



NRC Publications Archive Archives des publications du CNRC

High-throughput chemistry toward complex carbohydrates and carbohydrate-like compounds

Randell, Karla D.; Barkley, Angela; Arya, Prabhat

This publication could be one of several versions: author's original, accepted manuscript or the publisher's version. /
La version de cette publication peut être l'une des suivantes : la version prépublication de l'auteur, la version acceptée du manuscrit ou la version de l'éditeur.

Publisher's version / Version de l'éditeur:

Combinatorial Chemistry & High Throughput Screening, 5, 2, pp. 179-193, 2002-05

NRC Publications Record / Notice d'Archives des publications de CNRC:

<https://nrc-publications.canada.ca/eng/view/object/?id=8d43f475-ff36-43d2-80d3-00ad6375618a>

<https://publications-cnrc.canada.ca/fra/voir/objet/?id=8d43f475-ff36-43d2-80d3-00ad6375618a>

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at

<https://nrc-publications.canada.ca/eng/copyright>

READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L'accès à ce site Web et l'utilisation de son contenu sont assujettis aux conditions présentées dans le site

<https://publications-cnrc.canada.ca/fra/droits>

LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D'UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at

PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n'arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.



High-throughput Chemistry toward Complex Carbohydrates and Carbohydrate-like Compounds^a

Karla D. Randell,^b Angela Barkley^c and Prabhat Arya*

Chemical Biology Program, Steacie Institute for Molecular Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario, Canada, K1A 0R6

Abstract: In the area of peptide and nucleic acid chemistry and biology, high-throughput synthesis has played an important role in providing useful small-molecule-based chemical probes in understanding the structure and function relationships. The past several years, there has been a constant rise in interest toward understanding the biological roles and functions of another important class of biomolecules, i.e., carbohydrates and carbohydrate conjugates. Although at early stages, in recent years, several groups have developed high-throughput synthetic methods to obtain complex carbohydrates or carbohydrate-like small-molecules. The present review article summarizes some of these developments.

INTRODUCTION

Natural biopolymers, such as nucleic acids, proteins and glycoconjugates (i.e., glycoproteins and glycolipids) play important roles in fundamental aspects of cellular functions. These biomolecules are capable of storing as well as transmitting biological information that involves intra- and intercellular events. Unlike glycoconjugates, the biological roles and functions of nucleic acids and proteins are relatively well understood and are better appreciated by the scientific community. Nucleic acid and polypeptide-derived natural biopolymers are linear in nature and chemically defined, pure derivatives are readily obtained by automated synthetic methods.

In recent years, efforts are being made towards understanding the biological roles of glycoconjugates (commonly known as glycobiology) in the modulation of protein function, fertilization, chronic inflammation, immune responses and cancer metastasis [1]. It is now well accepted that glycoconjugates present on host cell surfaces provide specific binding sites for the attachment of bacterial and viral pathogens leading to infectious diseases. In addition, it has been demonstrated that the oligosaccharide moieties of several complex glycoconjugates present on tumor cell surfaces have unique structural features. These moieties are attractive targets for developing chemically well-defined, synthetic vaccines for cancer and the design of specific delivery of anticancer drugs on tumor cell surfaces [2].

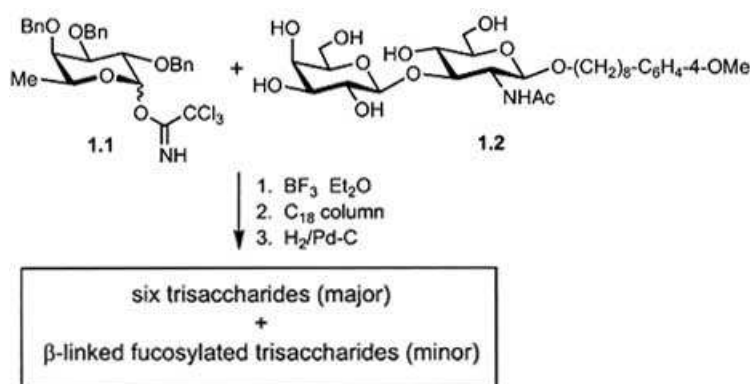
A major obstacle in the field of glycobiology is access to pure, chemically well defined complex carbohydrates and glycoconjugates. Unlike nucleic acids and polypeptides, these are non-linear molecules and the carbohydrate moieties

present tremendous challenges in developing their total syntheses. These polyhydroxy compounds contain an array of monosaccharide units and have a variety of glycosidic linkages between them. Each glycosidic linkage can exist in the α - or β -anomeric configuration. Therefore, carbohydrate syntheses requires many orthogonal protection-deprotection schemes and involve difficult glycosyl coupling reactions [3]. Recently, efforts have been made to develop automated syntheses of complex carbohydrates, but the progress to date has been slow [4]. The solution and solid-phase methodologies toward complex carbohydrates are not general in nature and several carbohydrate-derived coupling reactions do not produce the required products in a stereoselective manner.

The lack of chemical stability and bioavailability associated with complex carbohydrates and glycoconjugates has fuelled parallel developments in synthetic glycoconjugate mimics and in design and synthesis of inhibitors of oligosaccharide functions [5]. It is well established that despite the complexity of the oligosaccharide moieties of glycoconjugates, the terminal sugars (two to four residues) and their conformations are critical for biological activities. This not only reduces the chemical complexity of the synthetic target(s), but also makes possible the use of revolutionary synthetic strategies such as combinatorial chemistry, for rapid access to potential carbohydrate mimics.

