Davis, H. R., Jr. *Bioorg. Med. Chem. Lett.* **1998,** *8,* 313–318. doi:10.1016/S0960-894X(98)00008-0
59. Tsunoda, T.; Yamamiya, Y.; Itô, S. *Tetrahedron Lett.* **1993,** *34,* 1639–1642. doi:10.1016/0040-4039(93)85029-V
60. Cocinero, E. J.; Stanca-Kaposta, E. C.; Scanlan, E. M.; Gamblin, D. P.; Davis, B. G.; Simons, J. P. *Chem. – Eur. J.* **2008,** *14,* 8947–8955. doi:10.1002/chem.200800474
61. Deng, L.; Tsybina, P.; Gregg, K. J.; Mosi, R.; Zandberg, W. F.; Boraston, A. B.; Vocadlo, D. J. *Bioorg. Med. Chem.* **2013,** *21,* 4839–4845. doi:10.1016/j.bmc.2013.05.062
62. Otani, T.; Tsubogo, T.; Furukawa, N.; Saito, T.; Uchida, K.; Iwama, K.; Kanai, Y.; Yajima, H. *Bioorg. Med. Chem. Lett.* **2008,** *18,* 3582–3584. doi:10.1016/j.bmcl.2008.05.006
63. Donohoe, T. J.; Flores, A.; Bataille, C. J. R.; Churruca, F. *Angew. Chem.* **2009,** *121,* 6629–6632. doi:10.1002/ange.200902840
64. Nie, S.; Li, W.; Yu, B. *J. Am. Chem. Soc.* **2014,** *136,* 4157–4160. doi:10.1021/ja501460j
65. Audisio, D.; Methy-Gonnot, D.; Radanyi, C.; Renoir, J.-M.; Denis, S.; Sauvage, F.; Vergnaud-Gauduchon, J.; Brion, J.-D.; Messaoudi, S.; Alami, M. *Eur. J. Med. Chem.* **2014,** *83,* 498–507. doi:10.1016/j.ejmech.2014.06.067
66. Kawasaki, M.; Fusano, A.; Nigo, T.; Nakamura, S.; Ito, M. N.; Teranishi, Y.; Matsumoto, S.; Toda, H.; Nomura, N.; Sumiyoshi, T. *Bioorg. Med. Chem. Lett.* **2014,** *24,* 2689–2692. doi:10.1016/j.bmcl.2014.04.052
67. Marra, A.; Scherrmann, M.-C.; Dondoni, A.; Casnati, A.; Minari, P.; Ungaro, R. *Angew. Chem.* **1994,** *106,* 2533–2535. doi:10.1002/ange.19941062317
68. Dondoni, A.; Marra, A.; Scherrmann, M.-C.; Casnati, A.; Sansone, F.; Ungaro, R. *Chem. – Eur. J.* **1997,** *3,* 1774–1782. doi:10.1002/chem.19970031108
69. Ramesh, N. G.; Balasubramanian, K. K. *Tetrahedron* **1995,** *51,* 255–272. doi:10.1016/0040-4020(94)00939-R
70. Sobti, A.; Sulikowski, G. A. *Tetrahedron Lett.* **1994,** *35,* 3661–3664. doi:10.1016/S0040-4039(00)73065-6
71. Booma, C.; Balasubramanian, K. K. *Tetrahedron Lett.* **1993,** *34,* 6757–6760. doi:10.1016/S0040-4039(00)61694-5
72. Michigami, K.; Hayashi, M. *Tetrahedron* **2012,** *68,* 1092–1096. doi:10.1016/j.tet.2011.11.084
73. Guthrie, R. D.; Irvine, R. W.; Davison, B. E.; Henrick, K.; Trotter, J. *J. Chem. Soc., Perkin Trans. 2* **1981,** 468–472. doi:10.1039/P29810000468
74. Szarek, W. A.; Jarrell, H. C.; Jones, J. K. N. *Carbohydr. Res.* **1977,** *57,* C13–C16. doi:10.1016/S0008-6215(00)81946-1
75. Grynkiewicz, G.; Zamojski, A. *Synth. Commun.* **1978,** *8,* 491–496. doi:10.1080/00397917808063578
76. Guthrie, R. D.; Jenkins, I. D.; Yamasaki, R. *Aust. J. Chem.* **1982,** *35,* 1003–1018. doi:10.1071/CH9821003
77. Thanh Nga, T. T.; Ménage, C.; Bégué, J.-P.; Bonnet-Delpon, D.; Gantier, J. C.; Pradines, B.; Doury, J.-C.; Dinh Thac, T. *J. Med. Chem.* **1998,** *41,* 4101–4108. doi:10.1021/jm9810147
78. Gueyrard, D.; Rollin, P.; Thanh Nga, T. T.; Ourévitch, M.; Bégué, J.-P.; Bonnet-Delpon, D. *Carbohydr. Res.* **1999,** *318,* 171–179. doi:10.1016/S0008-6215(99)00089-0
79. Kato, K.; Mitsunobu, O. *J. Org. Chem.* **1970,** *35,* 4227–4229. doi:10.1021/jo00837a617
80. Camp, D.; Jenkins, I. D. *Aust. J. Chem.* **1990,** *43,* 161–168. doi:10.1071/CH9900161
81. Falconer, R. A.; Jablonkai, I.; Toth, I. *Tetrahedron Lett.* **1999,** *40,* 8663–8666. doi:10.1016/S0040-4039(99)01834-1
82. Ohnishi, Y.; Ichikawa, M.; Ichikawa, Y. *Bioorg. Med. Chem. Lett.* **2000,** *10,* 1289–1291. doi:10.1016/S0960-894X(00)00223-7
83. Jurczak, J.; Grynkiewicz, G.; Zamojski, A. *Carbohydr. Res.* **1975,** *39,* 147–150. doi:10.1016/S0008-6215(00)82648-8
84. Grynkiewicz, G.; Jurczak, J. *Carbohydr. Res.* **1975,** *43,* 188–191. doi:10.1016/S0008-6215(00)83985-3
85. Nishimura, Y.; Shitara, E.; Takeuchi, T. *Tetrahedron Lett.* **1999,** *40,* 2351–2354. doi:10.1016/S0040-4039(99)00184-7
86. Turner, J. J.; Wilschut, N.; Overkleeft, H. S.; Klaffke, W.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron Lett.* **1999,** *40,* 7039–7042. doi:10.1016/S0040-4039(99)01452-5
87. Petermichl, M.; Loscher, S.; Schobert, R. *Angew. Chem.* **2016,** *128,* 10276–10279. doi:10.1002/ange.201604912
88. Igarashi, Y.; Takagi, K.; Kajiura, T.; Furumai, T.; Oki, T. *J. Antibiot.* **1998,** *51,* 915–920. doi:10.7164/antibiotics.51.915
89. Guo, Y.; Sulikowski, G. A. *J. Am. Chem. Soc.* **1998,** *120,* 1392–1397. doi:10.1021/ja973348b
90. Ohkubo, M.; Nishimura, T.; Jona, H.; Honma, T.; Ito, S.; Morishima, H. *Tetrahedron* **1997,** *53,* 5937–5950. doi:10.1016/S0040-4020(97)00286-X
91. Ohkubo, M.; Nishimura, T.; Honma, T.; Nishimura, I.; Ito, S.; Yoshinari, T.; Arakawa, H.; Suda, H.; Morishima, H.; Nishimura, S. *Bioorg. Med. Chem. Lett.* **1999,** *9,* 3307–3312. doi:10.1016/S0960-894X(99)00595-8
92. Zembower, D. E.; Zhang, H.; Lineswala, J. P.; Kuffel, M. J.; Aytes, S. A.; Ames, M. M. *Bioorg. Med. Chem. Lett.* **1999,** *9,* 145–150. doi:10.1016/S0960-894X(98)00710-0
93. Voldoire, A.; Sancelme, M.; Prudhomme, M.; Colson, P.; Houssier, C.; Bailly, C.; Léonce, S.; Lambel, S. *Bioorg. Med. Chem.* **2001,** *9,* 357–365. doi:10.1016/S0968-0896(00)00251-0
94. Marminon, C.; Pierré, A.; Pfeiffer, B.; Pérez, V.; Léonce, S.; Renard, P.; Prudhomme, M. *Bioorg. Med. Chem.* **2003,** *11,* 679–687. doi:10.1016/S0968-0896(02)00532-1
95. Messaoudi, S.; Anizon, F.; Pfeiffer, B.; Prudhomme, M. *Tetrahedron* **2005,** *61,* 7304–7316. doi:10.1016/j.tet.2005.04.043
96. Messaoudi, S.; Anizon, F.; Peixoto, P.; David-Cordonnier, M.-H.; Golsteyn, R. M.; Léonce, S.; Pfeiffer, B.; Prudhomme, M. *Bioorg. Med. Chem.* **2006,** *14,* 7551–7562. doi:10.1016/j.bmc.2006.07.013
97. Hugon, B.; Pfeiffer, B.; Renard, P.; Prudhomme, M. *Tetrahedron Lett.* **2003,** *44,* 4607–4611. doi:10.1016/S0040-4039(03)00924-9
98. Hénon, H.; Messaoudi, S.; Hugon, B.; Anizon, F.; Pfeiffer, B.; Prudhomme, M. *Tetrahedron* **2005,** *61,* 5599–5614. doi:10.1016/j.tet.2005.03.101
99. Hénon, H.; Anizon, F.; Pfeiffer, B.; Prudhomme, M. *Tetrahedron* **2006,** *62,* 1116–1123. doi:10.1016/j.tet.2005.10.077

100. Hugon, B.; Anizon, F.; Bailly, C.; Golsteyn, R. M.; Pierré, A.; Léonce, S.; Hickman, J.; Pfeiffer, B.; Prudhomme, M. *Bioorg. Med. Chem.* **2007,** *15,* 5965–5980. doi:10.1016/j.bmc.2007.05.073
101. Wang, J.; Soundarajan, N.; Liu, N.; Zimmermann, K.; Naidu, B. N. *Tetrahedron Lett.* **2005,** *46,* 907–910. doi:10.1016/j.tetlet.2004.12.068
102. Szarek, W. A.; Depew, C.; Jarrell, H. C.; Jones, J. K. N. *J. Chem. Soc., Chem. Commun.* **1975,** 648–649. doi:10.1039/c39750000648
103. Hertel, L. W.; Grossmana, C. S.; Kroina, J. S.; Mineishi, S.; Chubb, S.; Novak, B.; Plunkett, W. *Nucleosides Nucleotides* **1989,** *8,* 951–955. doi:10.1080/07328318908054252
104. Kotra, L. P.; Xiang, Y.; Newton, M. G.; Schinazi, R. F.; Cheng, Y.-C.; Chu, K. C. *J. Med. Chem.* **1997,** *40,* 3635–3644. doi:10.1021/jm970275y
105. De Napoli, L.; Di Fabio, G.; Messere, A.; Montesarchio, D.; Piccialli, G.; Varra, M. *J. Chem. Soc., Perkin Trans. 1* **1999,** 3489–3493. doi:10.1039/a906195i
106. Downey, A. M.; Richter, C.; Pohl, R.; Mahrwald, R.; Hocek, M. *Org. Lett.* **2015,** *17,* 4604–4607. doi:10.1021/acs.orglett.5b02332
107. Seio, K.; Tokugawa, M.; Kaneko, K.; Shiozawa, T.; Masaki, Y. *Synlett* **2017,** *28,* 2014–2017. doi:10.1055/s-0036-1588445
108. Sugimura, H.; Koizumi, A.; Kiyohara, W. *Nucleosides, Nucleotides Nucleic Acids* **2003,** *22,* 727–729. doi:10.1081/NCN-120022620
109. Grochowski, E.; Jurczak, J. *Carbohydr. Res.* **1976,** *50,* C15–C16. doi:10.1016/S0008-6215(00)83868-9
110. Jurczak, J. *Carbohydr. Res.* **1982,** *104,* C18–C19. doi:10.1016/S0008-6215(00)82601-4
111. Nashed, E. M.; Grochowski, E.; Czyzewska, E. *Carbohydr. Res.* **1990,** *196,* 184–190. doi:10.1016/0008-6215(90)84118-E
112. Nicolaou, K. C.; Groneberg, R. D. *J. Am. Chem. Soc.* **1990,** *112,* 4085–4086. doi:10.1021/ja00166a082
113. Yang, D.; Kim, S. H.; Kahne, D. *J. Am. Chem. Soc.* **1991,** *113,* 4715–4716. doi:10.1021/ja00012a069
114. Hosokawa, S.; Ogura, T.; Togashi, H.; Tatsuta, K. *Tetrahedron Lett.* **2005,** *46,* 333–337. doi:10.1016/j.tetlet.2004.11.004
115. Grochowski, E.; Falent-Kwastowa, E. *J. Chem. Res., Synop.* **1978,** 300–301.
116. Grochowski, E.; Falent-Kwastowa, E. *Pol. J. Chem.* **1980,** *54,* 2229–2232.
117. Grochowski, E.; Stepowska, H. *Synthesis* **1988,** 795–797. doi:10.1055/s-1988-27710
118. Kunz, H.; Sager, W. *Helv. Chim. Acta* **1985,** *68,* 283–287. doi:10.1002/hlca.19850680134
119. Schörkhuber, W.; Zbiral, E. *Liebigs Ann. Chem.* **1980,** 1455–1469. doi:10.1002/jlac.198019800915
120. Besset, C.; Chambert, S.; Fenet, B.; Queneau, Y. *Tetrahedron Lett.* **2009,** *50,* 7043–7047. doi:10.1016/j.tetlet.2009.09.173
121. Larabi, M.-L.; Fréchou, C.; Demailly, G. *Tetrahedron Lett.* **1994,** *35,* 2175–2178. doi:10.1016/S0040-4039(00)76789-X
122. Kimura, J.; Kawashima, A.; Sugizaki, M.; Nemoto, N.; Mitsunobu, O. *J. Chem. Soc., Chem. Commun.* **1979,** 303–304. doi:10.1039/c39790000303

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Pathoblockers or antivirulence drugs as a new option for the treatment of bacterial infections

Matthew B. Calvert[1,2], Varsha R. Jumde[1,2] and Alexander Titz*[1,2,3]

Review

Open Access

Address:
[1]Chemical Biology of Carbohydrates, Helmholtz Institute for Pharmaceutical Research Saarland (HIPS), Helmholtz Centre for Infection Research (HZI), D-66123 Saarbrücken, Germany, [2]Deutsches Zentrum für Infektionsforschung (DZIF), Standort Hannover-Braunschweig, Germany and [3]Department of Pharmacy, Saarland University, Saarbrücken, Germany

Email:
Alexander Titz* - alexander.titz@helmholtz-hzi.de

* Corresponding author

Keywords:
antimicrobial resistance; bacterial adhesins; bacterial toxins; pathoblockers; quorum sensing

Beilstein J. Org. Chem. **2018,** *14,* 2607–2617.
doi:10.3762/bjoc.14.239

Received: 11 June 2018
Accepted: 20 September 2018
Published: 11 October 2018

This article is part of the Thematic Series "The glycosciences".

Guest Editor: A. Hoffmann-Röder

Abstract

The rapid development of antimicrobial resistance is threatening mankind to such an extent that the World Health Organization expects more deaths from infections than from cancer in 2050 if current trends continue. To avoid this scenario, new classes of anti-infectives must urgently be developed. Antibiotics with new modes of action are needed, but other concepts are also currently being pursued. Targeting bacterial virulence as a means of blocking pathogenicity is a promising new strategy for disarming pathogens. Furthermore, it is believed that this new approach is less susceptible towards resistance development. In this review, recent examples of anti-infective compounds acting on several types of bacterial targets, e.g., adhesins, toxins and bacterial communication, are described.

Review

1. Antimicrobial resistance crisis for bacterial infections

The current crisis caused by antimicrobial resistance [1,2] demands new strategies to fight infections. Antibiotics have served as life-saving drugs during the last 100 years and rescued the world from a situation where practically untreatable infections with high mortality rates were the norm. However, starting in the 1960s, the delusive belief that the available antibiotics were sufficiently effective to treat all infections led to a decline in the development of new antibiotics, with very few new antibiotics addressing a novel mode of action being brought to the market over the last four decades [3].

In parallel with the decline of new antibiotics, resistance towards these widely used drugs has evolved at a high pace and multi as well as extreme drug resistant (MDR/XDR) strains of pathogens are now commonplace. Exposure of bacteria to compounds directly acting on bacterial viability, such as antibiotics,

intrinsically leads to the development of resistance as a matter of microbial survival. This so-called selection pressure can lead to the overgrowth of the initial infective population with a resistant variant of the pathogen, rendering the antibiotic substance ineffective. Especially prevalent in the hospital setting, the abundance of resistance prevents efficient treatment of infected patients. The so-called ESKAPE pathogens, [4] *Enterococcus faecium, Staphylococcus aureus, Klebsiella pneumoniae, Acinetobacter baumanii, Pseudomonas aeruginosa,* and *Enterobacter* species, were initially identified as the most problematic ones. In 2017, an extended list of twelve pathogens, currently considered as those with the highest importance, was published by the WHO [5]. Emphasizing the current crisis, in 2017 one report described a patient infected with a pan-resistant *Klebsiella* strain, where no available drug was efficacious and the patient finally died from septic shock [6]. Therefore, new antibiotics and new alternative treatments are urgently needed.

2. Concept of antivirulence drugs or pathoblockers

Bacterial virulence is the prime determinant for the deterioration of an infected patient's health. Blocking bacterial virulence, or pathogenicity, is a new approach that has emerged over the last decade [7-9]. The pubmed.gov database yields 292 references on the topic (as of 06/08/2018), with an exponential increase over the years. Unfortunately, as the terms 'antivirulence' and 'pathoblocker' are often used interchangeably, many publications in the field are not found in this type of search, for example the pioneering review by Clatworthy et al. in 2007, entitled 'Targeting virulence: a new paradigm for antimicrobial therapy' [8], which has been cited approximately 800 times.

In sharp contrast to traditional antibiotics that kill or impair bacterial viability, this new approach aims to disarm the pathogen. Interfering with the interaction of the pathogen with its host in this way is believed to both reduce damage to the host and to enable the host to clear the microbe from its system. Furthermore, as antivirulence drugs do not kill, it is believed that the selection pressure for resistant mutants will be significantly reduced. In some cases, however, resistance has already been observed (e.g., through increased expression of efflux pumps to circumvent quorum quenching), with the likelihood of the appearance of resistance mechanisms seemingly dependent upon the importance of the targeted virulence factor to the pathogen [10].

3. Blocking adhesion and biofilm formation

Bacterial adhesion to the host's tissue is the initial step of every infection. In many cases, microbial adhesion is mediated by carbohydrate-binding proteins, so-called lectins, which recognize glycoconjugates on the surface of cells and tissue. Surface exposed glycoconjugates are highly abundant on all living cells and are generally referred to as the glycocalyx. Bacterial lectins act as adhesins with defined carbohydrate-binding specificities, in order to establish and maintain infection of the host's various tissues and organs. Therefore, the inhibition of this adhesion process using glycomimetics as pathoblockers has developed as an area of active research in the last two decades [11,12].

Uropathogenic *Escherichia coli* (UPEC) is a major cause for chronic and recurrent urinary tract infections. These bacteria employ lectins in order to attach to and invade bladder and kidney tissue, and to promote biofilm formation. Bladder-adhesive FimH is a mannose-specific lectin and the kidney-adhesive PapG binds galactosides. In a second indication, FimH also mediates the attachment of *E. coli* to the gut, inducing inflammation in Crohn's disease [13]. The crystal structure of FimH was published by Hultgren and Knight et al. in 1999 [14]. FimH is highly specific for α-D-mannoside ligands with this residue residing in a carbohydrate binding pocket with its α-linked substituent towards an adjacent cleft. This substituent, termed the aglycon, can also interact with the tyrosine gate formed by Tyr48 and Tyr137 [15,16], as well as form hydrogen bonds and electrostatic interactions with the Arg98/Glu50 salt bridge of the protein. Taking this coordination geometry into consideration for further ligand optimization, it was found important to focus on the aglycon part of the mannosides.