To meet the growing demand for economical synthesis of large numbers of diverse chemical compounds in a relatively short time, solid-phase synthesis and combinatorial synthesis are emerging technologies in the arena of medicinal chemistry [6]. Merrifield's conceptual solid-phase approach to peptide synthesis laid the foundation for the first set of combinatorial peptide libraries [7]. Over the years, solid-phase synthesis continues to undergo refinement and has been successfully extended to the synthesis of small organic molecules. Today combinatorial libraries are commonly used for the elucidation of structure-function relationships [8].

*Address correspondence to this author at the Chemical Biology Program, Steacie Institute for Molecular Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario, Canada, K1A 0R6; Tel. (613) 993 7014; Fax (613) 952 0068; e-mail: Prabhat.Arya@nrc.ca



Scheme 1.

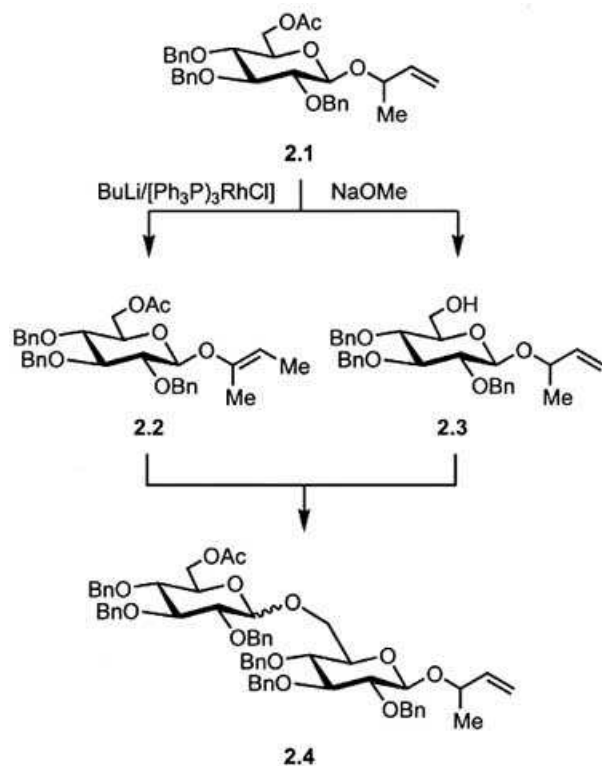
The branched nature of complex carbohydrates and the lack of highly stereoselective solid-phase reactions to obtain carbohydrate derivatives have precluded the rapid and efficient generation of oligosaccharide libraries either by solution or solid-phase synthesis. In recent years, several novel strategies in the generation of combinatorial libraries of oligosaccharides and of glycomimetics have been developed. Some of the major accomplishments made in this area are summarized in this article [9].

HIGH-THROUGHPUT SYNTHESIS OF OLIGOSACCHARIDES

High throughput synthesis of oligosaccharides is not a facile process. The polyvalent nature of carbohydrates, the requirement for many orthogonal protection and deprotection steps as well as the lack of a general glycosylation protocol

make the rapid assembly of carbohydrates a challenging process. Over the past few years, these challenges continue to be addressed resulting in novel approaches for the rapid assembly of oligosaccharides leading to combinatorial oligosaccharide libraries.

Initial approaches to oligosaccharide libraries sought to minimize the need for numerous orthogonally protected building blocks. As such, the first set of oligosaccharide libraries reported by Hindsgaul, *et al.*, utilized one major building block. This group used a random glycosylation strategy to rapidly attain di- and trisaccharides [10]. In their solution-phase approach, the major building block was a fully protected donor, activated with the trichloroacetimidate group while the acceptor contained six unprotected hydroxyl groups. Using this approach, a benzylated fucosyl donor, **1.1** (Scheme 1) reacted with a disaccharide acceptor, **1.2** to give a small library of all possible di- and trisaccharides. Six α-linked trisaccharides were the major products of the reaction while β-linked fucosylated trisaccharides were minor products. Only 30% of the disaccharide acceptor was fucosylated. Despite the fact that trisaccharides were rapidly attained and all the hydroxyl groups on the acceptor showed similar reactivity, the glycosylation was uncontrolled resulting in low yields.

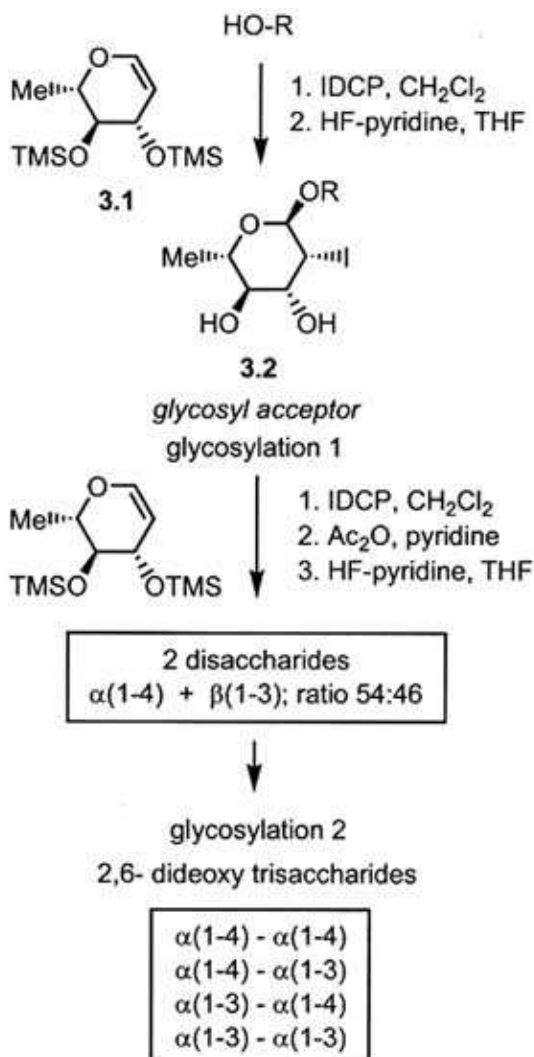


Scheme 2.