The attachment of lipophilic aglycons to an α-linked mannose residue was identified to increase the binding potency tremendously due to the opening of a lipophilic cleft on FimH, the tyrosine gate [15]. Various alkyl mannosides **1** (Figure 1) were analyzed and *n*-heptyl mannoside (**1b**) revealed the highest potency, as a result of it having the optimal length to bind to the tyrosine gate. Lindhorst and co-workers have demonstrated that mannosides with an extended aromatic aglycon could further improve the interaction as shown for compounds **2** and **3**. Their relative inhibitory potential (RIP), which is benchmarked with the reference methyl α-D-mannoside (**1a**) defined as RIP = 1, was increased up to 6900-fold [17].

The biphenyl mannosides (e.g., **4**, **5**) have subsequently been identified by the Ernst and Hultgren/Janetka groups as promising inhibitors of FimH-mediated bacterial adhesion in mice [18,19]. These compounds have been extensively optimized in many works published by both groups, culminating in the identification of mannophosphates as prodrugs to increase oral bioavailability [20] and mannose *C*-glycosides, such as compound **6** , demonstrating enhanced in vivo metabolic stability [21]. Such biphenyl mannoside-derived compounds are the current state of the art and are being further developed by the

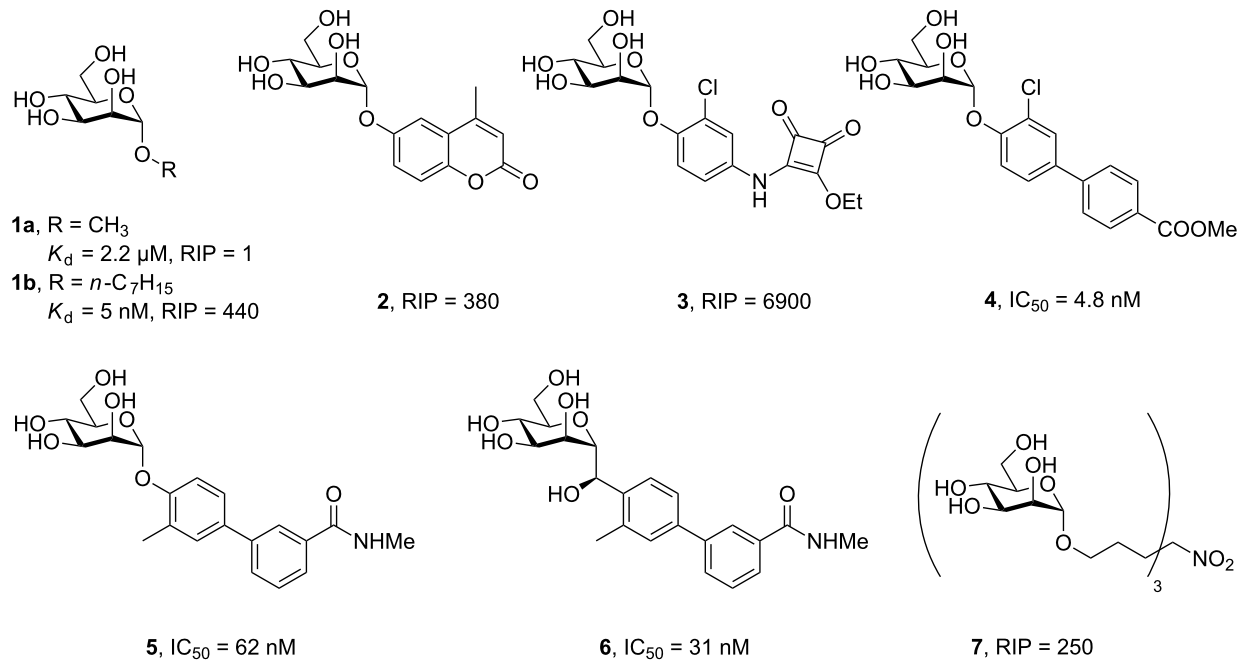


Figure 1: Mannosides as inhibitors of the lectin FimH from uropathogenic *Escherichia coli*.

company Fimbrion (St Louis, MO) in collaboration with GlaxoSmithKline.

In many cases, lectins have more than one carbohydrate binding site or are clustered in proximity. Therefore, multivalent display of lectin and ligand results in a higher avidity [22,23]. The Lindhorst group also synthesized and analyzed so-called glycoclusters, e.g., **7** (Figure 1), where the saccharide moiety is displayed in a multivalent fashion [17,24-26]. When this simple trimannosylated compound was tested in a whole cell ELISA, it was shown that the apparent binding affinity increases by a factor of 250 versus methyl α-D-mannoside, while the valency only increased by a factor three.

It should be noted that the full length FimH adhesin consists of two domains, a lectin and a pilin domain that are interconnected by a hinge region. Interestingly, in vitro binding studies have been performed with the lectin domain only. Recent works suggested that the conformation of the two domains influence the protein's affinity towards inhibitors and the biologically relevant state is a matter of ongoing research [27,28].

Another adhesin of uropathogenic *E. coli* is FmlH, which is located at the tip of F9 pili and binds β-D-galactosides with moderate potency. It could be shown that this lectin plays an important role in kidney-associated chronic UTIs, as its glycan receptor is abundantly expressed in this organ. In screening assays, 2-nitrophenyl galactoside (**8**) was identified displaying a dissociation constant of 10.6 μM (Figure 2). A detailed optimization program run by the Hultgren and Janetka groups yielded derivatives of *N*-acetyl galactosamine bearing biphenyl aglycons, such as compound **9**, as very potent ligands of this protein. Beyond blocking the binding site of FmlH on the pili of *E. coli*, these compounds proved effective at promoting eradication of bacteria from murine kidney in synergy with a mannoside for FimH [29].

P. aeruginosa is one of the highly resistant ESKAPE pathogens that, in addition to antimicrobial resistance, forms biofilms, a complex matrix of extracellular polysaccharides, polypeptides and DNA, which act as an additional protective barrier [30]. *P. aeruginosa* employs two lectins for biofilm formation and host–cell adhesion: proteins LecA and LecB [31,32] which are also important for mediating bacterial virulence in vivo [33]. Therefore, both LecA and LecB have served as targets for pathoblocker development [22,23,30,34,35]. As a result of the comparatively low affinity of both lectins towards their natural carbohydrate ligands (α-galactosides for LecA and α-fucosides and mannosides for LecB), numerous multivalent presentations have been developed with the aim to improve affinity based on avidity [22].

LecA recognizes aryl β-D-galactosides with moderate potency, e.g., compound **10** (Figure 2). However, attempts to optimize the potency by varying the aryl substitution resulted in a flat SAR with only little variation in potency among the substitu-

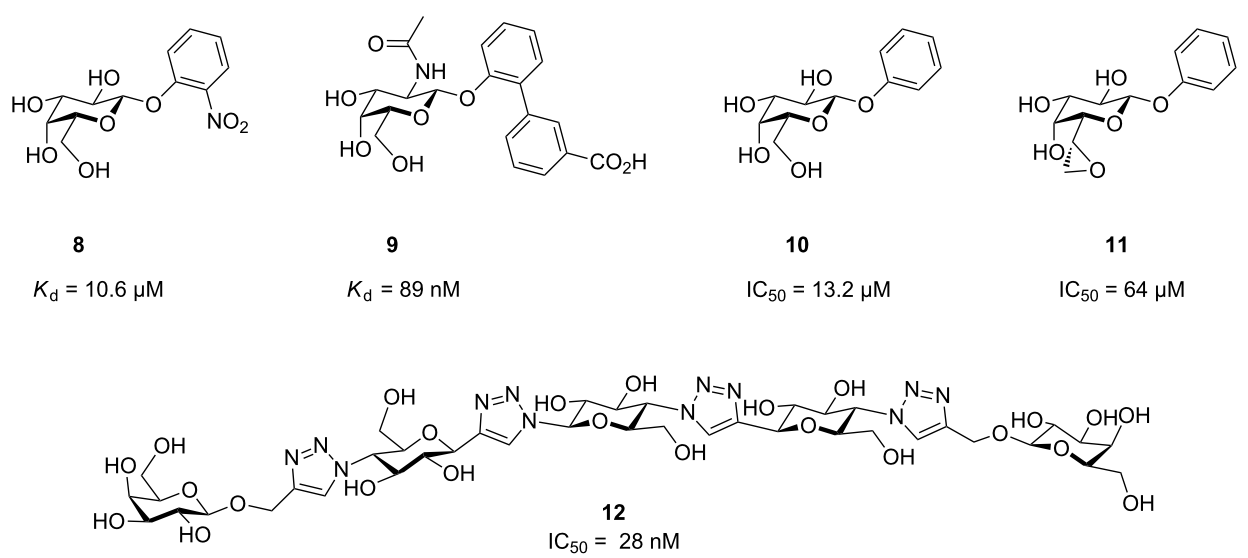


Figure 2: Galactosides targeting uropathogenic *Escherichia coli* FmlH (compounds **8** and **9**) and *Pseudomonas aeruginosa* LecA (compounds **10**–**12**).

ents analyzed [30,36-39]. Just recently, in an attempt to search for new pharmacophores, Titz et al. have reported the synthesis of the epoxyheptose derivative **11** targeting a cysteine residue of LecA with its electrophilic epoxide warhead [40]. It could be demonstrated that **11** is a covalent lectin inhibitor, which provided the first proof-of-concept for this new approach to lectin inhibition. To date, the most potent LecA inhibitor **12** has been designed by the Pieters group, where two galactoside moieties are optimally oriented in space to simultaneously bind to two of the four binding sites in LecA [41]. This optimal geometric match to LecA resulted in low nanomolar inhibition of LecA.

LecB has been studied in detail using multivalent and small molecule approaches. Interestingly, the sequence of LecB differs among clinical isolates of this highly variable pathogen, with some mutations in close proximity to the carbohydrate binding site, but carbohydrate-binding function is preserved across all lectins investigated [42,43]. The glycopeptide dendrimer **13** (Figure 3) showed potent inhibition of biofilm formation and synergistically acted with tobramycin to eradicate biofilm-embedded bacteria in vitro [44,45]. Also, the fucosylated tetravalent calixarene **14** proved a potent ligand to LecB (K_d = 48 nM) and showed beneficial effects in an acute murine pulmonary infection model following inhalative administration [46]. Despite its LecB-mediated in vivo activity, this compound had no effect on biofilms in vitro at concentrations up to 2000-fold above the K_d; a biofilm reduction by 80% could be achieved at concentrations as high as 100000-fold above K_d (5 mM). For a future systemic application, Titz et al. have developed small molecule LecB inhibitors derived from mannose and obtained potent monovalent inhibitors (compound **15**) of LecB-mediated bacterial adhesion [47]. The sulfonamide **15** and cinnamide **16** were developed to take advantage of interactions with a nearby shallow pocket, and indeed these compounds showed superior thermodynamics and kinetics of binding to LecB compared to mannose, resulting in a prolonged receptor residence time of several minutes [48]. In a complementary approach, glycomimetic *C*-glycoside **17** was obtained, aiming at improved metabolic stability and selectivity [49]. Both approaches were then combined into low molecular weight *C*-glycosidic sulfonamides, which resulted in very potent LecB and *P. aeruginosa* biofilm inhibitors with over 80% inhibition at a concentration of 100 μM [50]. Compound **18** of this series further showed very good in vitro stability against plasma and liver microsomes, absence of cytotoxicity, and excellent oral bioavailability in mice.

4. Direct toxin inhibition

Numerous bacteria secrete toxins that are responsible for acute virulence. Various small molecule and antibody approaches target the inhibition of bacterial toxins in order to antagonize bacterial virulence [51].

AB toxins are widespread among species and consist of a catalytically active A-domain and one or more units of a receptor-binding domain B. The B domain is responsible for binding to a cell-surface receptor, which engages in receptor-mediated cellular uptake. The AB toxin then migrates either via a classical endocytosis pathway, or via retrograde transport through the secretory pathway into the cytosol, where the A domain can exert its toxic property. AB_5 toxins are abundant in many pathogens and the B domain is a carbohydrate-binding domain

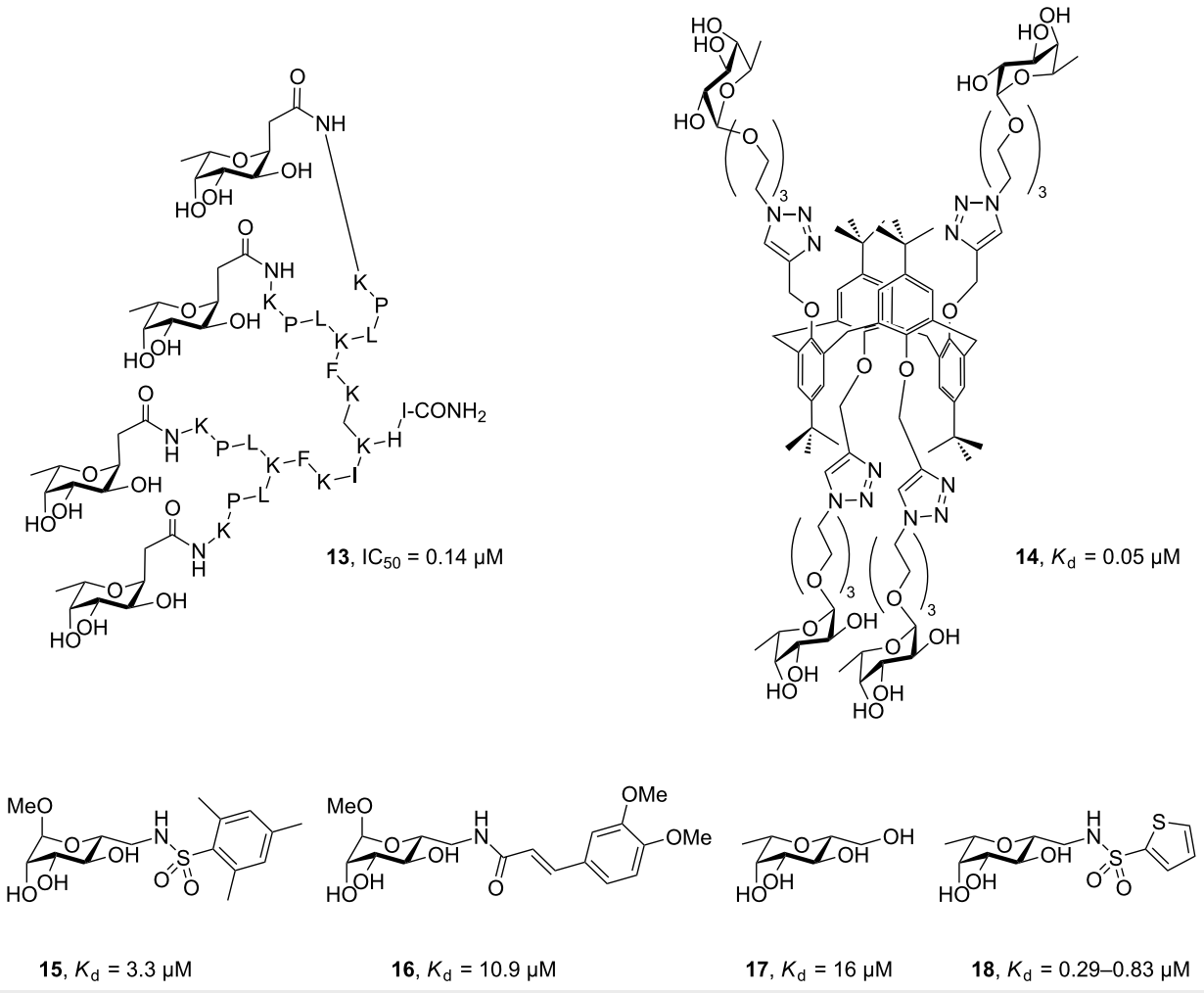


Figure 3: Mannosides and fucosides as inhibitors of *P. aeruginosa* LecB.

for cell-surface binding. Numerous inhibitors have been developed against AB toxins, targeting toxin transcription, assembly, receptor binding and enzyme function [51].

A set of antibodies against diverse toxins has recently been approved for therapeutic use, which demonstrates the scientific and medical feasibility of entering the market with an antivirulence drug. The monoclonal antibody bezlotoxumab binds to *C. difficile* toxin B and was approved for the prevention of infections with this intestinal pathogen in 2016 [52]. Obiltoxaximab [53] and raxibacumab [54] are two approved antibody treatments for inhalative anthrax that target the *Bacillus anthracis* toxin. It is likely that small molecules will also benefit from the knowhow obtained during the antibody-related clinical studies and it is probably only a matter of time before a small molecule drug is approved.

Pore forming toxins constitute another large set of virulence factors playing crucial roles in acute virulence [55]. *Staphylococcus aureus* infections are characterized by the toxic action of bacterial α-hemolysin, a pore forming toxin leading to hemolysis. The antibody MEDI4893, which blocks *S. aureus* α-hemolysin, is currently in phase II clinical trials [56]. Despite the challenges associated with the large size of the pore structures, small molecules have also been widely studied as antitoxins and there are examples at various stages of the discovery process [51]. An important example of this concept is the application of cyclodextrins as anti-infectives [57], with the ornithine-substituted compound **19** (Figure 4) being shown to be able to block various pore-forming toxins, as well as successfully preventing and treating infections by *S. aureus* in mice [58].

Enterohemorragic *E. coli* (EHEC) bacteria produce Shiga toxins Stx1 and Stx2 that belong to the group of AB_5 toxins. These Shiga toxins are the causative agents for bacterial virulence in the gut of the infected host and bind to the P blood group antigens that bear terminal Gal-α-1,4-Gal disaccharides. Blocking

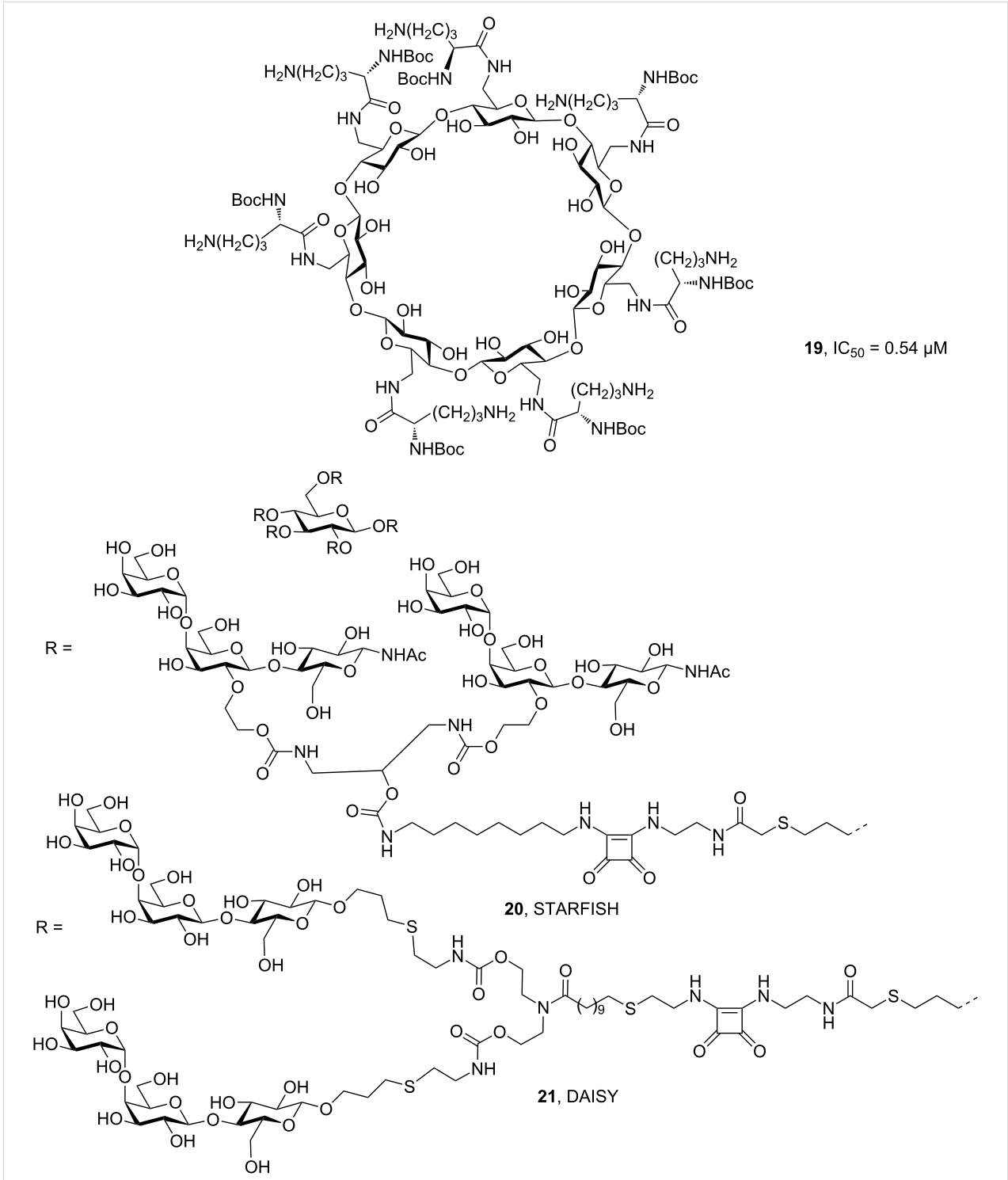


Figure 4: β-Cyclodextrin-based antitoxin **19** against *S. aureus* α-hemolysin and the decavalent Shiga toxin inhibitors STARFISH (**20**) and DAISY (**21**).