Boons, *et al.*, have also utilized one major building block in developing a novel latent-active glycosylation approach [11]. This major building block, 3-buten-2-yl glycoside, **2.1** (Scheme 2) can be readily converted into a glycosyl donor and into an acceptor. In this strategy, isomerization of **2.1** gives the glycosyl donor **2.2** while deprotection of the acetate group of **2.1** gives the glycosyl acceptor **2.3**. The disaccharide resulting from the glycosylation reaction between the donor and the acceptor was attained in 89% yield as an anomeric mixture. The methodology was used for the solution-phase synthesis of mixtures of linear and branched trisaccharide libraries in which the products were present in over 80% yield.

One monosaccharide building block was also utilized in the synthesis of a small, solution-phase library of 2,6-dideoxy trisaccharides by Ichikawa's group [12]. In their stereoselective yet non-regioselective strategy, the configuration at the anomeric position is controlled, and the regioisomers are generated. The building block, 6-deoxy-3,4-di-*O*-trimethylsilyl-L-glucal **3.1** (Scheme 3), was first

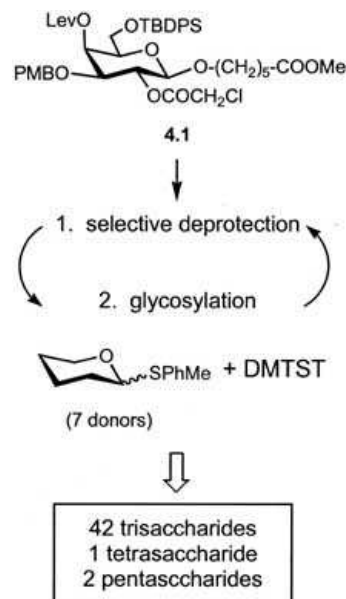
coupled to 6-trifluoroacetamidohexanol (R-OH) in the presence of iodonium di(sym-collidine)perchlorate (IDCP). The stereospecific reaction generated the α -glycoside with an iodo group at the 2-position. Deprotection of the trimethylsilyl groups (TMS) gave the glycosyl acceptor **3.2**, which can then undergo glycosylation with the glucal under IDCP catalysis. After two cycles of glycosylation, regioisomeric linear trisaccharides were obtained in 73% yields. The trisaccharides all contain 2-iodo groups which can either be reduced to give 2-deoxy derivatives or can undergo substitution resulting in further diversity.



Scheme 3.

The above solution-phase oligosaccharide library approaches gave rise to mixtures of anomers and/or regioisomers. In an effort to synthesize oligosaccharide libraries with defined structural features, Wong's group has designed a monosaccharide building block containing four selectively removable protecting groups [13]. The central glycosyl acceptor, **4.1** (Scheme 4) contained a chloroacetyl (ClAc), *p*-methoxybenzyl (PMB), levulinyl (Lev) and *tert*-butyldiphenylsilyl (TBDPS) groups with methyl 6-hydroxyhexanoate at the anomeric position. This acceptor was efficiently glycosylated in parallel, with seven thioglycoside donors in the presence of dimethyl(methylthio)sulfonium

triflate (DMTST) to give a solution-phase library of 45 protected oligosaccharides (tri-, tetra-, and pentasaccharides).



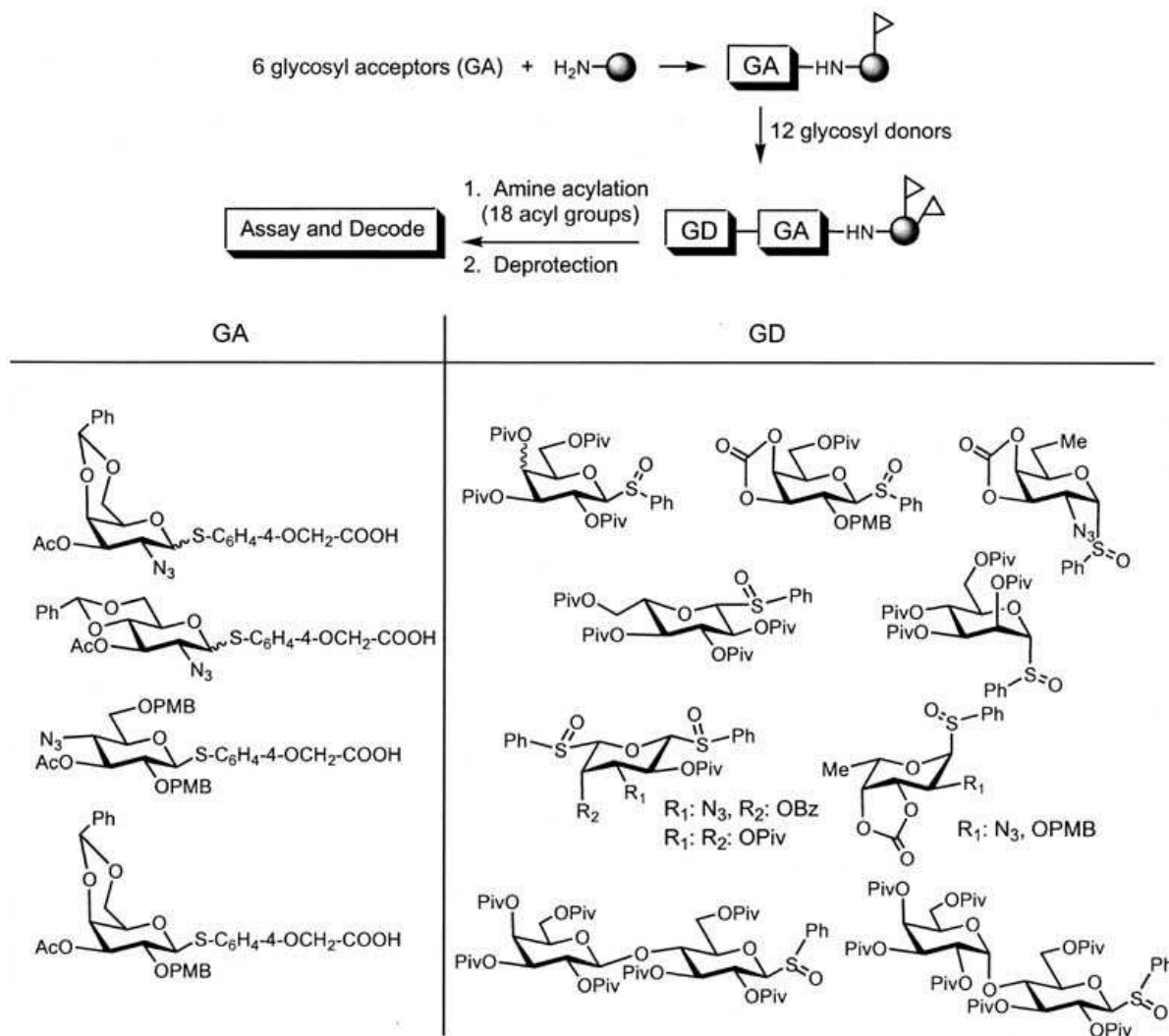
Scheme 4.