Stx1 with decavalent molecules STARFISH (**20**) and Stx1 and Stx2 with DAISY (**21**) resulted in a full protection of mice from the toxin [59,60]. Another set of compounds called SUPER TWIG bears P blood group antigens on the antennae of a carbosilane dendrimer and was developed as an intravenously applied scavenger of circulating Shiga toxins to prevent the most severe complications in these infections [61].

While not typically classed as toxins, bacterial proteolytic enzymes, such as collagenases or elastases, often account for host

cell damage and immune evasion. Janda and co-workers developed thiol-based small molecules targeting the active site zinc ion in *P. aeruginosa* elastase LasB showing prolonged survival in a *C. elegans* infection model [62]. Hydroxamic acid-containing molecules addressing the same enzyme were developed by the Hartmann group; these compounds showed a moderate reduction of biofilm formation resulting from a lowered release of the structural biofilm component extracellular DNA [63]. Recently, inhibitors of the clostridial collagenase were discovered that showed high selectivity for the bacterial enzyme over related host metalloproteases [64]. It is hoped that continued research in this area will lead to a complementary class of antivirulence drugs against *Clostridium difficile*, adding to the existing repertoire of clostridial AB antitoxins discussed previously.

5. Toxin secretion

A complementary approach to toxin inhibition is the interference with the ability of the bacterium to release the toxin into its environment, i.e., toxin secretion. Many different secretion systems exist in bacteria [65] and the Gram-negative specific type III secretion system (TTSS) is a focus of current research. TTSS is a major virulence determinant in a number of pathogens, including *P. aeruginosa*. In TTSS, toxins are secreted from the bacterial cytosol across the bacterial membranes and the extracellular environment through a needle-like structure into a host cell. The blockade of toxin secretion or needle assembly has been an active area of research, and small molecules as well as antibodies are currently being developed [30,66,67]. The TTSS needle tip protein PcrV was found to be a suitable target to prevent toxin secretion. The anti-PcrV antibody KB001 [68] and the bifunctional antibody MEDI3902 [69], which targets PcrV and the biofilm-associated exopolysaccharide psl, are both currently in phase II clinical trials.

6. Bacterial communication

Quorum sensing (QS) is employed by bacteria to communicate with each other in a given population [70]. In this regulatory mechanism, signal molecules (also known as autoinducers) are constantly secreted by each individual bacterium and at a defined population density the concentration of this molecular messenger reaches a threshold that activates quorum sensing-controlled processes (Figure 5). Many virulence traits are influenced by quorum sensing and thus developing methods to reduce virulence by interfering with bacterial communication is currently a topic of intense research efforts.

Quorum sensing exists in Gram-negative and Gram-positive bacteria. While Gram-positive bacteria often use peptides as signal molecules, Gram-negative bacteria employ *N*-acylhomoserine lactones (AHLs) with subtle differences in their chemical structure, as well as other types of autoinducers (Figure 5). Interestingly, the signaling molecule autoinducer-2 is used by both Gram-positive and Gram-negative species. Because highly structurally similar or even identical molecules are employed for bacterial signaling, it is obvious that bacteria also communi-

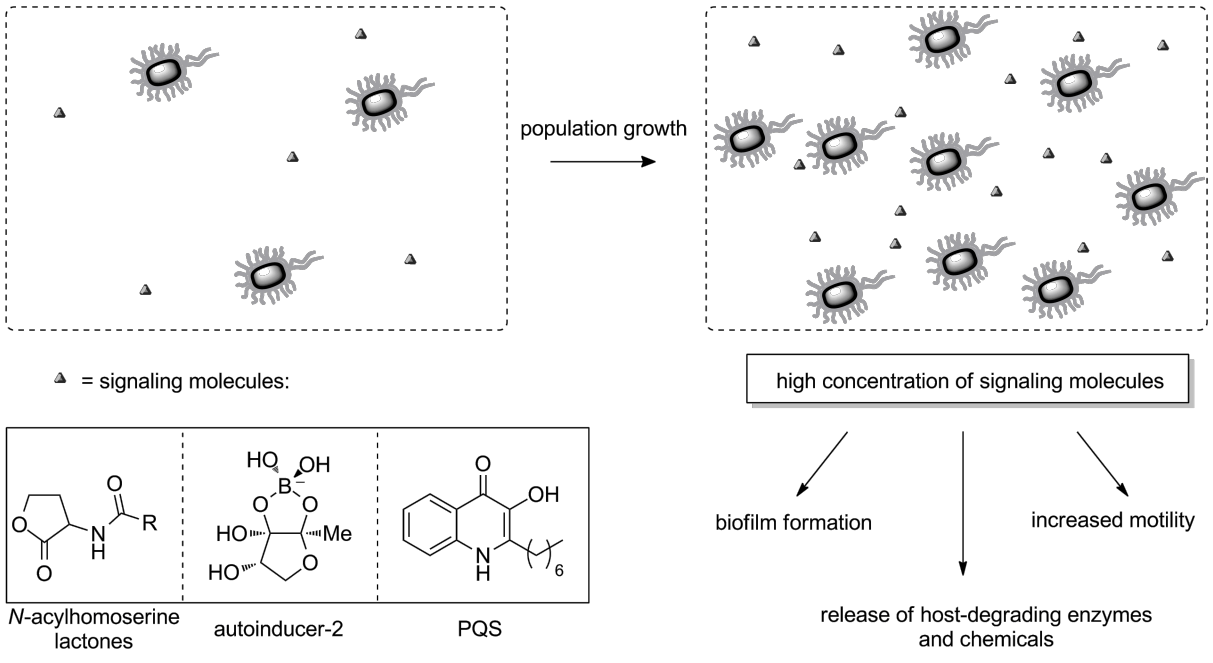


Figure 5: The mechanism of quorum sensing and representative signaling molecules.

cate between species, which can be of use for the bacterial community in co-infections. Often, multiple QS mechanisms exist within one species. For example in *P. aeruginosa*, four signaling systems have been identified to date, which are highly interconnected and mutually influence each other [30]. Some bacteria employ rather specific quorum sensing molecules, such as the Pseudomonas Quinolone Signal (PQS) and its biosynthetic precursors in *P. aeruginosa* some of which are also found in *Burkholderia* [71], two species that often co-infect patients for example in cystic fibrosis airways infections. By blocking QS processes, the release of virulence factors such as host degrading enzymes or chemicals, or the formation of bacterial biofilms, can be inhibited. Numerous reviews have detailed these processes addressing various QS pathways [30,72,73].

The antibiotic azithromycin (**22**), which does not have significant bactericidal activity for *P. aeruginosa*, but interferes with its quorum sensing pathways, was studied in a clinical trial (Figure 6) [74]. The macrolide antibiotic of natural origin, which does not resemble the structures of signal molecules, was shown to reduce the presence of quorum sensing molecules in vitro and in vivo. It prevented the selection of QS-mutants (*lasR*) that rapidly appear in untreated patients and outgrow wild-type bacteria as a result of a fitness advantage. Thus, it may be of help in acute infections to reduce virulence, as stated by Köhler et al.

Many approaches towards developing QS antagonists as tool compounds and drug candidates start from the natural QS signal molecules and mimic their structures. The PQS system of *P. aeruginosa* is particularly attractive and can be considered as a pathogen specific target. The biosynthesis of the PQS signal **23** involves a set of biosynthetic enzymes PqsABCDEH and its autocatalytic receptor PqsR (MvfR). Biaryl methanols (e.g., **24**) function as PqsD transition state analogues, and were shown to inhibit the enzyme and reduce bacterial biofilm formation [75]. Numerous approaches target the signal molecule receptor PqsR, and compounds such as **25**, **26** and **27** successfully inhibited virulence factor production, biofilm formation and virulence in an insect infection model or a murine model [76-78].

Conclusion

The current antimicrobial crisis poses an enormous challenge to society, and requires a joint effort for the development of novel anti-infectives. While there is an urgent need for new antibiotics with novel modes of action that avoid cross-resistance to established drug-resistant strains, the development of antivirulence drugs will address a promising new paradigm in antibacterial therapy, leading to a second anti-infective pillar.

It has to be emphasized that a concerted approach to new anti-infectives is of the utmost importance. Private and public research have to join forces to provide new treatments and to sustain a continuous supply of drugs with novel modes of action, necessary to maintain an arsenal that is able to treat MDR/XDR infections in the future. Some pathogens are well studied and numerous approaches have been developed, e.g., *P. aeruginosa* [30] and *S. aureus* [79,80]. It is, however, questionable why research on some of the most problematic pathogens discussed above, e.g., *Acinetobacter* or *Enterobacter*, is scarce and publications on their biology cover only a small fraction of the literature compared to the well-studied pathogens.

Adding to the prevalent resistance, recent reports [81,82] uncovered the abundance of various resistant pathogenic bacteria in proximity to antibiotic production facilities and their untreated sewage outlets in the Hyderabad area in India. One important factor for resistance development, also in industrialized countries, is the large scale exposure of organisms in the environment to released antibiotics. This can result from high drug concentrations in community waste water combined with

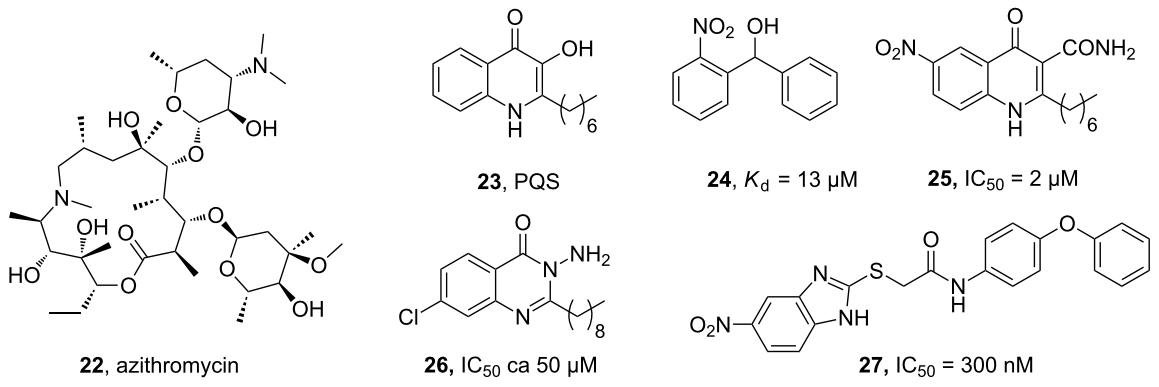


Figure 6: Inhibitors of bacterial quorum sensing.

ineffective drug clearance mechanisms, as well as from the excessive and inappropriate use of antibiotics in the commercial livestock breeding industry. Since numerous targets of classical antibiotics are conserved across a large number of bacterial species, it is obvious that resistance against antibiotics can also develop outside the patient in the environment. This process can take place in environmentally present pathogens, e.g., *P. aeruginosa*, but also in drug-exposed apathogenic bacteria, followed by horizontal transfer of the resistance gene into a pathogen, generating an uncontrollable risk for mankind. In contrast, it can be anticipated that resistance development of antivirulence compounds, which target specific mechanisms of the pathogen–host interplay, is absent outside a patient. Thus a reduced risk of resistance appearing both in the patient and in the environment would provide a benefit of antivirulence drugs over classical antibiotics.

It remains to be established whether antivirulence drugs will be sufficiently effective as a sole treatment, or if they will be used as adjuvants and co-application with antibiotics will be required. Antivirulence compounds dismantling biofilm-protected chronic pathogens or directly inhibiting bacterial factors of acute toxicity/virulence are likely to be successful as future therapies against the impending threat of highly antibiotic-resistant pathogens. In some cases, it was already shown that virulence blockers act synergistically in combination with antibiotics, for example against *P. aeruginosa* biofilms [44]. Therefore, it is likely that combinations of drugs will be applied for drug-resistant bacterial infections, as is currently the state of the art for many viral infections.

ORCID® iDs

Varsha R. Jumde - https://orcid.org/0000-0001-7677-7279
Alexander Titz - https://orcid.org/0000-0001-7408-5084

References

1. European Centre for Disease Prevention and Control. Antimicrobial resistance surveillance in Europe 2015. Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net). ECDC: Stockholm, 2017; https://ecdc.europa.eu/en/publications-data/antimicrobial-resistance-surveillance-europe-2015#no-link (accessed Sept 19, 2018).
2. Antimicrobial resistance: global report on surveillance 2014. World Health Organization; 2014.
3. Cooper, M. A.; Shlaes, D. *Nature* **2011,** *472,* 32. doi:10.1038/472032a
4. Rice, L. B. *J. Infect. Dis.* **2008,** *197,* 1079–1081. doi:10.1086/533452
5. WHO published list of bacteria for which new antibiotics are urgently needed, 2017. Available at: http://www.who.int/mediacentre/news/releases/2017/bacteria-antibiotics-needed/en/. (accessed May 15, 2017).
6. Chen, L.; Todd, R.; Kiehlbauch, J.; Walters, M.; Kallen, A. *Morb. Mortal. Wkly. Rep.* **2017,** *66,* 33. doi:10.15585/mmwr.mm6601a7
7. Dickey, S. W.; Cheung, G. Y. C.; Otto, M. *Nat. Rev. Drug Discovery* **2017,** *16,* 457–471. doi:10.1038/nrd.2017.23
8. Clatworthy, A. E.; Pierson, E.; Hung, D. T. *Nat. Chem. Biol.* **2007,** *3,* 541–548. doi:10.1038/nchembio.2007.24
9. Cegelski, L.; Marshall, G. R.; Eldridge, G. R.; Hultgren, S. J. *Nat. Rev. Microbiol.* **2008,** *6,* 17–27. doi:10.1038/nrmicro1818
10. Allen, R. C.; Popat, R.; Diggle, S. P.; Brown, S. P. *Nat. Rev. Microbiol.* **2014,** *12,* 300–308. doi:10.1038/nrmicro3232
11. Ernst, B.; Magnani, J. L. *Nat. Rev. Drug Discovery* **2009,** *8,* 661–677. doi:10.1038/nrd2852
12. Sharon, N. *Biochim. Biophys. Acta* **2006,** *1760,* 527–537. doi:10.1016/j.bbagen.2005.12.008
13. Mydock-McGrane, L. K.; Hannan, T. J.; Janetka, J. W. *Expert Opin. Drug Discovery* **2017,** *12,* 711–731. doi:10.1080/17460441.2017.1331216
14. Choudhury, D.; Thompson, A.; Stojanoff, V.; Langermann, S.; Pinkner, J.; Hultgren, S. J.; Knight, S. D. *Science* **1999,** *285,* 1061–1066. doi:10.1126/science.285.5430.1061
15. Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhrer, M.; Hung, C.-S.; Pinkner, J.; Slättegård, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. *Mol. Microbiol.* **2005,** *55,* 441–455. doi:10.1111/j.1365-2958.2004.04415.x
16. Fiege, B.; Rabbani, S.; Preston, R. C.; Jakob, R. P.; Zihlmann, P.; Schwardt, O.; Jiang, X.; Maier, T.; Ernst, B. *ChemBioChem* **2015,** *16,* 1235–1246. doi:10.1002/cbic.201402714
17. Sperling, O.; Fuchs, A.; Lindhorst, T. K. *Org. Biomol. Chem.* **2006,** *4,* 3913–3922. doi:10.1039/b610745a
18. Cusumano, C. K.; Pinkner, J. S.; Han, Z.; Greene, S. E.; Ford, B. A.; Crowley, J. R.; Henderson, J. P.; Janetka, J. W.; Hultgren, S. J. *Sci. Transl. Med.* **2011,** *3,* 109ra115. doi:10.1126/scitranslmed.3003021
19. Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Lüthi, C.; Scharenberg, M.; Bezençon, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. *J. Med. Chem.* **2010,** *53,* 8627–8641. doi:10.1021/jm101011y
20. Kleeb, S.; Jiang, X.; Frei, P.; Sigl, A.; Bezençon, J.; Bamberger, K.; Schwardt, O.; Ernst, B. *J. Med. Chem.* **2016,** *59,* 3163–3182. doi:10.1021/acs.jmedchem.5b01923
21. Mydock-McGrane, L.; Cusumano, Z.; Han, Z.; Binkley, J.; Kostakioti, M.; Hannan, T.; Pinkner, J. S.; Klein, R.; Kalas, V.; Crowley, J.; Rath, N. P.; Hultgren, S. J.; Janetka, J. W. *J. Med. Chem.* **2016,** *59,* 9390–9408. doi:10.1021/acs.jmedchem.6b00948
22. Bernardi, A.; Jiménez-Barbero, J.; Casnati, A.; De Castro, C.; Darbre, T.; Fieschi, F.; Finne, J.; Funken, H.; Jaeger, K.-E.; Lahmann, M.; Lindhorst, T. K.; Marradi, M.; Messner, P.; Molinaro, A.; Murphy, P. V.; Nativi, C.; Oscarson, S.; Penadés, S.; Peri, F.; Pieters, R. J.; Renaudet, O.; Reymond, J.-L.; Richichi, B.; Rojo, J.; Sansone, F.; Schäffer, C.; Turnbull, W. B.; Velasco-Torrijos, T.; Vidal, S.; Vincent, S.; Wennekes, T.; Zuilhof, H.; Imberty, A. *Chem. Soc. Rev.* **2013,** *42,* 4709–4727. doi:10.1039/C2CS35408J
23. Cecioni, S.; Imberty, A.; Vidal, S. *Chem. Rev.* **2015,** *115,* 525–561. doi:10.1021/cr500303t
24. Hartmann, M.; Lindhorst, T. K. *Eur. J. Org. Chem.* **2011,** 3583–3609. doi:10.1002/ejoc.201100407
25. Dubber, M.; Sperling, O.; Lindhorst, T. K. *Org. Biomol. Chem.* **2006,** *4,* 3901–3912. doi:10.1039/B610741A
26. Lindhorst, T. K.; Dubber, M. *Carbohydr. Res.* **2015,** *403,* 90–97. doi:10.1016/j.carres.2014.06.032