The differences in reactivities within a class of glycosyl donors such as *p*-methylphenyl thioglycosides have also been exploited in the search for facile oligosaccharide library generation. Wong's group have correlated the reactivity of thioglycosides containing a variety of electron donating or electron withdrawing groups with the chemical shift of the anomeric proton [14]. This has led to the development of a computerized database, OptiMer for selecting differentially protected thioglycoside building blocks for solution-phase sequential one pot glycosylation assembly as shown in Scheme 4. Using this approach, an oligosaccharide library of 33 tri- and tetrasaccharides was rapidly assembled using selectively protected monosaccharides and some disaccharides [15].

Takahashi's group has also synthesized a combinatorial library of trisaccharides by solution-phase one-pot glycosylation [16]. In this library, a variety of bromo glycosides, phenylthio glycosides and 2-bromoethyl glycosides of glucose, galactose and mannose were selectively activated with either silver triflate or *N*-iodosuccinimide/triflic acid and rapidly assembled on a QUEST 210 manual synthesizer to give a library of 72 trisaccharides (Scheme not shown).

To date, the successful solution-phase approaches to combinatorial oligosaccharide libraries certainly rivals solid-phase library generation. Thus far, only two reports of combinatorial oligosaccharide libraries on solid support have been published. This can be attributed to the additional challenges of solid support-derived approaches together with the inherent nature of oligosaccharides.

For solid-phase carbohydrate library generation, Kahne's group used anomeric sulfoxides as glycosyl donors [17]. The

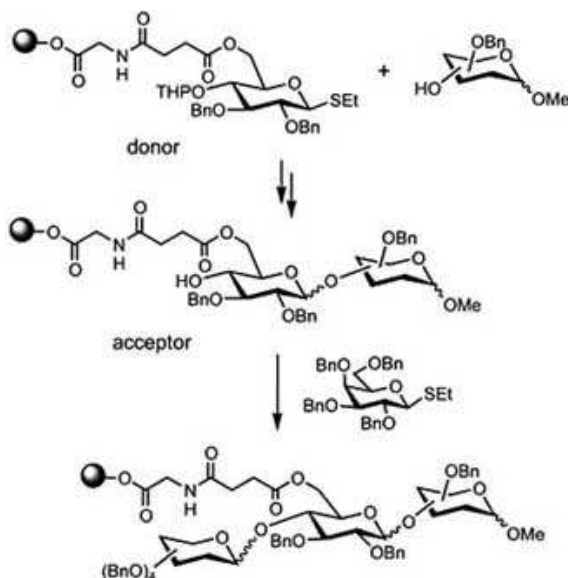


Scheme 5.

sulfoxides can be glycosylated at low temperatures with efficient stereocontrol coupling reaction for a range of donor-acceptor pairs. A split-mix library approach with chemical encoding at each combinatorial step was used. The glycosyl acceptors were separately attached to TentaGel amine resin and twelve glycosyl donors were then coupled followed by N-acylation (**Scheme 5**). After three steps, a tagged library of approximately 1300 di- and trisaccharides was produced on beads. Each bead contained a single carbohydrate and this on-bead library was screened against the *Bauhinia purpurea* lectin.

The second solid-phase library was synthesized by Zhu and Boons [18]. In this synthesis, a thioethyl glycoside that can function as a donor or an acceptor was immobilized onto glycine derivatized TentaGel resin through a succinimidyl linker. In order to eliminate any oligomeric side products during glycosylation, the tetrahydropyranyl (THP) group was used to protect the C4-hydroxy group which was unmasked in the next step to generate the acceptor (**Scheme 6**). A small library of 12 trisaccharides was synthesized using a split-mix approach to generate anomeric mixtures.