27. Mayer, K.; Eris, D.; Schwardt, O.; Sager, C. P.; Rabbani, S.; Kleeb, S.; Ernst, B. *J. Med. Chem.* **2017,** *60,* 5646–5662. doi:10.1021/acs.jmedchem.7b00342
28. Schwartz, D. J.; Kalas, V.; Pinkner, J. S.; Chen, S. L.; Spaulding, C. N.; Dodson, K. W.; Hultgren, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2013,** *110,* 15530–15537. doi:10.1073/pnas.1315203110
29. Kalas, V.; Hibbing, M. E.; Maddirala, A. R.; Chugani, R.; Pinkner, J. S.; Mydock-McGrane, L. K.; Conover, M. S.; Janetka, J. W.; Hultgren, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2018,** *115,* E2819–E2828. doi:10.1073/pnas.1720140115
30. Wagner, S.; Sommer, R.; Hinsberger, S.; Lu, C.; Hartmann, R. W.; Empting, M.; Titz, A. *J. Med. Chem.* **2016,** *59,* 5929–5969. doi:10.1021/acs.jmedchem.5b01698
31. Diggle, S. P.; Stacey, R. E.; Dodd, C.; Cámara, M.; Williams, P.; Winzer, K. *Environ. Microbiol.* **2006,** *8,* 1095–1104. doi:10.1111/j.1462-2920.2006.001001.x
32. Tielker, D.; Hacker, S.; Loris, R.; Strathmann, M.; Wingender, J.; Wilhelm, S.; Rosenau, F.; Jaeger, K.-E. *Microbiology (London, U. K.)* **2005,** *151,* 1313–1323. doi:10.1099/mic.0.27701-0
33. Chemani, C.; Imberty, A.; de Bentzmann, S.; Pierre, M.; Wimmerová, M.; Guery, B. P.; Faure, K. *Infect. Immun.* **2009,** *77,* 2065–2075. doi:10.1128/IAI.01204-08
34. Sommer, R.; Joachim, I.; Wagner, S.; Titz, A. *Chimia* **2013,** *67,* 286–290. doi:10.2533/chimia.2013.286
35. Titz, A. Carbohydrate-Based Anti-Virulence Compounds against Chronic Pseudomonas aeruginosa Infections with a Focus on Small Molecules. In *Topics in Medicinal Chemistry;* Seeberger, P. H.; Rademacher, C., Eds.; Springer: Berlin Heidelberg, 2014; Vol. 12 Carbohydrates as Drugs, pp 169–186. doi:10.1007/7355_2014_44
36. Joachim, I.; Rikker, S.; Hauck, D.; Ponader, D.; Boden, S.; Sommer, R.; Hartmann, L.; Titz, A. *Org. Biomol. Chem.* **2016,** *14,* 7933–7948. doi:10.1039/C6OB01313A
37. Casoni, F.; Dupin, L.; Vergoten, G.; Meyer, A.; Ligeour, C.; Géhin, T.; Vidal, O.; Souteyrand, E.; Vasseur, J.-J.; Chevolot, Y.; Morvan, F. *Org. Biomol. Chem.* **2014,** *12,* 9166–9179. doi:10.1039/C4OB01599A
38. Kadam, R. U.; Garg, D.; Schwartz, J.; Visini, R.; Sattler, M.; Stocker, A.; Darbre, T.; Reymond, J.-L. *ACS Chem. Biol.* **2013,** *8,* 1925–1930. doi:10.1021/cb400303w
39. Rodrigue, J.; Ganne, G.; Blanchard, B.; Saucier, C.; Giguère, D.; Shiao, T. C.; Varrot, A.; Imberty, A.; Roy, R. *Org. Biomol. Chem.* **2013,** *11,* 6906–6918. doi:10.1039/c3ob41422a
40. Wagner, S.; Hauck, D.; Hoffmann, M.; Sommer, R.; Joachim, I.; Müller, R.; Imberty, A.; Varrot, A.; Titz, A. *Angew. Chem., Int. Ed.* **2017,** *56,* 16559–16564. doi:10.1002/anie.201709368
41. Pertici, F.; de Mol, N. J.; Kemmink, J.; Pieters, R. J. *Chem. – Eur. J.* **2013,** *19,* 16923–16927. doi:10.1002/chem.201303463
42. Sommer, R.; Wagner, S.; Varrot, A.; Nycholat, C. M.; Khaledi, A.; Haussler, S.; Paulson, J. C.; Imberty, A.; Titz, A. *Chem. Sci.* **2016,** *7,* 4990–5001. doi:10.1039/C6SC00696E
43. Boukerb, A. M.; Decor, A.; Ribun, S.; Tabaroni, R.; Rousset, A.; Commin, L.; Buff, S.; Doléans-Jordheim, A.; Vidal, S.; Varrot, A.; Imberty, A.; Cournoyer, B. *Front. Microbiol.* **2016,** *7,* No. 811. doi:10.3389/fmicb.2016.00811
44. Michaud, G.; Visini, R.; Bergmann, M.; Salerno, G.; Bosco, R.; Gillon, E.; Richichi, B.; Nativi, C.; Imberty, A.; Stocker, A.; Darbre, T.; Reymond, J.-L. *Chem. Sci.* **2016,** *7,* 166–182. doi:10.1039/C5SC03635F
45. Johansson, E. M. V.; Crusz, S. A.; Kolomiets, E.; Buts, L.; Kadam, R. U.; Cacciarini, M.; Bartels, K.-M.; Diggle, S. P.; Cámara, M.; Williams, P.; Loris, R.; Nativi, C.; Rosenau, F.; Jaeger, K.-E.; Darbre, T.; Reymond, J.-L. *Chem. Biol.* **2008,** *15,* 1249–1257. doi:10.1016/j.chembiol.2008.10.009
46. Boukerb, A. M.; Rousset, A.; Galanos, N.; Méar, J.-B.; Thepaut, M.; Grandjean, T.; Gillon, E.; Cecioni, S.; Abderrahmen, C.; Faure, K.; Redelberger, D.; Kipnis, E.; Dessein, R.; Havet, S.; Darblade, B.; Matthews, S. E.; de Bentzmann, S.; Guéry, B.; Cournoyer, B.; Imberty, A.; Vidal, S. *J. Med. Chem.* **2014,** *57,* 10275–10289. doi:10.1021/jm500038p
47. Hauck, D.; Joachim, I.; Frommeyer, B.; Varrot, A.; Philipp, B.; Möller, H. M.; Imberty, A.; Exner, T. E.; Titz, A. *ACS Chem. Biol.* **2013,** *8,* 1775–1784. doi:10.1021/cb400371r
48. Sommer, R.; Hauck, D.; Varrot, A.; Wagner, S.; Audfray, A.; Prestel, A.; Möller, H. M.; Imberty, A.; Titz, A. *ChemistryOpen* **2015,** *4,* 756–767. doi:10.1002/open.201500162
49. Sommer, R.; Exner, T. E.; Titz, A. *PLoS One* **2014,** *9,* e112822. doi:10.1371/journal.pone.0112822
50. Sommer, R.; Wagner, S.; Rox, K.; Varrot, A.; Hauck, D.; Wamhoff, E.-C.; Schreiber, J.; Ryckmans, T.; Brunner, T.; Rademacher, C.; Hartmann, R. W.; Brönstrup, M.; Imberty, A.; Titz, A. *J. Am. Chem. Soc.* **2018,** *140,* 2537–2545. doi:10.1021/jacs.7b11133
51. Garland, M.; Loscher, S.; Bogyo, M. *Chem. Rev.* **2017,** *117,* 4422–4461. doi:10.1021/acs.chemrev.6b00676
52. Wilcox, M. H.; Gerding, D. N.; Poxton, I. R.; Kelly, C.; Nathan, R.; Birch, T.; Cornely, O. A.; Rahav, G.; Bouza, E.; Lee, C.; Jenkin, G.; Jensen, W.; Kim, Y.-S.; Yoshida, J.; Gabryelski, L.; Pedley, A.; Eves, K.; Tipping, R.; Guris, D.; Kartsonis, N.; Dorr, M.-B. *N. Engl. J. Med.* **2017,** *376,* 305–317. doi:10.1056/NEJMoa1602615
53. Greig, S. L. *Drugs* **2016,** *76,* 823–830. doi:10.1007/s40265-016-0577-0
54. Migone, T.-S.; Subramanian, G. M.; Zhong, J.; Healey, L. M.; Corey, A.; Devalaraja, M.; Lo, L.; Ullrich, S.; Zimmerman, J.; Chen, A.; Lewis, M.; Meister, G.; Gillum, K.; Sanford, D.; Mott, J.; Bolmer, S. D. *N. Engl. J. Med.* **2009,** *361,* 135–144. doi:10.1056/NEJMoa0810603
55. Dal Peraro, M.; van der Goot, F. G. *Nat. Rev. Microbiol.* **2016,** *14,* 77–92. doi:10.1038/nrmicro.2015.3
56. Yu, X.-Q.; Robbie, G. J.; Wu, Y.; Esser, M. T.; Jensen, K.; Schwartz, H. I.; Bellamy, T.; Hernandez-Illas, M.; Jafri, H. S. *Antimicrob. Agents Chemother.* **2017,** *61,* e01020–16. doi:10.1128/AAC.01020-16
57. Karginov, V. A. *Curr. Opin. Pharmacol.* **2013,** *13,* 717–725. doi:10.1016/j.coph.2013.08.007
58. Ragle, B. E.; Karginov, V. A.; Bubeck Wardenburg, J. *Antimicrob. Agents Chemother.* **2010,** *54,* 298–304. doi:10.1128/AAC.00973-09
59. Mulvey, G. L.; Marcato, P.; Kitov, P. I.; Sadowska, J.; Bundle, D. R.; Armstrong, G. D. *J. Infect. Dis.* **2003,** *187,* 640–649. doi:10.1086/373996
60. Kitov, P. I.; Sadowska, J. M.; Mulvey, G.; Armstrong, G. D.; Ling, H.; Pannu, N. S.; Read, R. J.; Bundle, D. R. *Nature* **2000,** *403,* 669–672. doi:10.1038/35001095
61. Nishikawa, K.; Matsuoka, K.; Kita, E.; Okabe, N.; Mizuguchi, M.; Hino, K.; Miyazawa, S.; Yamasaki, C.; Aoki, J.; Takashima, S.; Yamakawa, Y.; Nishijima, M.; Terunuma, D.; Kuzuhara, H.; Natori, Y. *Proc. Natl. Acad. Sci. U. S. A.* **2002,** *99,* 7669–7674. doi:10.1073/pnas.112058999
62. Zhu, J.; Cai, X.; Harris, T. L.; Gooyit, M.; Wood, M.; Lardy, M.; Janda, K. *Chem. Biol.* **2015,** *22,* 483–491. doi:10.1016/j.chembiol.2015.03.012

63. Kany, A. M.; Sikandar, A.; Yahiaoui, S.; Haupenthal, J.; Walter, J.; Empting, M.; Köhnke, J.; Hartmann, R. W. *ACS Chem. Biol.* **2018,** *13,* 2449–2455. doi:10.1021/acschembio.8b00257
64. Schönauer, E.; Kany, A. M.; Haupenthal, J.; Hüsecken, K.; Hoppe, I. J.; Voos, K.; Yahiaoui, S.; Elsässer, B.; Ducho, C.; Brandstetter, H.; Hartmann, R. W. *J. Am. Chem. Soc.* **2017,** *139,* 12696–12703. doi:10.1021/jacs.7b06935
65. Green, E. R.; Mecsas, J. Bacterial Secretion Systems: An Overview. In *Virulence Mechanisms of Bacterial Pathogens,* 5th ed.; Kudva, I.; Cornick, N.; Plummer, P.; Zhang, Q.; Nicholson, T.; Bannantine, J.; Bellaire, B., Eds.; ASM Press: Washington, DC, 2016; pp 215–239. doi:10.1128/microbiolspec.VMBF-0012-2015
66. Duncan, M. C.; Linington, R. G.; Auerbuch, V. *Antimicrob. Agents Chemother.* **2012,** *56,* 5433–5441. doi:10.1128/AAC.00975-12
67. Anantharajah, A.; Mingeot-Leclercq, M.-P.; Van Bambeke, F. *Trends Pharmacol. Sci.* **2016,** *37,* 734–749. doi:10.1016/j.tips.2016.05.011
68. Milla, C. E.; Chmiel, J. F.; Accurso, F. J.; VanDevanter, D. R.; Konstan, M. W.; Yarranton, G.; Geller, D. E. *Pediatr. Pulmonol.* **2014,** *49,* 650–658. doi:10.1002/ppul.22890
69. DiGiandomenico, A.; Keller, A. E.; Gao, C.; Rainey, G. J.; Warrener, P.; Camara, M. M.; Bonnell, J.; Fleming, R.; Bezabeh, B.; Dimasi, N.; Sellman, B. R.; Hilliard, J.; Guenther, C. M.; Datta, V.; Zhao, W.; Gao, C.; Yu, X.-Q.; Suzich, J. A.; Stover, C. K. *Sci. Transl. Med.* **2014,** *6,* 262r. doi:10.1126/scitranslmed.3009655
70. Bassler, B. L.; Losick, R. *Cell* **2006,** *125,* 237–246. doi:10.1016/j.cell.2006.04.001
71. Tashiro, Y.; Yawata, Y.; Toyofuku, M.; Uchiyama, H.; Nomura, N. *Microbes Environ.* **2013,** *28,* 13–24. doi:10.1264/jsme2.ME12167
72. Galloway, W. R. J. D.; Hodgkinson, J. T.; Bowden, S. D.; Welch, M.; Spring, D. R. *Chem. Rev.* **2011,** *111,* 28–67. doi:10.1021/cr100109t
73. Geske, G. D.; O'Neill, J. C.; Blackwell, H. E. *Chem. Soc. Rev.* **2008,** *37,* 1432–1447. doi:10.1039/b703021p
74. Köhler, T.; Perron, G. G.; Buckling, A.; van Delden, C. *PLoS Pathog.* **2010,** *6,* e1000883. doi:10.1371/journal.ppat.1000883
75. Storz, M. P.; Maurer, C. K.; Zimmer, C.; Wagner, N.; Brengel, C.; de Jong, J. C.; Lucas, S.; Müsken, M.; Häussler, S.; Steinbach, A.; Hartmann, R. W. *J. Am. Chem. Soc.* **2012,** *134,* 16143–16146. doi:10.1021/ja3072397
76. Lu, C.; Maurer, C. K.; Kirsch, B.; Steinbach, A.; Hartmann, R. W. *Angew. Chem., Int. Ed.* **2014,** *53,* 1109–1112. doi:10.1002/anie.201307547
77. Starkey, M.; Lepine, F.; Maura, D.; Bandyopadhaya, A.; Lesic, B.; He, J.; Kitao, T.; Righi, V.; Milot, S.; Tzika, A.; Rahme, L. *PLoS Pathog.* **2014,** *10,* e1004321. doi:10.1371/journal.ppat.1004321
78. Ilangovan, A.; Fletcher, M.; Rampioni, G.; Pustelny, C.; Rumbaugh, K.; Heeb, S.; Cámara, M.; Truman, A.; Chhabra, S. R.; Emsley, J.; Williams, P. *PLoS Pathog.* **2013,** *9,* e1003508. doi:10.1371/journal.ppat.1003508
79. Kumar, K.; Chopra, S. *J. Antimicrob. Chemother.* **2013,** *68,* 1465–1470. doi:10.1093/jac/dkt045
80. David, M. Z.; Dryden, M.; Gottlieb, T.; Tattevin, P.; Gould, I. M. *Int. J. Antimicrob. Agents* **2017,** *50,* 303–307. doi:10.1016/j.ijantimicag.2017.05.006
81. Waxman, H. A.; Corr, B. Waste from pharmaceutical plants in India and China promotes antibiotic-resistant superbugs. Statnews: Boston, MA, 2016; https://www.statnews.com/2016/10/14/superbugs-antibiotic-resistance-india-china (accessed Sept 19, 2018).
82. Lübbert, C.; Baars, C.; Dayakar, A.; Lippmann, N.; Rodloff, A. C.; Kinzig, M.; Sörgel, F. *Infection* **2017,** *45,* 479–491. doi:10.1007/s15010-017-1007-2

License and Terms

BEILSTEIN JOURNAL OF ORGANIC CHEMISTRY

Lectins of *Mycobacterium tuberculosis* – rarely studied proteins

Katharina Kolbe*[1], Sri Kumar Veleti[1], Norbert Reiling[2,3] and Thisbe K. Lindhorst[4]

Review

Open Access

Address:
[1]Tuberculosis Research Section, Laboratory of Clinical Immunology and Microbiology, National Institute of Allergy and Infectious Diseases, 33 North Drive, Bethesda, 20892, MD, United States, [2]Microbial Interface Biology, Research Center Borstel, Leibniz Lung Center, Parkallee 22, 23845 Borstel, Germany, [3]German Center for Infection Research (DZIF), Borstel Site, 23845 Borstel, Germany and [4]Otto Diels Institute of Organic Chemistry, Christiana Albertina University of Kiel, Otto-Hahn-Platz 3–4, 24118 Kiel, Germany

Email:
Katharina Kolbe* - katharina.kolbe@nih.gov

* Corresponding author

Keywords:
adhesion; carbohydrates; fimbriae; lectins; *Mycobacterium tuberculosis*; pili

Beilstein J. Org. Chem. **2019,** *15,* 1–15.
doi:10.3762/bjoc.15.1

Received: 17 September 2018
Accepted: 29 November 2018
Published: 02 January 2019

This article is part of the Thematic Series "The glycosciences".

Associate Editor: S. Flitsch

Abstract

The importance of bacterial lectins for adhesion, pathogenicity, and biofilm formation is well established for many Gram-positive and Gram-negative bacteria. However, there is very little information available about lectins of the tuberculosis-causing bacterium, *Mycobacterium tuberculosis* (*Mtb*). In this paper we review previous studies on the carbohydrate-binding characteristics of mycobacteria and related *Mtb* proteins, discussing their potential relevance to *Mtb* infection and pathogenesis.

Introduction

More than 135 years after the discovery of *Mycobacterium tuberculosis* (*Mtb*) by Robert Koch [1], tuberculosis (TB) is still one of the world's deadliest communicable diseases [2]. TB is theoretically curable and preventable, especially since effective antibiotics have been available since the 1940s [3-5]. However, the World Health Organization (WHO) reported 1.6 million fatalities worldwide from tuberculosis in 2017, with more than 10 million annual new cases, and an overall estimated global burden of almost 1.7 billion latently infected people [2]. The fight against this primarily pulmonary disease is strongly influenced by localized poverty and the efficiency of regional health care systems, and nowadays is further complicated by the rapidly increasing prevalence of antibiotic-resistant *Mtb* strains [2]. To successfully combat this disease, it is important to improve our understanding of *Mtb* biology and identify new drug targets and anti-*Mtb* strategies.

Mtb bacteria are mainly transmitted by inhalation of aerosolized droplets released from infected patients by coughing. The infection process is initiated by contact between inhaled bacteria and

host cells within the alveolar airspace. The main target cells of *Mtb* bacteria are primarily alveolar macrophages, which internalize the pathogen through phagocytosis [6]. These innate immune cells initiate a number of responses to limit bacterial replication and spread with the ultimate goal of eradicating the pathogen. However, *Mtb* has evolved successful strategies to survive, replicate and persist within macrophages for days, months or even years, including highly-specialized metabolic pathways for nutrient acquisition and stress-responsive processes for protection against the immune system [7-12]. In this regard, invasion of alveolar macrophages is considered as one of the seminal steps in *Mtb* infection. However, within the alveolar space of the lung, epithelial cells are present in far larger numbers than macrophages. The first cells that *Mtb* encounters are therefore most likely alveolar epithelial cells. Previous work has indeed shown that alveolar type II pneumocytes can also become infected with *Mtb* bacteria in vitro and in vivo [13-17]. Furthermore, dendritic cells and neutrophils internalize *Mtb* bacteria and are important key players in the immune response against this pathogen [18-20].

Bacterial invasion of host cells is a complex process, which is initiated by interactions between host and bacterial cell surface structures. As shown in previous studies, host cells can bind to mycobacterial cell wall carbohydrates via a class of surface-localized or secreted proteins known as lectins, and these interactions strongly contribute to bacterial adhesion and uptake, and are also associated with the capability of *Mtb* to survive, replicate, and persist within macrophages [21-25]. Ubiquitous in both eukaryotes and prokaryotes, lectins comprise a subclass of glycan-binding proteins most commonly associated with intercellular binding, cell–cell recognition, intracellular protein trafficking, and toxin activity [26]. Lectins typically possess high carbohydrate ligand specificity, enabling precise control over protein–target contacts and associated downstream processes. Lectins are often easily identified based on the primary amino acid sequence alone, due to the presence of conserved lectin-associated domains (carbohydrate-recognition domains; CRDs) [27]. Well known lectin examples within the innate immune system include the DC-specific intercellular adhesion molecule 3-grabbing nonintegrin (DC-SIGN) [28,29], the dendritic cell-specific C-type lectins (Dectin-1, Dectin-2) [30,31], the macrophage inducible C-type lectin (Mincle) [32-34], the macrophage C-type lectin (MCL) [35,36], and the mannose receptor (MR) [37,38] (Figure 1). Since the importance of host lectins in *Mtb* infection has already been studied and reviewed in detail

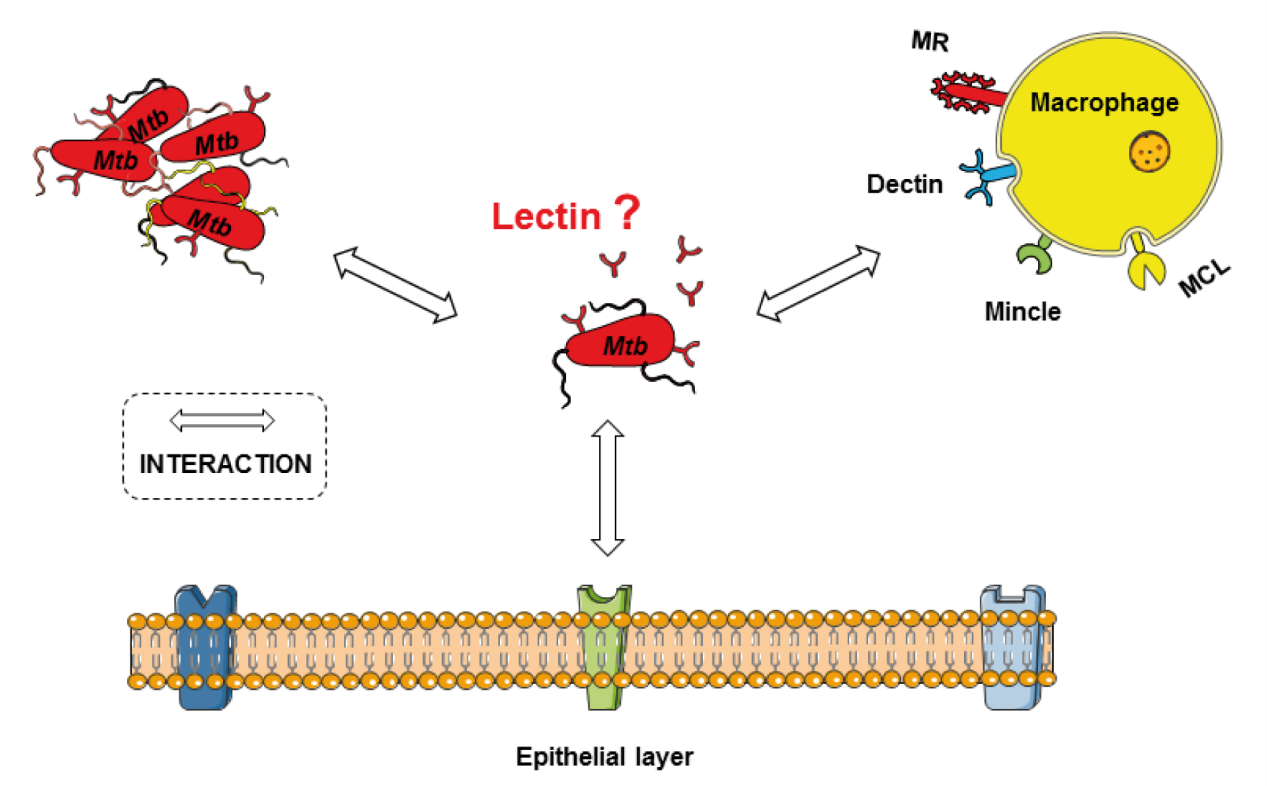


Figure 1: Immune cells (e.g., macrophages) and epithelial cells express lectins on the cell surface (e.g., dendritic cell-specific C-type lectins (Dectin), the macrophage inducible C-type lectin (Mincle), the macrophage C-type lectin (MCL), and the mannose receptor (MR)), which recognize carbohydrates of the *Mtb* cell wall. These proteins contribute to bacterial adhesion and uptake, as well as intracellular survival of the pathogen. The relevance of mycobacterial lectins for host–pathogen interactions has been poorly studied and is the focus of this review.

[21,39], we focus this review on the rarely studied mycobacterial lectins and their roles in recognizing glycosides on the surfaces of host immune and epithelial cells.

Review

Glycosides on the surfaces of mycobacteria and their host cell

Eukaryotic cells exhibit a diverse array of glycoconjugates on their cell surfaces, together known as the glycocalyx (Figure 2). Carbohydrate moieties of the eukaryotic glycocalyx mainly exist in the form of oligosaccharide chains covalently linked to proteins or lipids. The most prevalent oligosaccharide modifications of glycocalyx proteins are *N*-glycans (asparagine-linked) and *O*-glycans (serine- or threonine-linked), while glycosphingolipids are the major subclass of glycosylated lipids in the cell membrane of human cells (Figure 2). While many core elements of glycocalyx oligosaccharides are conserved between host proteins and cell types, for example the invariant *N*-acetyl-D-glucosamine or *N*-acetyl-D-galactosamine residues that attach *N*- or *O*-glycans, respectively, to the peptide side chains, the large variety and possible permutations of "capping" residues (for example D-mannopyranosides, D-galactopyranosides, L-fucopyranosides and sialic acids) that comprise the most terminal, and therefore most accessible for lectin recognition, oligosaccharide regions contribute to a vast diversity of possible glycocalyx structures [40-42]. It is known, for example, that the carbohydrate composition of the glycocalyx is a major determinant of cell type, function, and developmental state, and can have serious pathogenic consequences in the event of dysregulation [43].

The glycoside composition of the mycobacterial cell wall differs strongly from the glycocalyx of eukaryotic cells (Figure 2). The bacterial cell membrane is surrounded by a peptidoglycan layer (PG) consisting of multiple, parallel glycan chains of alternating (1→4)-linked subunits of *N*-acetyl-β-D-

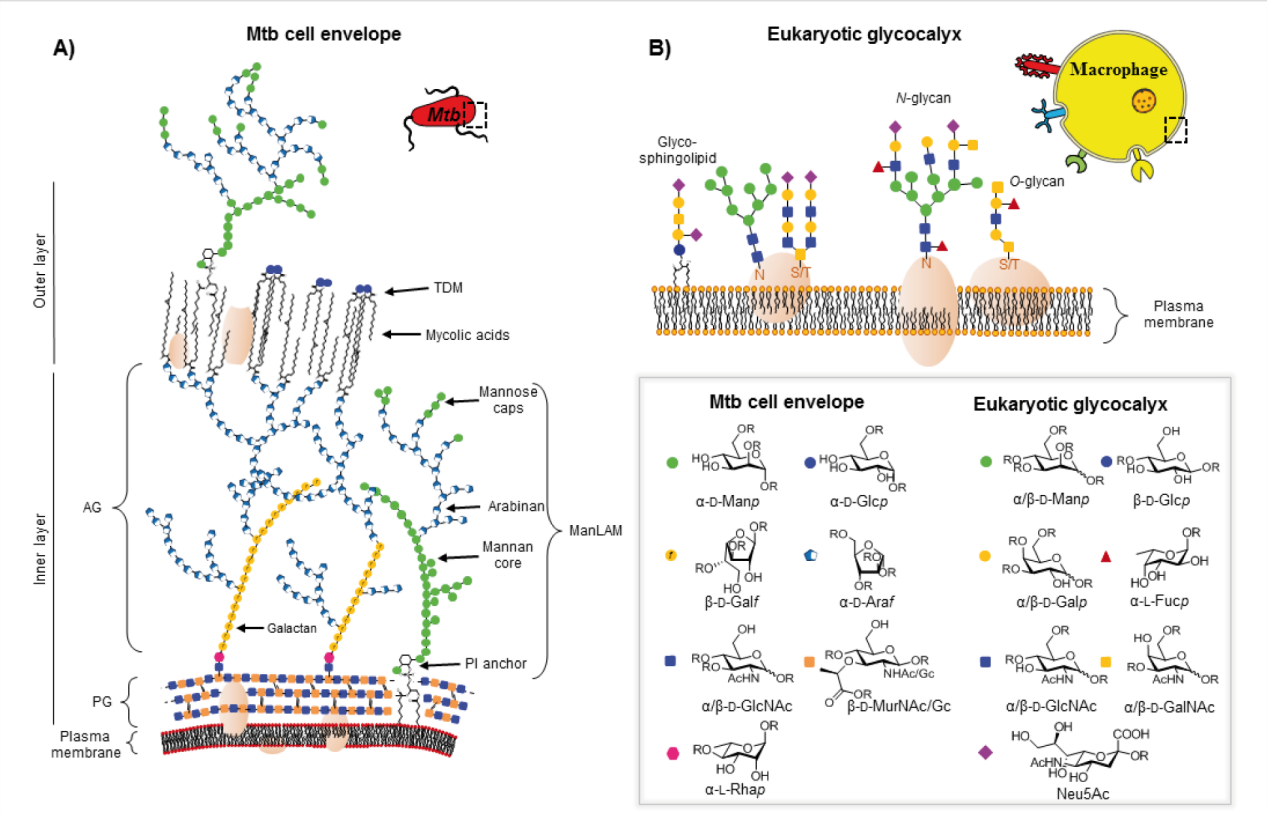


Figure 2: Both mycobacteria and mammalian host cells possess unique subsets of glycosides on their cell surfaces. The main carbohydrates of the *Mtb* cell envelope, a multi-layered structure composed of a mycolyl–arabinogalactan (AG)–peptidoglycan (PG) complex, the lipoglycans lipomannan (LM), and mannosylated lipoarabinomannan (ManLAM), as well as glycolipids, such as trehalose 6,6'-dimycolate (TDM) and trehalose 6-monomycolate (TMM), are α-D-mannopyranosides (α-D-Man*p*), α-D-glucopyranosides (α-D-Glc*p*), α-D-galactofuranosides (α-D-Gal*f*), α-D-arabinofuranosides (α-D-Ara*f*), α-L-rhamnopyranosides (α-L-Rha*p*), *N*-acetyl-α-D-glucosamine (α-D-GlcNAc), *N*-acetyl-β-D-glucosamine (β-D-GlcNAc), and *N*-acetyl- or *N*-glycolyl-β-D-muramic acid (β-D-MurNAc/Gc) residues. The eukaryotic glycocalyx, composed of various glycolipids and glycoproteins, contains D-mannopyranosides (D-Man*p*), D-glucospyranosides (D-Glc*p*), D-galactopyranosides (D-Gal*p*), L-fucopyranosides (L-Fuc*p*), *N*-acetyl-D-glucosamine (D-GlcNAc), *N*-acetyl-D-galactosamine (D-GalNAc), and sialic acid residues, such as *N*-acetyl-D-neuraminic acid (Neu5Ac). While most of the internal glycosides in eukaryotic oligosaccharides are β-linked, terminally localized carbohydrates are often attached via an α-glycosidic bond. (R = H or glycosidic linkage; the figure of the *Mtb* cell wall was originally published in the thesis of K. Kolbe [12] and has been slightly modified for this article).

glucosamine and *N*-acetyl- or *N*-glycolyl-β-D-muramic acid, crosslinked via short conserved oligopeptide stems [44,45]. The PG is covalently attached to the galactan chain of arabinogalactan (AG) by a unique phosphodiester linkage stemming from the 6-OH of a PG muramic acid [46]. AG is the major polysaccharide of the mycobacterial cell envelope and is composed of α-D-arabinosides and β-D-galactosides, both in the relatively uncommon furanose form [47]. The primary hydroxy groups of the terminal arabinofuranoside residues are esterified with mycolic acids forming the basis of the outer lipid layer [48]. The major lipoglycans found in the mycobacterial cell envelope are lipoarabinomannan (LAM) and its precursor lipomannan (LM), both of which consist of a phosphatidyl-*myo*-inositol core structure, glycosylated at the 2-position of *myo*-inositol [49-51]. The oligosaccharide of LM consists exclusively of linear (1→6)-linked and (1→2)-branched α-mannopyranosides [22], while in LAM the mannan structure is elongated by highly (1→2), (1→3) and (1→5)-branched α-D-arabinofuranoside-containing polymers [52]. LAM can further be peripherally modified, also known as ”capping“, the nature of which differs between mycobacterial species. In pathogenic mycobacteria, such as *Mtb*, LAM is capped to various degrees with one to three α-D-mannopyranosides [53], while the fast growing non-pathogenic species *Mycobacterium smegmatis* (*M. smegmatis*) contains inositol phosphate-capped LAM (PILAM) [54]. In addition to lipoglycans, various free, noncovalently associated glycolipids are present in the mycobacterial cell wall, such as the mycolic acid diester trehalose 6,6'-dimycolate (TDM) and its precursor trehalose 6-monomycolate (TMM) [55]. Mycobacteria therefore possess α-D-mannopyranosides, α-D-arabinofuranosides, α-D-glucopyranosides, α-D-galactofuranosides, and their associated oligomeric forms as surface-exposed carbohydrates accessible to extracellular protein recognition. While manno- and glucopyranosides are also present in the eukaryotic glycocalyx, galactofuranosides, arabinofuranosides, and the (1→1)-linked glucose disaccharide trehalose are unique to the mycobacterial cell wall. The occurrence of galactose in the furanose form is restricted to bacteria [56], protozoa [57], and fungi [58], and totally absent in mammals. D-Arabinofuranose can only be found in prokaryotes, for example in Gram-negative bacteria where it is a cytoplasmic intermediate in the biosynthesis of 3-deoxy-D-manno-octulosonic acid (KDO), an essential carbohydrate of the cell wall lipopolysaccharide (LPS) [59]. As a surface-localized carbohydrate, however, D-arabinofuranoside has been exclusively detected in the bacterial suborder of the *Corynebacterineae,* to which the mycobacteria belong [60]. Cell wall-localized D-trehalose is likewise restricted to *Corynebacterineae* [61,62].

In summary, both mycobacteria and mammalian host cells possess unique subsets of surface-exposed carbohydrates, which could function as ligands for putative host- or self-lectins, in processes such as interbacterial aggregation or host–pathogen interactions.

Bacterial lectins

The existence of bacterially-expressed lectins has been known since the first half of the 20th century. Many of these bacterial lectins were originally detected based on their ability to agglutinate red blood cells. Their primary function, however, is to facilitate adhesion of bacteria to host cells or to contribute to interactions among bacteria, which is crucial for the formation of well-organized superstructures such as biofilms. In contrast to eukaryotic lectins, bacterial lectins commonly occur in the form of filamentous protein appendages projecting from their surface, known as fimbriae and pili [63]. Fimbriae are present in high numbers (100–400) on bacterial surfaces, have a diameter of 5–7 nm and can extend hundreds of nanometers in length. Pili, on the other hand, are thicker, longer, and less abundant. Most bacteria encode multiple lectins, each with different carbohydrate specificities [63]. The most intensely studied bacterial lectins are the mannose-specific FimH of type 1 fimbriae and the galabiose-specific PapG of P fimbriae, expressed by *Enterobactericea*, such as *Escherichia coli* (*E. coli*). While type 1-fimbrial expression of *E. coli* is associated with urinary tract infections, the presence of P fimbriae is connected to colonization of the kidney [64,65]. Inhibition of carbohydrate–lectin interactions by antiadhesive drugs is an emerging anti-infective therapeutic approach, particularly in light of increasing rates of bacterial resistance to traditional antibiotics. α-D-Mannosides containing aromatic aglycons, which act as FimH antagonists, for example, have been successfully used to significantly reduce the severity of *E. coli* infections of the urinary tract in mice [66]. Furthermore, preliminary clinical trials with D-mannose indicate promising effects of this monosaccharide on controlling urinary tract infections in humans, presumably through interference with lectin-associated pathogen–host adhesion [67,68]. Besides facilitation of adhesion, some bacterial lectins are also known to act as toxins. The secreted pertussis toxin, for example, is a lectin and an important virulence factor of *Bordetella pertussis* [69-71], the bacterial pathogen responsible for the respiratory disease pertussis, or whooping cough. While no reports exist to date, inhibiting the adhesion of the pertussis toxin to host–cell surface carbohydrates using carbohydrate ligand mimics might permit reduction of the pathogenicity of the toxin and thereby severity of disease.

Mycobacterial lectins

The first experimental evidence of the existence of mycobacterial lectins was described in 1989, when Kundu et al. isolated a 12–14 kDa protein with lectin properties (subsequently named “mycotin’) from the culture supernatant of non-pathogenic

M. smegmatis. This protein was able to agglutinate human A, B and O erythrocytes [72], and the detected hemagglutination could be inhibited by different carbohydrates. The polysaccharide arabinogalactan isolated from *M. smegmatis*, composed of α-D-arabinofuranosides and β-D-galactofuranosides, as well as the monosaccharide D-arabinose, were both found to reverse agglutination, while the α-L-arabinofuranoside-, β-L-arabinopyranoside-, and β-D-galactopyranoside-containing larch wood arabinogalactan and the corresponding L-arabinose monosaccharide were ineffective. Furthermore, the yeast polysaccharide mannan, composed of linear α(1→6)-linked, and α(1→2)- and α(1→3)-branched mannopyranosides, and the glycoside *p*-nitrophenyl α-D-mannopyranoside showed even higher hemagglutination inhibitory potency. These initial experiments let to the assumptions that mycotin is a secreted D-arabinoside- and α-D-mannopyranoside-binding lectin [72]. The relatively high glycoside concentrations (in the mM range) used in this hemagglutination inhibition assays; however, indicate that the tested mono- and polysaccharides are not the optimal or native ligands for this lectin. Other mycobacteria, like *Mtb*, have since been found to contain molecules immunologically related to mycotin on their cell surface [73]. Furthermore, adhesion of *Mtb* to mouse peritoneal macrophages was inhibited using antimycotin antibodies, which led to the assumption that mycotin-like molecules are involved in the interaction of *Mtb* with macrophages and might play a role in *Mtb* infections [73]. However, the 35 kDa cell wall-localized mycotin-like protein identified in *Mtb* in this study was not further characterized and it is still unclear where it is encoded in the bacterial genome. Cell surface-localized mycobacterial lectins and their corresponding ligands have been further investigated using cellular aggregation assays [12,74]. Mycobacteria are known to form large clumps, especially in stationary liquid culture, and it is postulated that lectin–glycan interactions may be at least partially responsible for this aggregation. Anton et al. identified several monosaccharides able to disperse mycobacterial clumps and inhibit bacterial cellular aggregation when added to pure cultures, including D-arabinose (both *M. smegmatis* and *Mtb*), D-xylose, inositol, and D-glucose (*M. smegmatis* only). The impact of D-glucose on *M. smegmatis* aggregation was studied in more detail, where an inhibitory effect of methyl β-D-glucoside, but not methyl α-D-glucoside, was observed [74]. However, the related lectins that mediate self-aggregation have not been isolated or further analyzed to date.

These preliminary findings suggest that *Mtb* has the capacity to express a D-arabinose-specific lectin involved in aggregation processes, and a mycotin-like protein important for adhesion of mycobacteria to macrophages (Table 1).