The challenges of solid-phase oligosaccharide library generation continue to be addressed. A number of solid-phase approaches have been described in the literature for oligosaccharides with the potential for library generation. To overcome the use of an insoluble support, a soluble polyethylene glycol (PEG) polymer has been used to maximize the use of solution-phase glycosylation protocols as well as taking advantage of solid-phase purification for the rapid assembly of oligosaccharides [19]. Ito's group used the soluble PEG support and a hydrophobic tag at the reducing end of the oligosaccharide for ready purification. In this approach, an orthogonal glycosylation strategy, employing a combination of thioglycoside and glycosyl fluoride was used for the rapid assembly of a tetrasaccharide. New linkers have also been developed, such as the thiol linker for α -mannose and α -fucose glycosides [20] and a ring closing metathesis based linker that generates *O*-allyl glycosides upon cleavage from the resin [21]. Derivatives such as *n*-pentenyl glycosides have been demonstrated to be adaptable to solid-phase carbohydrate library generation [22] while the synthesis of β -(1-4)- and β -(1-6)- linked oligosaccharides using glycosyl phosphates in combination



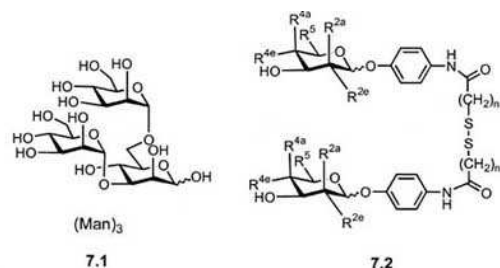
Scheme 6.

with a versatile octenediol linker can be successfully accomplished [23]. Other solid-phase approaches described in the literature include the glycal assembly method for the synthesis of polymer bound thioethyl glycosyl donors for the synthesis of β -linked oligosaccharides [24]; the synthesis of thio-oligosaccharides by nucleophilic substitution of triflate activated glycosides by resin-bound sugar-1-thiolate containing unprotected hydroxyl groups [25], and the use of a novel photocleavable aglycon linker for the synthesis of a dodecasaccharide with high stereospecificity [26].

The solution-phase and solid-phase library generation for oligosaccharides discussed above differ fundamentally from

Lehn's approach to combinatorial carbohydrate library generation [27]. This dynamic combinatorial library (DCL) strategy involves the transient formation of compounds *in situ*, using reversible reactions, in the presence of a receptor (adaptive combinatorial library) which then selects for the ligand(s) that possess the highest affinity. Alternatively, the receptor can be added after equilibration of the constituents (pre-equilibrated dynamic library) [28].

To demonstrate the feasibility of this approach, a dynamic combinatorial library was generated against the lectin, concanavalin A (Con A). This plant lectin, specifically binds a branched trimannoside unit **7.1** (see table



Compound	$\alpha\beta$	R^{2a}	R^{2e}	R^{4a}	R^{4e}	R^5	n
man/Man	α	OH	H	H	OH	CH_2OH	3
GalC2/GalC2	β	H	OH	OH	H	CH_2OH	2
GalC3/GalC3	β	H	OH	OH	H	CH_2OH	3
Glc/Glc	β	H	OH	H	OH	CH_2OH	2
Ara/Ara	β	H	OH	OH	H	H	
2Xyl/Xyl	β	H	OH	H	OH	H	2

Scheme 7.

in **Scheme 7**) located on certain glycoproteins. The chemoselective thiol-disulfide reaction was selected in which appropriate spacer between monosaccharide units can be generated. The monosaccharides, D-mannose, D-glucose, D-galactose, L-arabinose and D-xylose selected were derivatized with the phenylamido group, **7.2** attached to two linkers that differed by a CH₂ group. Both homodimers (pH \geq 7) and heterodimers (in the presence of dithiothreitol) were generated and libraries were produced in the presence of Con A or Con A was added after pre-equilibration. For these libraries, Con A was immobilized onto sepharose beads for ready isolation of bound compounds. A bis-mannose compound was preferentially selected by the receptor, which demonstrated the potential for such simultaneous library generation and screening capabilities of this approach [27].

CARBOHYDRATE SCAFFOLDS

Over the past 40 years carbohydrates have been shown to play vital roles in a vast number of biological recognition events. It is because of their complex structures that they were not studied extensively, and yet, it is this complexity that enables carbohydrates to be involved in such a wide range of processes. It is also due to this complexity that synthesis of oligosaccharide libraries has not been transferred to automated technologies in the same manner as peptide and oligonucleotide synthesis. However, the ability to diversify a carbohydrate template has been used increasingly in the drug discovery process [29]. The polyfunctional nature of a carbohydrate unit is ideal for use as a scaffold to produce libraries of compounds that are not necessarily glycoconjugates. Previous work had demonstrated the validity of this approach in the design of nonpeptide somatostatin mimics [30]. The first solid-phase library based on a carbohydrate scaffold containing three sites of diversity

was reported in 1998 by Sofia *et al.* [31]. The scaffold contained a functional triad consisting of a carboxylic acid moiety, a free hydroxyl group and a protected amino acid group on the glucoside **F1** (**Figure 1**). This derivatized monosaccharide was then coupled to an amino acid functionalized trityl TentaGel resin. Using IRORI radiofrequency tagged split-pool methodology [32], sixteen 48-member libraries were prepared from eight amino acids, six isocyanates and eight carboxylic acids. The libraries were analyzed by LC/MS and found to have greater than 80% purity. These libraries were dubbed "universal pharmacophore mapping libraries" by their authors.

The solid-phase synthesis of a 48 member library of a β -disaccharide on Rink amide resin was achieved using disaccharide **F2** [33]. The use of phenylsulfonyl 2-deoxy-2-trifluoroacetamido glycopyranosides as glycosyl donors afforded the disaccharide scaffold stereospecifically and in high yields. The disaccharide was derivatized using six isocyanates and eight carboxylic acids. The same donor molecule and its galactosamine counterpart were used in the synthesis of moenomycin A dissaccharide templated libraries **F3** [34]. Diversity was introduced at C-1, C-3 and C-2', and a library of 1300 disaccharides was successfully prepared. The library was screened for both inhibition of bacterial cell wall biosynthesis and inhibition of bacterial growth, and a novel class of potent inhibitors of both processes was identified.