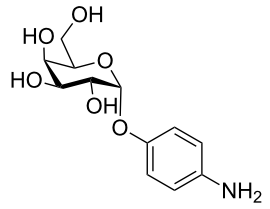
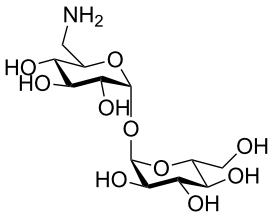
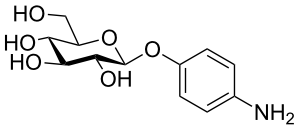
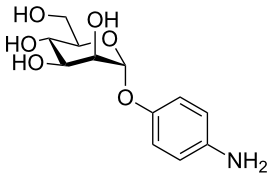
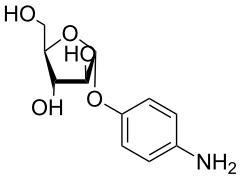
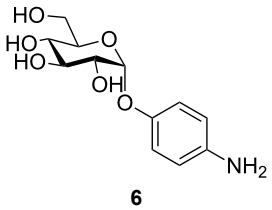
More recently, microtiter plate-based adherence assays were used to further support the carbohydrate-dependent adhesion characteristics of *Mtb* [12]. The author observed stronger adhesion of *Mtb* H37Rv in wells functionalized with α-D-galactopyranoside **1**, or the Actinobacteria-specific cell wall disaccharide D-trehalose (**2**), compared to β-D-glucopyranoside **3** or α-D-mannopyranoside **4**. In contrast to the results described by V. Anton et al. [74], the bacteria did not adhere to surfaces functionalized with the D-arabinoside derivative **5**. However, in the synthetic structure **5**, arabinose is fixed in the furanose form, while the unmodified D-arabinose, applied by V. Anton et al., is mainly present in the pyranose form. Thus, the results might not be contradictory, but rather suggest that an arabinopyranose-, but not arabinofuranoside-binding lectin might be present in the mycobacterial cell envelope. *M. bovis* BCG bacteria showed divergent and much broader adhesion characteristics with strong binding to α-D-galactopyranoside **1**, trehalose (**2**), β-D-glucopyranoside **3**, α-D-mannopyranoside **4** and D-arabinofuranoside **5**, but not α-D-glucopyranoside **6** (Table 2) [12].

In general, stronger adhesion was detected for carbohydrate derivatives with aromatic aglycon moieties compared to aliphatic aglycons (structures not shown) [12]. These results are similar to previous observations with other lectins. Adhesion and inhibition studies with the fimbrial lectin FimH of *E. coli* bacteria, for example, also revealed higher affinities of glycosides

Table 1: Identification and characterization of the lectin mycotin and inhibition studies of bacterial agglutination have provided initial insights into carbohydrate specificity, sub-cellular location and functions of putative mycobacterial lectins.

Mycobacterial species	Carbohydrate specificity	Potential location of the lectin	Potential lectin function
Mtb	unknown (maybe mannosides)	cell wall	interaction with macrophages
	D-arabinose	cell surface	agglutination
M. smegmatis	D-arabinoside, α-D-mannopyranoside	supernatant	interaction with macrophages
	D-arabinose, D-xylose, inositol, methyl β-D-glucoside	cell surface	agglutination

Table 2: In the thesis of K. Kolbe various sugar derivatives were immobilized in 96 well microtiter plates via an amino group. Bacterial adhesion was studied using GFP-expressing mycobacterial strains. Stronger fluorescence intensities detected after incubation and washing steps was correlated to a higher amount of bacteria, and therefore stronger adhesion. The experiments verified carbohydrate-dependent adhesion characteristics of *M. bovis* BCG bacteria and *Mtb* H37Rv bacteria. The carbohydrate binding specificity strongly varied between the two investigated mycobacterial species. (+++: very strong adhesion, ++: strong adhesion, +: adhesion, −: no adhesion).

Immobilized carbohydrate derivatives	Adhesion *Mtb*	Adhesion BCG
 **1**	+++	+
 **2**	+++	++
 **3**	−	++
 **4**	−	++
 **5**	−	++
 **6**	−	−

carrying an aromatic aglycon compared to derivatives with aliphatic aglycon portions. This finding can be attributed to π-interactions with tyrosine residues located at the rim of the carbohydrate binding pocket (Figure 3) [64,75].

Importantly, adhesion of *Mtb* was observed to both mycobacterial- and host-specific carbohydrates indicating that cell surface-localized mycobacterial lectins may be involved in mediating both inter-bacterial and bacteria–host interactions. Furthermore,

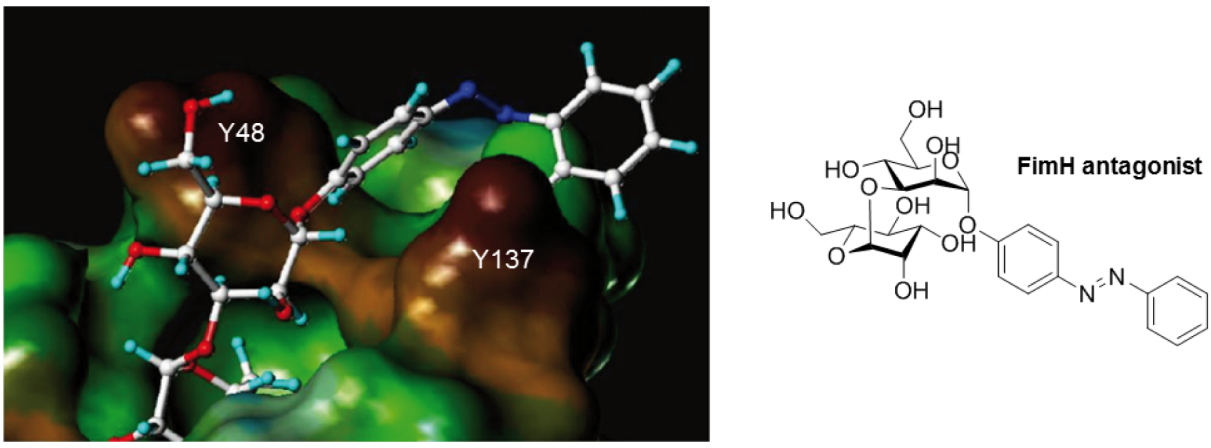


Figure 3: Structure of FimH CRD with a docked azobenzene mannobioside showing the aromatic aglycon and the tyrosine residues, Y48 and Y137, of the protein in close proximity. The figure is a slightly modified version of an image originally published by V. Chandrasekaran et al. [76].

the carbohydrate specificity of *Mtb* adhesion appears to differ significantly from BCG, suggesting that lectins may constitute a contributing factor to the differences in human pathogenicity observed between the two species [12].

The presence of mycobacterial lectins was further supported by Abhinav et al. using in silico genome analysis. A bioinformatics homology-based search of lectin-encoding gene regions in 30 fully or partially sequenced mycobacterial genomes identified 94 potential glycan-binding proteins. The number of detected potential lectins, which ranged from one to six per strain, and their phylogenetic association to established lectin families strongly varied depending on the mycobacterial species in question [77]. These results are consistent with the varying carbohydrate-binding characteristics observed between different mycobacterial species, as described above [12,74]. While three potential glycan-binding proteins were identified in the *Mtb* (H37Rv) genome in this study (Table 3) [77], Singh et al., using a different suite of bioinformatic tools, identified eleven, of which nine were annotated as potential lectins [78]. However, most of the proteins encoded by these genes have yet to be biochemically characterized, precluding further functional predictions. Exceptions are the secreted 13 kDa large lectin from *Mtb*, sMTL-13 [79,80], and the heparin-binding hemagglutinin (HBHA) [81-85], which have been previously studied in detail (see below). We subsequently discuss the association of the nine putative *Mtb* lectins identified by Singh et al. and Abhinav et al. with established lectin families [77,78], such as agglutinin-like sequences (ALS), mannose-sensitive hemagglutinin (MSHA), C-type lectins, and R-type lectins. Furthermore, the filamentous hemagglutinin (FHA) and the heparin-binding hemagglutinin (HBHA) as glycosaminoglycan-binding protein families are also discussed (Table 3, Figure 4).

Agglutinin-like sequences

Based on bioinformatic analysis, the two *Mtb* gene products of Rv1753 and Rv2082 were reported to have 27% and 25% amino acid sequence similarity to the ALS1 gene from *Candida albicans*, which encodes the candida adhesin [78]. This lectin is cell surface-localized and mediates adherence of the fungus to endothelial and epithelial cells [86,87]. Fucose-containing glycans were detected as potential carbohydrate ligands for the

Table 3: Eleven *Mtb* genes were predicted based on in silico genome analysis to encode for glycan-binding proteins, as reported by Singh et al. The three genes in bold (Rv2075, Rv1419, Rv0475) were also identified by Abhinav et al., using different bioinformatics methods. Only two of the encoded proteins have been biochemically characterized to date.

Lectin family	Gene ID	*Mtb* protein
agglutinin like sequences (ALS)	Rv2082, Rv1753	–
mannose sensitive hemagglutinin (MSHA)	Rv2813, Rv3659	–
C-type lectin	**Rv2075**	–
R-type lectin	**Rv1419**	**sMTL-13**
filamentous hemagglutinin (FHA)	Rv0355, Rv1917, Rv3343, Rv3350	–
heparin-binding hemagglutinin (HBHA)	**Rv0475**	**HBHA**

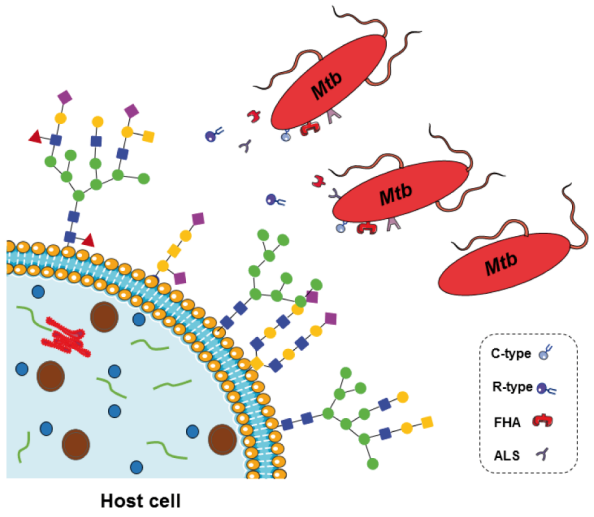


Figure 4: Computer-based genome analysis supports the existence of mycobacterial glycan-binding proteins, which can be associated with known lectin and glycosaminoglycan-binding protein families, including agglutinin-like sequences (ALS), mannose-sensitive hemagglutinin (MSHA), C-type lectins, and R-type lectins, filamentous hemagglutinin (FHA) and heparin-binding hemagglutinin (HBHA). However, hitherto there is only limited information concerning expression, cell localization and function of mycobacterial lectins and glycosaminoglycan-binding proteins.

ALS1 protein [88]. Intriguingly, Rv1753 is described as essential for in vitro growth of *Mtb*, as detected by transposon mutagenesis studies [89,90]. However, no further biochemical or genetic data are available for either Rv1753 or Rv2082, and an associated ALS-like lectin function is only speculation.

Mannose-sensitive hemagglutinin

Two *Mtb* gene products, encoded by Rv2813 and Rv3659, were classified as MSHA-like proteins, with the highest amino acid similarity directed to MshM (41%) and MshE (26%), respectively, of the marine bacterium *Pseudomonas haloplanktis* [78]. These genes encode for proteins involved in assembly of type IV pili (T4P) [91]. Since bacterial lectins are often located at the terminal ends of pili or fimbriae, this homology is of potential interest as it indicates that *Mtb* might express carbohydrate-binding pili on the cell surface (discussed further below).

C-Type lectin

C-Type lectins are one of the largest and most diverse lectin families, including the *Mtb*-recognizing eukaryotic host immune receptors DC-SIGN, Dectin-1/2, Mincle, MCL, and MR, mentioned before. These lectins bind carbohydrates in a calcium-dependent manner. The ligand specificity is highly diverse, including fucosides, mannosides, glucosides, *N*-acetylglucosamines, galactosides, and *N*-acetylgalactosamines. While some of the C-type lectins are known to be secreted, others are membrane-associated proteins. They often oligomerize into homodimers, homotrimers, and higher-ordered oligomers, which increases their avidity for multivalent ligands. C-Type lectins play key roles in cell–cell interactions, such as host–pathogen interactions, and phagocytosis [92]. The *Mtb* gene product of Rv2075c shows partial amino acid sequence similarity to mannose-specific C-type lectins from *Caenorhabditis elegans*, *Mus musculus*, and *Homo sapiens* (see Figure 5 for partial secondary structure prediction and alignment with the human C-type mannose receptor 2) [77,78], and is predicted to be localized to the outer membrane [93]. While Rv2075c orthologues have been identified in all tested *Mtb* strains (*Mtb* H37Ra, *Mtb* H37Rv, *Mtb* KZN 1435, *Mtb* KZN 4207, *Mtb* CDC1551), no homologous gene was identified in the *Mycobacterium africanum* strain GM041182 [77]. TB in humans is primarily caused by *Mtb*, but can also be a consequence of infection with *Mycobacterium africum,* which is currently limited to West Africa [94]. Thus, the potential C-type lectin of *Mtb* might not be essential for a typical TB infection in humans. However, cell localization and function need to be investigated in further detail.

R-Type lectin

R-Type lectins are classified as lectins containing a carbohydrate-recognition domain similar to the CRD in ricin, a toxin of the poisonous plant *Ricinus communis*. R-type lectins have been detected in plants, animals, and bacteria. Plant R-type lectins often contain a separate subunit functioning as a toxin. Furthermore, ricin-type lectin domains have been found in glycosyltransferases as well as in bacterial hydrolases [95]. The *Mtb* gene product of Rv1419 shows 41% amino acid sequence similarity to R-type lectins and encodes the *Mtb* protein sMTL-13 (see Figure 5 for secondary structure prediction and alignment with the ricin B-like lectin from *Streptomyces olivaceoviridis*) [77,78]. This secreted protein was crystallized in 2010 by Patra et al., however, a three-dimensional structure has yet to be resolved [79]. Recently Nogueira et al. detected high titers of IgG antibodies against sMTL-13 in sera from TB patients, a response found to be diminished following successful antituberculosis therapy [80]. The results underline that mycobacterial lectins are expressed in vivo and might be important for *Mtb* infections. Furthermore, anti-sMTL-13 antibodies could serve as a biomarker of disease treatment progression. The exact function of sMTL-13 and its ligand specificity are, however, still unknown. As described before some R-type lectins exhibit toxin activity. Until recently *Mtb* was regarded as a bacteria that does not express toxins [96-98]. In 2014 Danilchanka et al. challenged this paradigm by discovering that the secreted C-terminal domain of the outer membrane channel protein CpnT acts as a toxin [99]. Thus, it might be conceivable that certain *Mtb* lectins could also have toxin function.

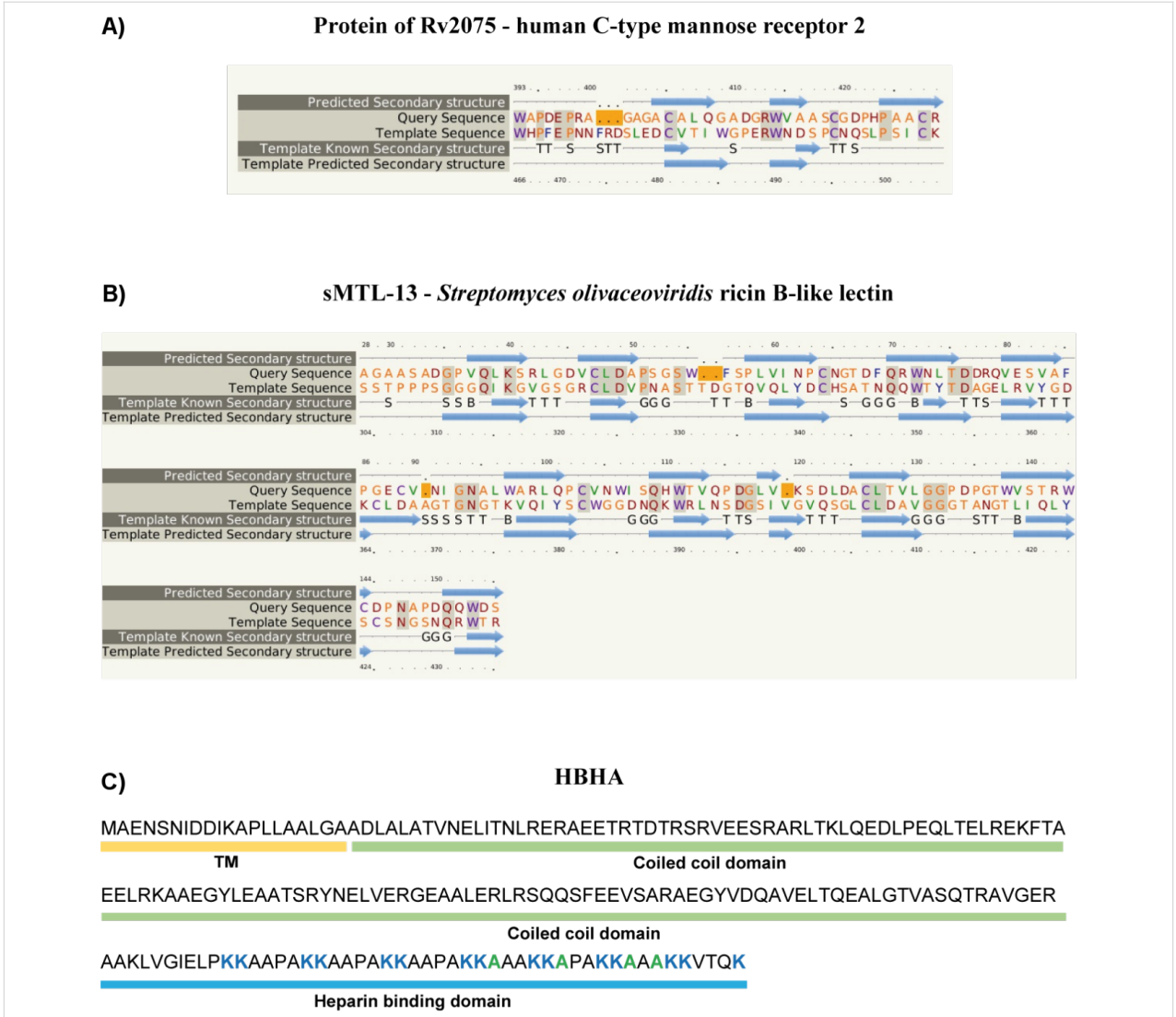


Figure 5: Amino acid sequence and secondary structure alignments of *Mtb* proteins encoded by Rv1419 and Rv2075 to known proteins were determined using Phyre[2]. A) The amino acid sequence (aa393-428) of Rv2075 showed with 97% confidence sequence similarity to the C-type lectin domain of the human C-type mannose receptor 2, the sequence identity was 28%. Identical amino acids are highlighted in grey, amino acids with a small/polar side chain: orange, hydrophobic side chain: green, charged side chain: red, aromatic amino acids and cysteine: violet. β-Sheets of the secondary structure are shown as blue arrows. B) sMTL-13 (aa28-155), encoded by Rv1419, showed with 100% confidence sequence similarity to a ricin B-like lectin of *Streptomyces olivaceoviridis*, the sequence identity was 22%. C) Known domain structure of HBHA (Rv0475): Transmembrane domain (TM), coiled coil domain, and heparin binding domain. Amino acids involved in heparin bind are colored in blue (lysine) and green (alanine).

Filamentous hemagglutinin

One of the most well-characterized FHAs is expressed by *Bordetella pertussis*. The FHA of this pathogen is both surface-exposed and secreted. It functions as an adhesin, where it recognizes and binds to sulfated glycolipids on epithelial host–cell surfaces. The ability of the bacteria to attach to and infect the epithelium of the upper respiratory tract is essential in the pathogenesis of the pertussis organism, underlining the crucial role of this lectin in bacterial physiology [100]. FHA also promotes the formation of biofilms by mediating cell–substrate and interbacterial adhesions [101]. Singh et al. reported that the products of four genes of the *Mtb* strain H37Rv: Rv0355, Rv1917, Rv3343, and Rv3350, show varying levels of amino acid sequence similarities to FHA of *Bordetella pertussis* [78]. However, there is no reported biochemical evidence to date of similar lectin functions for any of these proteins.

Heparin-binding hemaglutanin

The HBHA encoded by Rv0475 is the most well-characterized glycan-binding protein in *Mtb*. Using biophysical and biochemical methods the domain structure of HBHA has been determined, and includes a canonical lysine-rich C-terminal heparin binding domain (see Figure 5) [83,102-105], which has been shown to bind sulfated glycoconjugates like heparin, facili-

tating the adhesion of mycobacteria to epithelial cells, but not to macrophages [81-85]. Furthermore, this transmembrane protein has been associated with mycobacterial aggregation [82,85]. BALB/c mice infected with either wild-type or HBHA-deficient *Mtb* displayed equivalent bacterial lung colonization, but the HBHA-deficient mutant showed reduced dissemination to other regions of the body relative to wild type, suggesting that HBHA plays an important role in extrapulmonary spread [84]. It has also been shown that antibodies directed against HBHA can limit adhesion of mycobacteria to epithelial cells in vitro and in vivo [80,83]. Interestingly, anti-HBHA antibodies have been detected in the sera of TB patients [82]. Thus, a humoral immune response to HBHA might also be connected to a reduced dissemination of *Mtb* from human lungs.

Apart from the potential lectins predicted by in silico genome analysis, a C-type lectin-like carbohydrate binding domain was recently identified to be present in the arabinofuranosyltransferase EmbC (Rv3793), which is involved in the LAM biosynthesis of the *Mtb* cell wall [106]. However, the known function of this protein in arabinogalactan biosynthesis suggests the lectin-like domain to be more associated with catalysis and/or substrate recognition, rather than in a canonical interbacterial or host–pathogen lectin–carbohydrate adhesion role.

As described above, only limited data exists concerning expression, subcellular localization and physiological functions of mycobacterial lectins and glycosaminoglycan-binding proteins to date. However, agglutination-inhibition and adhesion assays, genome analyses, and immunological studies have provided the first indications that glycan-binding proteins might be important mediators of TB infections and *Mtb* pathogenesis. Detection of appendages on the mycobacterial surface, as extensively reviewed by Ramsugit et al. [107], further supports the possible existence of carbohydrate-binding proteins on the cell surface of *Mtb*, since bacterial lectins are often located at the terminal end of fimbriae or pili.

Mycobacterial pili

Mycobacteria have traditionally been regarded as a non-piliated genus; however, recently, studies using transmission electron microscopy (TEM) and atomic force microscopy (AFP) have identified long appendages on the surfaces of *M. smegmatis* and *Mtb*, which could be identified as pili [108-110]. Two different pili types were detected for *Mtb* bacteria (Figure 6). Interestingly, type IV pili are expressed by broth-grown *Mtb*, while curli-like pili are mainly produced by bacilli cultured on solid media [108,111].

Curli-like pili

Curli pili are classified as coiled, non-branching proteins with a typical β-sheet-rich structure, 4–6 nm wide and with aggregative properties. These cell surface structures are produced by several members of the *Enterobacteriaceae* family [112]. The *Mtb* curli-like pili (MTP) encoded by Rv3312A, although currently disputed [113], are 2–3 nm in diameter, have a similar

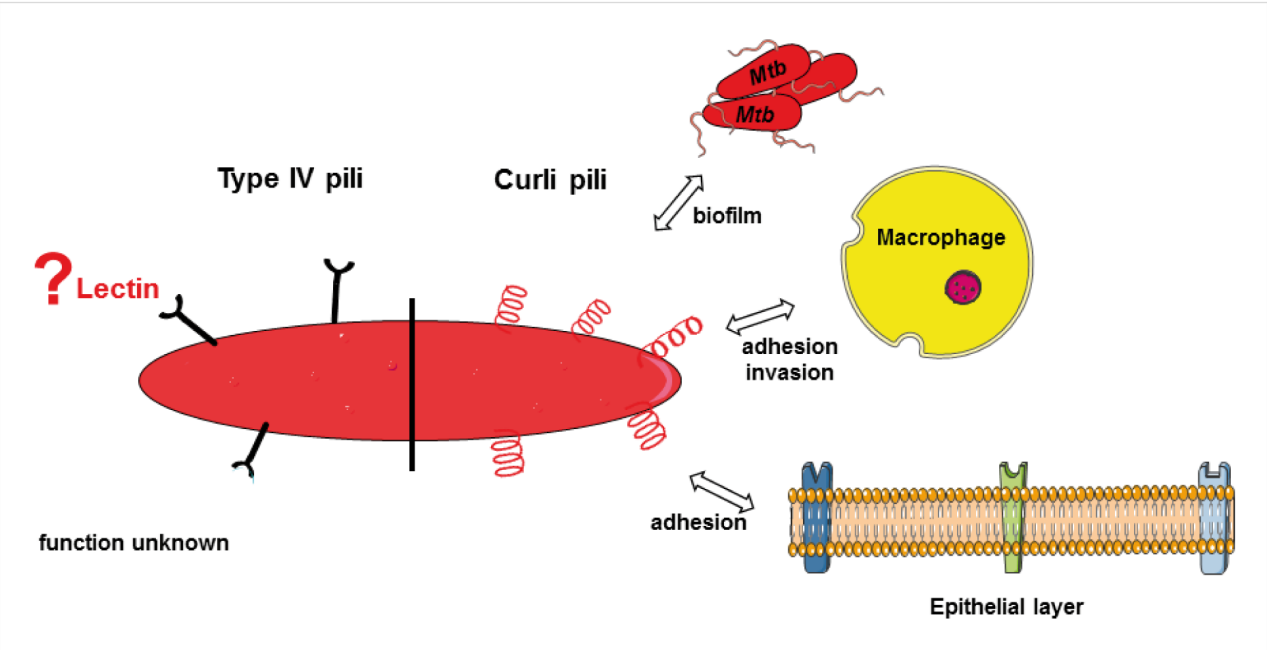


Figure 6: Recently, pili were detected on the cell surface of *Mtb*, which were classified as curli and type IV pili (T4P). While the expression of curli pili is associated with biofilm formation and adhesion to macrophages and epithelial cells, the function of T4P has not yet been examined. Pili often have carbohydrate-binding activity. Whether mycobacterial pili are associated with lectin functions is, however, not known to date.

ultrastructure to curli pili of *E. coli* or *Salmonella* species [114,115], but lack primary sequence homology and the typical β-sheet secondary structure of curli pili from these latter species [108,116]. The *mtp* gene is present in all strains of the *Mtb* complex (MTBC), but absent in non-tuberculosis mycobacteria and other respiratory pathogens [117]. IgG antibodies have been detected in sera of TB patients indicating that MTP are produced during human TB infections [108,118]. Ramsugit et al. studied the adhesive characteristics of MTP using an MTP-deficient (*mtp*-null mutant) strain of *Mtb* and an MTP-overexpressing complemented strain. It was shown that MTP is associated with *Mtb* aggregation and biofilm formation in vitro [116]. The importance of these interactions in patients, however, has yet to be confirmed, as the association of mycobacterial biofilms with bacterial pathogenesis has not yet been conclusively shown in vivo. Besides mediating interactions among mycobacterial cells, MTP has been shown to play a role in *Mtb* adhesion and invasion of A549 pulmonary epithelial cells and THP-1 macrophages [107,119]. Furthermore, an impact of MTP on histopathology in a mouse model of infection has previously been described [113]. Elsewhere, using purified proteins, Alteri et al. detected laminin as a ligand for MTP [108]. While the exact structure recognized by MTP has yet to be determined, laminin is a glycoprotein and so it is conceivable that MTP binds to mono- or oligosaccharide constituents of this protein.

Type IV pili

Type IV pili (T4P) are surface-exposed fibers that mediate many functions in both Gram-positive and Gram-negative bacteria, including motility, adhesion to host cells, biofilm formation, DNA uptake, and protein secretion [120-128]. *Mtb* expresses T4P that appear by electron microscopy as rope-like bundles on the cell surface. Mature T4P are encoded by a seven gene operon, the expression of which is up-regulated during contact with A549 epithelial cells and within macrophages [111,129]. However, their significance in *Mtb* pathogenicity has hitherto not been further investigated. Interestingly, one of the T4P-associated genes is Rv3659, previously identified by in silico genome analysis as coding for a potential mycobacterial lectin (see above) [78]. Although the related protein is most likely involved in pili assembly, it is not inconceivable that T4P have carbohydrate-binding characteristics and are involved in adhesion processes, although this has yet to be proven. The hypothesis is supported by the fact that T4P of other bacteria, for example bundle-forming pili from *E. coli* [130], were shown to have lectin function before.

Conclusion

Lectins are known to play a fundamental role in mediating and regulating numerous biological processes which are initiated by specific carbohydrate recognition. Much effort has been dedicated to the synthesis of specific lectin ligands in order to study and manipulate lectins. On the other hand, intensive work has been spent on the identification and characterization of lectins. Also in microbe–host cell interactions, specific carbohydrate–lectin interactions are the key to adhesion, microbial colonization as well as to infection. For *Mycobacterium tuberculosis* it is known that the macrophage-associated lectins Dectin and Mincle, for example, specifically interact with *Mtb* cell surface glycans, which in many parts differ significantly from the carbohydrates found in eukaryotic cells. However, in spite of the fact that *Mycobacterium tuberculosis* has been the subject of intense research since its discovery in 1882, many details of carbohydrate–protein interactions in *Mtb* infections are still to be discovered. Significant advances have been made in our fundamental understanding of this bacterium in recent years, but several genes annotated in the *Mtb* genome are still classified as coding for “uncharacterized”, “unknown” or “hypothetical” proteins [131-133] including many of the putative *Mtb* lectins and indeed, *Mtb* lectins have been poorly studied in mycobacteria. This account has thus focused on reviewing the available knowledge on *Mtb* lectins, which are a promising field of research with a diagnostic and therapeutic perspective in the field of tuberculosis. Agglutination-inhibition and adhesion assays, as well as immunological studies have indeed provided the first indications that lectins might play an important and as yet underappreciated role in TB infections, underscoring the necessity of more research into these protein families.

List of Abbreviations

AG: arabinogalactan; ALS: agglutinin-like sequences; D-Ara*f*: D-arabinofuranoside; CpnT: outer membrane channel protein; CRD: carbohydrate-recognition domain; DC-SIGN: dendritic cell-specific intercellular adhesion molecule 3-grabbing nonintegrin; dectin: dendritic cell-specific C-type lectin; FHA: filamentous hemagglutinin; L-Fuc*p*: L-fucopyranoside; D-Gal*f*: D-galactofuranoside; D-Gal*p*: D-galactopyranoside; D-GlcNAc: *N*-acetyl-D-glucosamine; D-Glc*p*: D-glucopyranoside; HBHA: heparin-binding hemagglutinin; KDO: 3-deoxy-D-manno-octulosonic acid; LAM: lipoarabinomannan; LM: lipomannan; LPS: lipopolysaccharide; D-Man*p*: D-mannopyranoside; MCL: macrophage C-type lectin; Mincle: macrophage inducible C-type lectin; MR: mannose receptor; MSHA: mannose-sensitive hemagglutinin; *Mtb*: *Mycobacterium tuberculosis*; MTP: *Mtb* curli-like pili; *M. smegmatis*: *Mycobacterium smegmatis*; D-MurNAc/Gc: *N*-acetyl- or *N*-glycolyl-D-muramic acid; Neu5Ac: *N*-acetyl-D-neuraminic acid; PG: peptidoglycan; L-Rha*p*: L-rhamnopyranosides; sMTL: secreted 13 kDa large lectin from *Mtb*; TB: tuberculosis; TDM: trehalose 6,6'-dimycolate; TMM: trehalose 6-monomycolate; T4P: type IV pili.

Acknowledgements

K. K. thanks Dr. Gareth A. Prosser for proofreading.

ORCID® iDs

Norbert Reiling - https://orcid.org/0000-0001-6673-4291
Thisbe K. Lindhorst - https://orcid.org/0000-0001-6788-4224

References

1. Koch, R. *Berl. Klin. Wochenschr.* **1882,** 428–445.
2. *Global tuberculosis report 2018;* World Health Organization, 2018.
3. Schatz, A.; Bugle, E.; Waksman, S. A. *Proc. Soc. Exp. Biol. Med.* **1944,** *55,* 66–69. doi:10.3181/00379727-55-14461
4. Wassersug, J. D. *N. Engl. J. Med.* **1946,** *235,* 220–229. doi:10.1056/nejm194608152350704
5. Marshall, G. *Br. Med. J.* **1949,** *1,* 382–386.
6. Cohen, S. B.; Gern, B. H.; Delahaye, J. L.; Adams, K. N.; Plumlee, C. R.; Winkler, J. K.; Sherman, D. R.; Gerner, M. Y.; Urdahl, K. B. *Cell Host Microbe* **2018,** *24,* 439–446.e4. doi:10.1016/j.chom.2018.08.001
7. Pieters, J. *Cell Host Microbe* **2008,** *3,* 399–407. doi:10.1016/j.chom.2008.05.006
8. Russell, D. G. *Immunol. Rev.* **2011,** *240,* 252–268. doi:10.1111/j.1600-065x.2010.00984.x
9. Flynn, J. L.; Chan, J. *Curr. Opin. Immunol.* **2003,** *15,* 450–455. doi:10.1016/s0952-7915(03)00075-x
10. Gupta, A.; Kaul, A.; Tsolaki, A. G.; Kishore, U.; Bhakta, S. *Immunobiology* **2012,** *217,* 363–374. doi:10.1016/j.imbio.2011.07.008
11. Russell, D. G.; VanderVen, B. C.; Lee, W.; Abramovitch, R. B.; Kim, M.-j.; Homolka, S.; Niemann, S.; Rohde, K. H. *Cell Host Microbe* **2010,** *8,* 68–76. doi:10.1016/j.chom.2010.06.002
12. Kolbe, K. New carbohydrate derivatives as tools to bind and metabolically label strains of the *Mycobacterium tuberculosis* complex. Ph.D. Thesis, Christiana Albertina University of Kiel and Leibniz-Center for Medicine and Biosciences, Germany, 2016.
13. McDonough, K. A.; Kress, Y. *Infect. Immun.* **1995,** *63,* 4802–4811.
14. Bermudez, L. E.; Goodman, J. *Infect. Immun.* **1996,** *64,* 1400–1406.
15. Fine, K. L.; Metcalfe, M. G.; White, E.; Virji, M.; Karls, R. K.; Quinn, F. D. *Cell. Microbiol.* **2012,** *14,* 1402–1414. doi:10.1111/j.1462-5822.2012.01804.x
16. Hernández-Pando, R.; Jeyanathan, M.; Mengistu, G.; Aguilar, D.; Orozco, H.; Harboe, M.; Rook, G. A. W.; Bjune, G. *Lancet* **2000,** *356,* 2133–2138. doi:10.1016/s0140-6736(00)03493-0
17. Harriff, M. J.; Cansler, M. E.; Toren, K. G.; Canfield, E. T.; Kwak, S.; Gold, M. C.; Lewinsohn, D. M. *PLoS One* **2014,** *9,* e97515. doi:10.1371/journal.pone.0097515
18. Dallenga, T.; Repnik, U.; Corleis, B.; Eich, J.; Reimer, R.; Griffiths, G. W.; Schaible, U. E. *Cell Host Microbe* **2017,** *22,* 519–530.e3. doi:10.1016/j.chom.2017.09.003
19. Dallenga, T.; Schaible, U. E. *Pathog. Dis.* **2016,** *74,* ftw012. doi:10.1093/femspd/ftw012
20. Mihret, A. *Virulence* **2012,** *3,* 654–659. doi:10.4161/viru.22586
21. Lugo-Villarino, G.; Hudrisier, D.; Tanne, A.; Neyrolles, O. *Eur. J. Microbiol. Immunol.* **2011,** *1,* 25–40. doi:10.1556/eujmi.1.2011.1.6
22. Briken, V.; Porcelli, S. A.; Besra, G. S.; Kremer, L. *Mol. Microbiol.* **2004,** *53,* 391–403. doi:10.1111/j.1365-2958.2004.04183.x
23. Vergne, I.; Gilleron, M.; Nigou, J. *Front. Cell. Infect. Microbiol.* **2015,** *4,* No. 187. doi:10.3389/fcimb.2014.00187
24. Kleinnijenhuis, J.; Oosting, M.; Joosten, L. A. B.; Netea, M. G.; Van Crevel, R. *Clin. Dev. Immunol.* **2011,** No. 405310. doi:10.1155/2011/405310
25. Hossain, M. M.; Norazmi, M.-N. *BioMed Res. Int.* **2013,** No. 179174. doi:10.1155/2013/179174
26. Taylor, M. E.; Drickamer, K.; Schnaar, R. L.; Etzler, M. E.; Varki, A. Chapter 28: Discovery and Classification of Glycan-Binding Proteins. In *Essentials of Glycobiology,* 3rd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2017.
27. Lis, H.; Sharon, N. *Chem. Rev.* **1998,** *98,* 637–674. doi:10.1021/cr940413g
28. Tailleux, L.; Pham-Thi, N.; Bergeron-Lafaurie, A.; Herrmann, J.-L.; Charles, P.; Schwartz, O.; Scheinmann, P.; Lagrange, P. H.; de Blic, J.; Tazi, A.; Gicquel, B.; Neyrolles, O. *PLoS Med.* **2005,** *2,* e381. doi:10.1371/journal.pmed.0020381
29. Tailleux, L.; Schwartz, O.; Herrmann, J.-L.; Pivert, E.; Jackson, M.; Amara, A.; Legres, L.; Dreher, D.; Nicod, L. P.; Gluckman, J. C.; Lagrange, P. H.; Gicquel, B.; Neyrolles, O. *J. Exp. Med.* **2003,** *197,* 121–127. doi:10.1084/jem.20021468
30. Yadav, M.; Schorey, J. S. *Blood* **2006,** *108,* 3168–3175. doi:10.1182/blood-2006-05-024406
31. Yonekawa, A.; Saijo, S.; Hoshino, Y.; Miyake, Y.; Ishikawa, E.; Suzukawa, M.; Inoue, H.; Tanaka, M.; Yoneyama, M.; Oh-hora, M.; Akashi, K.; Yamasaki, S. *Immunity* **2014,** *41,* 402–413. doi:10.1016/j.immuni.2014.08.005
32. Lang, R. *Front. Immunol.* **2013,** *4,* No. 5. doi:10.3389/fimmu.2013.00005
33. Ishikawa, E.; Ishikawa, T.; Morita, Y. S.; Toyonaga, K.; Yamada, H.; Takeuchi, O.; Kinoshita, T.; Akira, S.; Yoshikai, Y.; Yamasaki, S. *J. Exp. Med.* **2009,** *206,* 2879–2888. doi:10.1084/jem.20091750
34. Matsunaga, I.; Moody, D. B. *J. Exp. Med.* **2009,** *206,* 2865–2868. doi:10.1084/jem.20092533
35. Lobato-Pascual, A.; Saether, P. C.; Fossum, S.; Dissen, E.; Daws, M. R. *Eur. J. Immunol.* **2013,** *43,* 3167–3174. doi:10.1002/eji.201343752
36. Richardson, M. B.; Williams, S. J. *Front. Immunol.* **2014,** *5,* No. 288. doi:10.3389/fimmu.2014.00288
37. Schlesinger, L. S.; Hull, S. R.; Kaufman, T. M. *J. Immunol.* **1994,** *152,* 4070–4079.
38. Schlesinger, L. S. *J. Immunol.* **1993,** *150,* 2920–2930.
39. Goyal, S.; Klassert, T. E.; Slevogt, H. *Med. Microbiol. Immunol.* **2016,** *205,* 513–535. doi:10.1007/s00430-016-0470-1
40. Stanley, P.; Taniguchi, N.; Aebi, M. Chapter 9: *N*-Glycans. In *Essentials of Glycobiology,* 3rd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2017.
41. Brockhausen, I.; Stanley, P. Chapter 10: O-GalNAc Glycans. In *Essentials of Glycobiology,* 3rd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2017.
42. Schnaar, R. L.; Kinoshita, T. Chapter 11: Glycosphingolipids. In *Essentials of Glycobiology,* 3rd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Stanley, P.; Hart, G. W.; Aebi, M.; Darvill, A. G.; Kinoshita, T.; Packer, N. H., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2017.