Another approach to combinatorial libraries used an orthogonally protected thioglucoside scaffold [35]. The protecting groups included *tert*-butyldiphenylsilyl (TBDPS), 1-ethoxy ethyl (OEE) and allyl (OAl) which was converted to *n*-propyl. An important feature of the scaffold was the use of a functionalized thioglycoside, which served as a linker for immobilizing the compound on aminomethyl

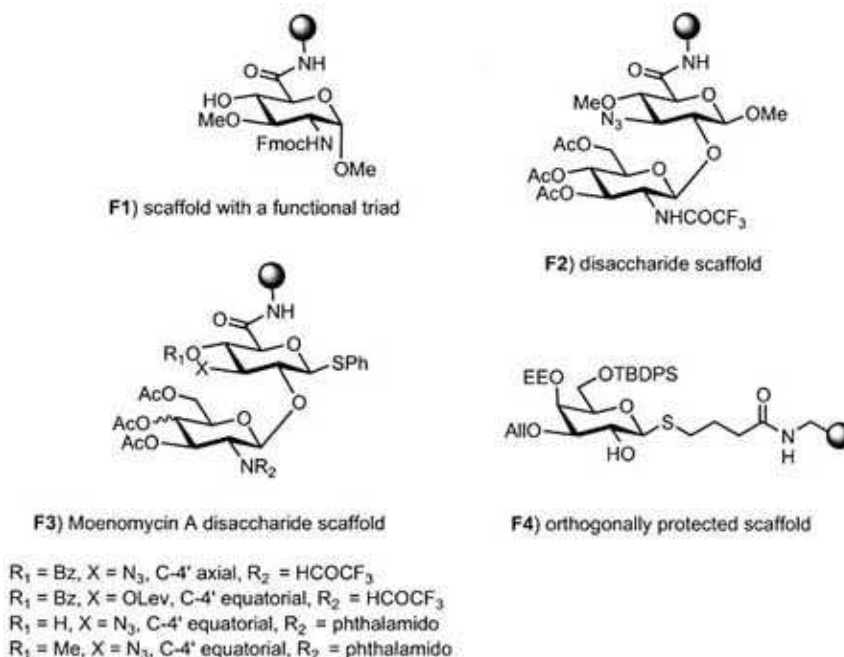
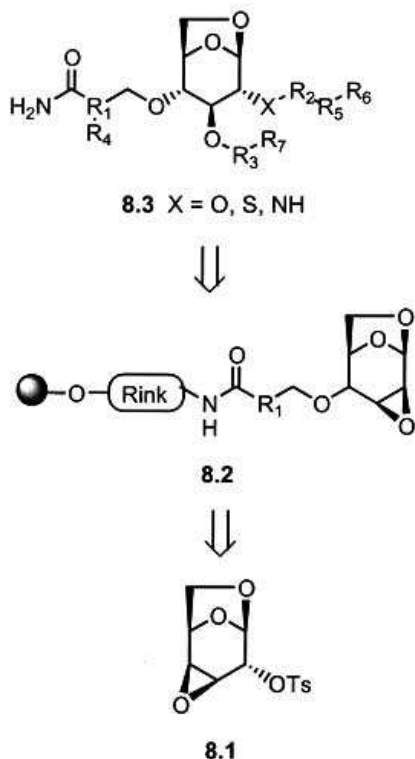


Fig. (1).



Scheme 8.

polystyrene resin and later as a glycosyl donor to diversify the anomeric center. This methodology was extended to a galactopyranose scaffold **F4** containing five sites of diversity [36]. Using sequential deprotection and alkylation protocols, an array of structurally diverse compounds was successfully synthesized.

Silva and Sofia have recently reported the synthesis of a uridine-mannose scaffold based on tunicamycin [37]. The key step in the synthesis of the β -L-mannoside derivative involved the use of Crich's modification of the sulfoxide glycosylation. The authors are investigating the synthesis of tunicamycin analogs containing two points of diversity at the C-6 azide and C-4 hydroxyl group with the expectations of finding out more about the mechanism of action of the tunicamycins. Ghosh, *et al.*, have reported the synthesis of a

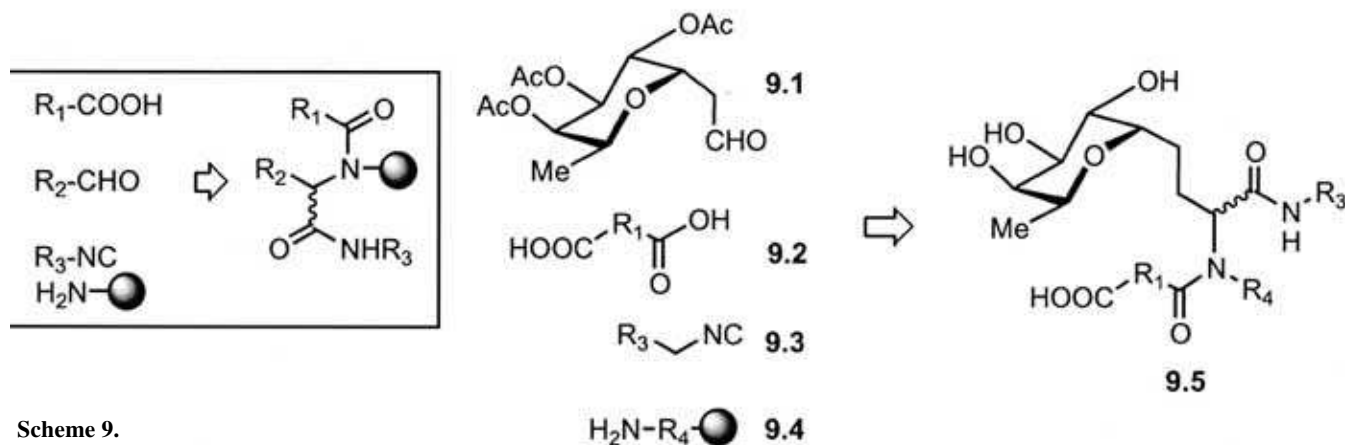
scaffold based on 2-acetamido-2-deoxy-D-glucose containing an azide group at C-6, a hydroxyl group at C-4 and a carboxylic acid functionality at C-1 [38]. The authors allude to a 12000 compound library, but no details are given.