43. Shurer, C. R.; Colville, M. J.; Gupta, V. K.; Head, S. E.; Kai, F.; Lakins, J. N.; Paszek, M. J. *ACS Biomater. Sci. Eng.* **2018,** *4,* 388–399. doi:10.1021/acsbiomaterials.7b00037
44. Wietzerbin, J.; Das, B. C.; Petit, J. F.; Lederer, E.; Leyh-Bouille, M.; Ghuysen, J. M. *Biochemistry* **1974,** *13,* 3471–3476. doi:10.1021/bi00714a008
45. Schleifer, K. H.; Kandler, O. *Bacteriol. Rev.* **1972,** *36,* 407–477.
46. McNeil, M.; Daffe, M.; Brennan, P. J. *J. Biol. Chem.* **1990,** *265,* 18200–18206.
47. Daffe, M.; Brennan, P. J.; McNeil, M. *J. Biol. Chem.* **1990,** *265,* 6734–6743.
48. McNeil, M.; Daffe, M.; Brennan, P. J. *J. Biol. Chem.* **1991,** *266,* 13217–13223.
49. Hunter, S. W.; Brennan, P. J. *J. Biol. Chem.* **1990,** *265,* 9272–9279.
50. Korduláková, J.; Gilleron, M.; Puzo, G.; Brennan, P. J.; Gicquel, B.; Mikušová, K.; Jackson, M. *J. Biol. Chem.* **2003,** *278,* 36285–36295. doi:10.1074/jbc.m303639200
51. Hsu, F.-F.; Turk, J.; Owens, R. M.; Rhoades, E. R.; Russell, D. G. *J. Am. Soc. Mass Spectrom.* **2007,** *18,* 479–492. doi:10.1016/j.jasms.2006.10.020
52. Nigou, J.; Gilleron, M.; Puzo, G. *Biochimie* **2003,** *85,* 153–166. doi:10.1016/s0300-9084(03)00048-8
53. Chatterjee, D.; Khoo, K.-H. *Cell. Mol. Life Sci.* **2001,** *58,* 2018–2042. doi:10.1007/pl00000834
54. Khoo, K.-H.; Dell, A.; Morris, H. R.; Brennan, P. J.; Chatterjee, D. *J. Biol. Chem.* **1995,** *270,* 12380–12389. doi:10.1074/jbc.270.21.12380
55. Barry, C. E., III; Lee, R. E.; Mdluli, K.; Sampson, A. E.; Schroeder, B. G.; Slayden, R. A.; Yuan, Y. *Prog. Lipid Res.* **1998,** *37,* 143–179. doi:10.1016/s0163-7827(98)00008-3
56. Lindberg, B. *Adv. Carbohydr. Chem. Biochem.* **1990,** *48,* 279–318. doi:10.1016/s0065-2318(08)60033-5
57. de Lederkremer, R. M.; Colli, W. *Glycobiology* **1995,** *5,* 547–552. doi:10.1093/glycob/5.6.547
58. Notermans, S.; Veeneman, G. H.; van Zuylen, C. W. E. M.; Hoogerhout, P.; van Boom, J. H. *Mol. Immunol.* **1988,** *25,* 975–979. doi:10.1016/0161-5890(88)90003-x
59. Levin, D. H.; Racker, E. *J. Biol. Chem.* **1959,** *234,* 2532–2539.
60. McNeil, M.; Wallner, S. J.; Hunter, S. W.; Brennan, P. J. *Carbohydr. Res.* **1987,** *166,* 299–308. doi:10.1016/0008-6215(87)80065-4
61. Elbein, A. D.; Pan, Y. T.; Pastuszak, I.; Carroll, D. *Glycobiology* **2003,** *13,* 17R–27R. doi:10.1093/glycob/cwg047
62. Berg, S.; Kaur, D.; Jackson, M.; Brennan, P. J. *Glycobiology* **2007,** *17,* 35R–56R. doi:10.1093/glycob/cwm010
63. Esko, J. D.; Sharon, D. Chapter 37: Microbial lectins: hemagglutinins, adhesins, and toxins. In *Essentials of Glycobiology,* 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009.
64. Hartmann, M.; Lindhorst, T. K. *Eur. J. Org. Chem.* **2011,** 3583–3609. doi:10.1002/ejoc.201100407
65. Lane, M. C.; Mobley, H. L. T. *Kidney Int.* **2007,** *72,* 19–25. doi:10.1038/sj.ki.5002230
66. Mydock-McGrane, L. K.; Cusumano, Z. T.; Janetka, J. W. *Expert Opin. Ther. Pat.* **2016,** *26,* 175–197. doi:10.1517/13543776.2016.1131266
67. Kranjčec, B.; Papeš, D.; Altarac, S. *World J. Urol.* **2014,** *32,* 79–84. doi:10.1007/s00345-013-1091-6
68. Barclay, J.; Veeratterapillay, R.; Harding, C. *BMJ [Br. Med. J.]* **2017,** *359,* j5193. doi:10.1136/bmj.j5193
69. Sandros, J.; Rozdzinski, E.; Zheng, J.; Cowburn, D.; Tuomanen, E. *Glycoconjugate J.* **1994,** *11,* 501–506. doi:10.1007/bf00731300
70. Stein, P. E.; Boodhoo, A.; Armstrong, G. D.; Heerze, L. D.; Cockle, S. A.; Klein, M. H.; Read, R. J. *Nat. Struct. Mol. Biol.* **1994,** *1,* 591–596. doi:10.1038/nsb0994-591
71. Witvliet, M. H.; Burns, D. L.; Brennan, M. J.; Poolman, J. T.; Manclark, C. R. *Infect. Immun.* **1989,** *57,* 3324–3330.
72. Kundu, M.; Basu, J.; Chakrabarti, P. *FEBS Lett.* **1989,** *256,* 207–210. doi:10.1016/0014-5793(89)81749-1
73. Goswami, S.; Sarkar, S.; Basu, J.; Kundu, M.; Chakrabarti, P. *FEBS Lett.* **1994,** *355,* 183–186. doi:10.1016/0014-5793(94)01203-2
74. Anton, V.; Rougé, P.; Daffé, M. *FEMS Microbiol. Lett.* **1996,** *144,* 167–170. doi:10.1111/j.1574-6968.1996.tb08525.x
75. Knight, S. D.; Bouckaert, J. *Top. Curr. Chem.* **2009,** *288,* 67–107. doi:10.1007/128_2008_13
76. Chandrasekaran, V.; Kolbe, K.; Beiroth, F.; Lindhorst, T. K. *Beilstein J. Org. Chem.* **2013,** *9,* 223–233. doi:10.3762/bjoc.9.26
77. Abhinav, K. V.; Sharma, A.; Vijayan, M. *Proteins: Struct., Funct., Bioinf.* **2013,** *81,* 644–657. doi:10.1002/prot.24219
78. Singh, D. D.; Chandran, D.; Jeyakani, J.; Chandran, N. *Protein Pept. Lett.* **2007,** *14,* 683–691. doi:10.2174/092986607781483813
79. Patra, D.; Srikalaivani, R.; Misra, A.; Singh, D. D.; Selvaraj, M.; Vijayan, M. *Acta Crystallogr., Sect. F: Struct. Biol. Cryst. Commun.* **2010,** *66,* 1662–1665. doi:10.1107/s1744309110042892
80. Nogueira, L.; Cardoso, F. C.; Mattos, A. M.; Bordignon, J.; Figueiredo, C. P.; Dahlstrom, P.; Frota, C. C.; Duarte dos Santos, C. N.; Chalhoub, M.; Cavada, B. S.; Teixeira, H. C.; Oliveira, S. C.; Barral-Netto, M.; Báfica, A. *Eur. J. Immunol.* **2010,** *40,* 744–753. doi:10.1002/eji.200939747
81. Menozzi, F. D.; Rouse, J. H.; Alavi, M.; Laude-Sharp, M.; Muller, J.; Bischoff, R.; Brennan, M. J.; Locht, C. *J. Exp. Med.* **1996,** *184,* 993–1001. doi:10.1084/jem.184.3.993
82. Menozzi, F. D.; Bischoff, R.; Fort, E.; Brennan, M. J.; Locht, C. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95,* 12625–12630. doi:10.1073/pnas.95.21.12625
83. Menozzi, F. D.; Reddy, V. M.; Cayet, D.; Raze, D.; Debrie, A.-S.; Dehouck, M.-P.; Cecchelli, R.; Locht, C. *Microbes Infect.* **2006,** *8,* 1–9. doi:10.1016/j.micinf.2005.03.023
84. Pethe, K.; Alonso, S.; Biet, F.; Delogu, G.; Brennan, M. J.; Locht, C.; Menozzi, F. D. *Nature* **2001,** *412,* 190–194. doi:10.1038/35084083
85. Esposito, C.; Marasco, D.; Delogu, G.; Pedone, E.; Berisio, R. *Biochem. Biophys. Res. Commun.* **2011,** *410,* 339–344. doi:10.1016/j.bbrc.2011.05.159
86. Loza, L.; Fu, Y.; Ibrahim, A. S.; Sheppard, D. C.; Filler, S. G.; Edwards, J. E., Jr. *Yeast* **2004,** *21,* 473–482. doi:10.1002/yea.1111
87. Modrzewska, B.; Kurnatowski, P. *Ann. Parasitol.* **2015,** *61,* 3–9.
88. Donohue, D. S.; Ielasi, F. S.; Goossens, K. V. Y.; Willaert, R. G. *Mol. Microbiol.* **2011,** *80,* 1667–1679. doi:10.1111/j.1365-2958.2011.07676.x
89. Sassetti, C. M.; Boyd, D. H.; Rubin, E. J. *Mol. Microbiol.* **2003,** *48,* 77–84. doi:10.1046/j.1365-2958.2003.03425.x
90. Lamichhane, G.; Zignol, M.; Blades, N. J.; Geiman, D. E.; Dougherty, A.; Grosset, J.; Broman, K. W.; Bishai, W. R. *Proc. Natl. Acad. Sci. U. S. A.* **2003,** *100,* 7213–7218. doi:10.1073/pnas.1231432100
91. Fullner, K. J.; Mekalanos, J. J. *Infect. Immun.* **1999,** *67,* 1393–1404.

92. Cummings, R. D.; McEver, R. P. Chapter 31: C-type Lectins. In *Essentials of Glycobiology,* 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009.
93. Song, H.; Sandie, R.; Wang, Y.; Andrade-Navarro, M. A.; Niederweis, M. *Tuberculosis* **2008,** *88,* 526–544. doi:10.1016/j.tube.2008.02.004
94. de Jong, B. C.; Antonio, M.; Gagneux, S. *PLoS Negl. Trop. Dis.* **2010,** *4,* e744. doi:10.1371/journal.pntd.0000744
95. Cummings, R. D.; Etzler, M. E. Chapter 28: R-Type Lectins. In *Essentials of Glycobiology,* 2nd ed.; Varki, A.; Cummings, R. D.; Esko, J. D.; Freeze, H. H.; Stanley, P.; Bertozzi, C. R.; Hart, G. W.; Etzler, M. E., Eds.; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, 2009.
96. Gordon, S. V.; Bottai, D.; Simeone, R.; Stinear, T. P.; Brosch, R. *BioEssays* **2009,** *31,* 378–388. doi:10.1002/bies.200800191
97. Mukhopadhyay, S.; Nair, S.; Ghosh, S. *FEMS Microbiol. Rev.* **2012,** *36,* 463–485. doi:10.1111/j.1574-6976.2011.00302.x
98. Forrellad, M. A.; Klepp, L. I.; Gioffré, A.; Sabio y García, J.; Morbidoni, H. R.; de la Paz Santangelo, M.; Cataldi, A. A.; Bigi, F. *Virulence* **2013,** *4,* 3–66. doi:10.4161/viru.22329
99. Danilchanka, O.; Sun, J.; Pavlenok, M.; Maueröder, C.; Speer, A.; Siroy, A.; Marrero, J.; Trujillo, C.; Mayhew, D. L.; Doornbos, K. S.; Muñoz, L. E.; Herrmann, M.; Ehrt, S.; Berens, C.; Niederweis, M. *Proc. Natl. Acad. Sci. U. S. A.* **2014,** *111,* 6750–6755. doi:10.1073/pnas.1400136111
100. Hannah, J. H.; Menozzi, F. D.; Renauld, G.; Locht, C.; Brennan, M. J. *Infect. Immun.* **1994,** *62,* 5010–5019.
101. Serra, D. O.; Conover, M. S.; Arnal, L.; Sloan, G. P.; Rodriguez, M. E.; Yantorno, O. M.; Deora, R. *PLoS One* **2011,** *6,* e28811. doi:10.1371/journal.pone.0028811
102. Delogu, G.; Brennan, M. J. *J. Bacteriol.* **1999,** *181,* 7464–7469.
103. Esposito, C.; Carullo, P.; Pedone, E.; Graziano, G.; Del Vecchio, P.; Berisio, R. *FEBS Lett.* **2010,** *584,* 1091–1096. doi:10.1016/j.febslet.2010.02.044
104. Esposito, C.; Pethoukov, M. V.; Svergun, D. I.; Ruggiero, A.; Pedone, C.; Pedone, E.; Berisio, R. *J. Bacteriol.* **2008,** *190,* 4749–4753. doi:10.1128/jb.01988-07
105. Squeglia, F.; Ruggiero, A.; De Simone, A.; Berisio, R. *Protein Sci.* **2018,** *27,* 369–380. doi:10.1002/pro.3346
106. Alderwick, L. J.; Lloyd, G. S.; Ghadbane, H.; May, J. W.; Bhatt, A.; Eggeling, L.; Fütterer, K.; Besra, G. S. *PLoS Pathog.* **2011,** *7,* e1001299. doi:10.1371/journal.ppat.1001299
107. Ramsugit, S.; Pillay, M. *Arch. Microbiol.* **2015,** *197,* 737–744. doi:10.1007/s00203-015-1117-0
108. Alteri, C. J.; Xicohténcatl-Cortes, J.; Hess, S.; Caballero-Olín, G.; Girón, J. A.; Friedman, R. L. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104,* 5145–5150. doi:10.1073/pnas.0602304104
109. Velayati, A. A.; Farnia, P.; Masjedi, M. R. *Int. J. Mycobact.* **2012,** *1,* 57–58. doi:10.1016/j.ijmyco.2012.04.002
110. Hosseini, H.; Fooladi, A. A. I.; Arjomandzadegan, M.; Emami, N.; Bornasi, H. *Asian Pac. J. Trop. Med.* **2014,** *7,* S199–S203. doi:10.1016/s1995-7645(14)60232-7
111. Alteri, C. Novel pili of *Mycobacterium tuberculosis*. Ph.D. Thesis, University of Arizona, Tuscon, USA, 2005.
112. Epstein, E. A.; Chapman, M. R. *Cell. Microbiol.* **2008,** *10,* 1413–1420. doi:10.1111/j.1462-5822.2008.01148.x
113. Mann, K. M.; Pride, A. C.; Flentie, K.; Kimmey, J. M.; Weiss, L. A.; Stallings, C. L. *Microbiology (London, U. K.)* **2016,** *162,* 1784–1796. doi:10.1099/mic.0.000368
114. Olsén, A.; Jonsson, A.; Normark, S. *Nature* **1989,** *338,* 652–655. doi:10.1038/338652a0
115. Collinson, S. K.; Emödy, L.; Müller, K. H.; Trust, T. J.; Kay, W. W. *J. Bacteriol.* **1991,** *173,* 4773–4781. doi:10.1128/jb.173.15.4773-4781.1991
116. Ramsugit, S.; Guma, S.; Pillay, B.; Jain, P.; Larsen, M. H.; Danaviah, S.; Pillay, M. *Antonie van Leeuwenhoek* **2013,** *104,* 725–735. doi:10.1007/s10482-013-9981-6
117. Naidoo, N.; Ramsugit, S.; Pillay, M. *Tuberculosis* **2014,** *94,* 338–345. doi:10.1016/j.tube.2014.03.004
118. Naidoo, N.; Pillay, B.; Bubb, M.; Pym, A.; Chiliza, T.; Naidoo, K.; Ndung'u, T.; Kasprowicz, V. O.; Pillay, M. *Tuberculosis* **2018,** *109,* 80–84. doi:10.1016/j.tube.2018.01.007
119. Ramsugit, S.; Pillay, M. *Jpn. J. Infect. Dis.* **2014,** *67,* 476–478. doi:10.7883/yoken.67.476
120. Aas, F. E.; Løvold, C.; Koomey, M. *Mol. Microbiol.* **2002,** *46,* 1441–1450. doi:10.1046/j.1365-2958.2002.03265.x
121. Mattick, J. S. *Annu. Rev. Microbiol.* **2002,** *56,* 289–314. doi:10.1146/annurev.micro.56.012302.160938
122. Kirn, T. J.; Bose, N.; Taylor, R. K. *Mol. Microbiol.* **2003,** *49,* 81–92. doi:10.1046/j.1365-2958.2003.03546.x
123. Burrows, L. L. *Mol. Microbiol.* **2005,** *57,* 878–888. doi:10.1111/j.1365-2958.2005.04703.x
124. Reguera, G.; McCarthy, K. D.; Mehta, T.; Nicoll, J. S.; Tuominen, M. T.; Lovley, D. R. *Nature* **2005,** *435,* 1098–1101. doi:10.1038/nature03661
125. Han, X.; Kennan, R. M.; Parker, D.; Davies, J. K.; Rood, J. I. *J. Bacteriol.* **2007,** *189,* 5022–5033. doi:10.1128/jb.00138-07
126. Burrows, L. L. *Annu. Rev. Microbiol.* **2012,** *66,* 493–520. doi:10.1146/annurev-micro-092611-150055
127. Craig, L.; Pique, M. E.; Tainer, J. A. *Nat. Rev. Microbiol.* **2004,** *2,* 363–378. doi:10.1038/nrmicro885
128. Piepenbrink, K. H.; Sundberg, E. J. *Biochem. Soc. Trans.* **2016,** *44,* 1659–1666. doi:10.1042/bst20160221
129. Danelishvili, L.; Yamazaki, Y.; Selker, J.; Bermudez, L. E. *PLoS One* **2010,** *5,* e10474. doi:10.1371/journal.pone.0010474
130. Hyland, R. M.; Sun, J.; Griener, T. P.; Mulvey, G. L.; Klassen, J. S.; Donnenberg, M. S.; Armstrong, G. D. *Cell. Microbiol.* **2008,** *10,* 177–187. doi:10.1111/j.1462-5822.2007.01028.x
131. Mazandu, G. K.; Mulder, N. J. *Int. J. Mol. Sci.* **2012,** *13,* 7283–7302. doi:10.3390/ijms13067283
132. Modlin, S. J.; Gunasekaran, D.; Zlotnicki, A. M.; Elghraoui, A.; Kuo, N.; Chan, C. K.; Valafar, F. *bioRxiv* **2018**. doi:10.1101/358986
133. Doerks, T.; van Noort, V.; Minguez, P.; Bork, P. *PLoS One* **2012,** *7,* e34302. doi:10.1371/journal.pone.0034302