The search for potent aminoglycoside mimics lead Wong, *et al.*, to rationally design a library with an aminoglucopyranoside core. The aminoglucopyranoside contains a 1,3-hydroxyamine motif at the anomeric position and has been synthesized using a parallel solution-phase approach [39]. The small molecule library was screened for binding to RNA and several compounds showed high affinity for the target.

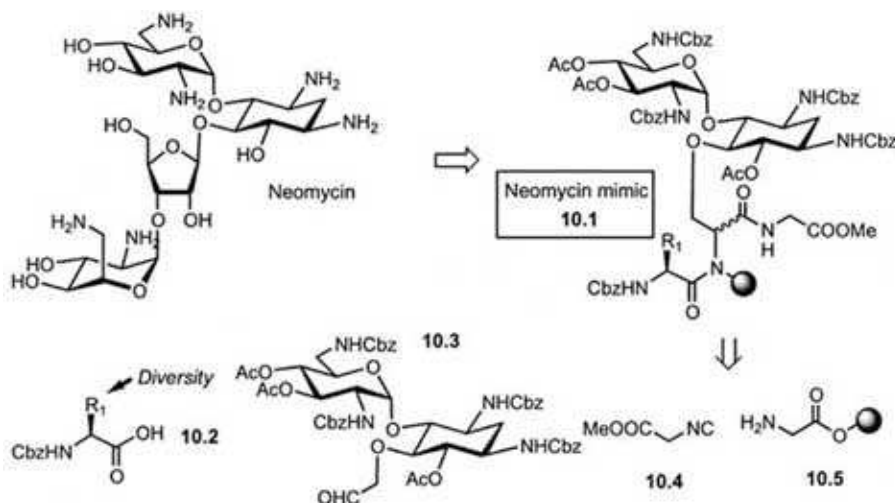
The solid-phase synthesis of levoglucosan derivatives starting with 1,6:3,4-di-anhydro-2-*O*-*p*-toluenesulfonyl- β -D-galactopyranose **8.1** (Scheme 8) was reported in 1998 by a group of scientists from Novartis [40]. The scaffold **8.2** was diversified at C-2, C-3 and C-4 with various alkoxy, thioxy or amino groups. These groups were further derivatized using acylations or Pd-mediated coupling reactions to give compounds with the framework of **8.3**. The key step was the opening of the stable epoxide, and the lack of complex protecting schemes is noteworthy. These researchers also reported the opening of several other levoglucosan epoxides using oxygen, nitrogen and sulfur nucleophiles under mild acidic or basic conditions [41]. These compounds are further amenable to diversification at several sites. A review of their efforts has subsequently been published [42].

MULTIPLE COMPONENT CONDENSATION (MCC)

Ugi's multi-component reaction involving the condensation of an isocyanide, aldehyde, amine and a carboxylic acid results in α -acylamino amide derivatives. This reaction allows for the generation of a large number of compounds and has been adapted to solid-phase synthesis [43]. This reaction has been employed in the synthesis of a library of C-glycosides [44]. As shown in Scheme 9, a C-fucose aldehyde **9.1** was reacted with eight diacids **9.2**, two isocyanides **9.3** and Rink amide resin derivatized with five different amino acids **9.4**. The library of sialyl Lewis^x mimetics **9.5** was obtained rapidly and in high purity. Wong's group has utilized this methodology to synthesize a library of aminoglycoside antibiotic mimetics on a soluble polyethyleneglycol (PEG) polymer based on compound **10.1**



Scheme 9.

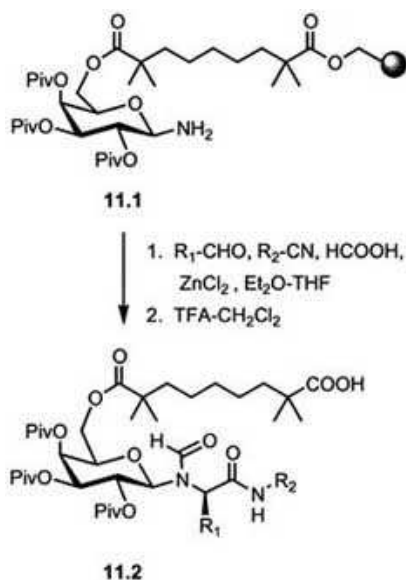


Scheme 10.

(Scheme 10) [45]. The library was designed to have diversity introduced in the amino group **10.2**, while keeping the neamine moiety **10.3** constant because it is critical for the inhibition of HIV RNA transactivator protein.

Researchers at Bayer have shown that carbohydrate building blocks containing aldehyde, amino, carboxylic acid and isocyanide groups can be incorporated into the Ugi MCC [46]. They were also able to condense four carbohydrate building blocks resulting in a four fold-glycosylated amino acid derivative. These glycomimetics may be used as probes for various biological interactions as well as give rise to new lead structures for drug development.

The first stereoselective synthesis using Ugi-MCC on solid-phase was carried out using an *O*-pivaloylated galactosylamine **11.1** (Scheme 11) linked to Wang resin via a $\alpha,\alpha,\alpha',\alpha'$ -tetramethyl azelaic acid spacer. The compounds **11.2** were obtained with high diastereoselectivity, which were verified by cleaving the *N*-glycosidic bond, releasing

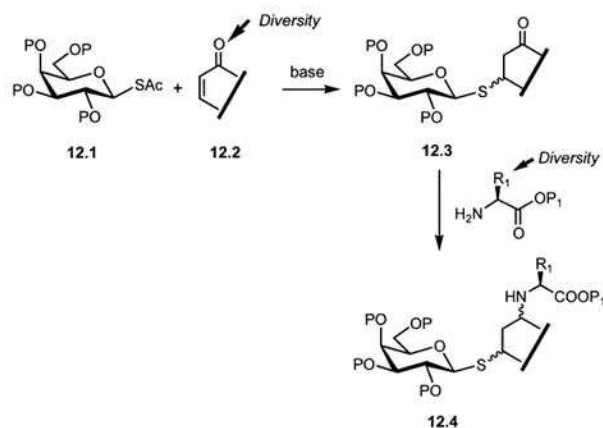


Scheme 11.

the amino acid derivatives, and comparing with the authentic compounds prepared by solution synthesis [47].

GLYCOHYBRIDS

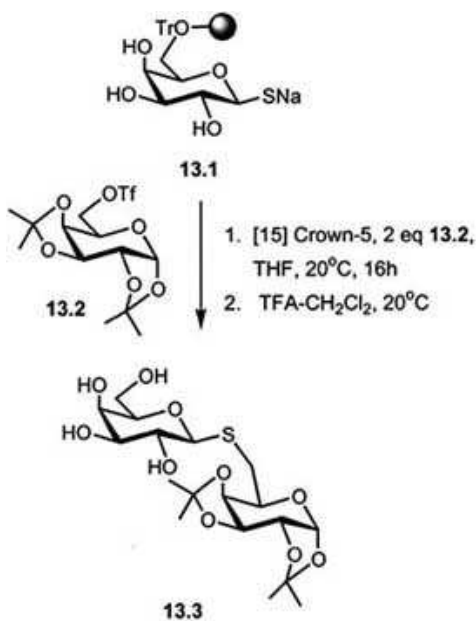
Using a novel solid-phase extraction purification strategy, a 1-thio- β -D-galactopyranoside library was synthesized in solution by Hindsgaul *et al.* [48]. The hydroxyl groups of a galactosyl building block **12.1** (Scheme 12) were protected as their hydrophobic *O*-laurates, which allowed adsorption of the products onto C_{18} silica and enabled the reagents and by-products of each reaction to be washed away using methanol, followed by elution of the products with pentane. The thiogalactoside was reacted with



Scheme 12.

a series of Michael acceptors and an α -chloro ketone. The resulting ketones **12.3** were either reduced to the corresponding alcohol or reductively aminated with different amino acids. A library of thirty compounds, **12.4**, each containing a mixture of four diastereomers was produced. This library was screened for inhibitors of β -galactosidase from *E. coli*. One compound was found to be a competitive inhibitor.

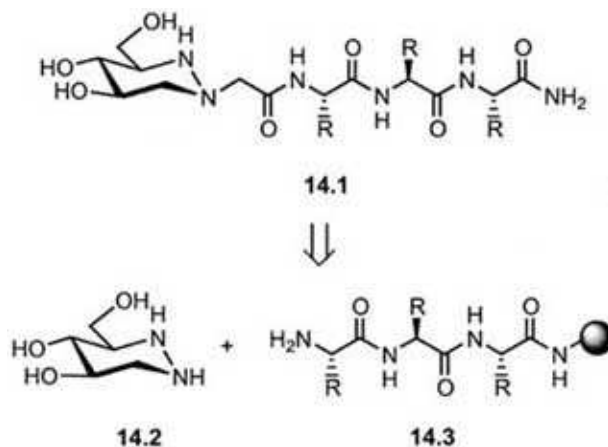
A second method for the production of thioglycosides coupled a highly reactive resin-bound nucleophilic sugar-1-thioate **13.1** (for an example see **Scheme 13**), without protecting groups, to a triflate-activated glycoside **13.2** [49]. The reactions proceeded with a high degree of stereoselectivity and in high yields to give the desired disaccharides **13.3**.



Scheme 13.

Another report on the synthesis of 1-thioglycosides involved direct *S*-alkylation of 4-*O*-triflates and primary tosyl ether derivatives of several pyranosides using iminophosphorane bases [50]. The reactions were carried out in solution and on solid-phase to give thio-disaccharides stereoselectively and in high yield.

Many known inhibitors of glycosidases are carbohydrate molecules containing nitrogen in the ring [51]. The solid-phase synthesis of piperidine based carbohydrate mimics containing 4-hydroxypiperidine-3-carboxylic acid was achieved in 1997 [52]. Recently, a library of 125 compounds **14.1** (**Scheme 14**) containing 1-azafagomine **14.2** linked to a tripeptide **14.3** via acetic acid was synthesized using split and mix peptide library synthesis approach [53]. The sublibraries were screened for inhibition of β -glucosidase, α -glucosidase and glycogen phosphorylase. The sublibrary containing hydroxyproline at the *N*-terminus was deconvoluted to identify the compound responsible for the inhibition of β -glucosidase. The most potent inhibitor found contained three hydroxyproline residues.



Scheme 14.

GLYCOSYLATED AMINO ACID BUILDING BLOCKS

Many glycosylated *N*-fluorenyl-methoxycarbonyl (Fmoc) amino acid pentafluorophenyl esters (OPfp) (**Figure 2**) or free Fmoc-amino acids have been used in the solid-phase synthesis of glycopeptides. An assortment of solid supports have also been used to produce parallel arrays of stereoselectively linked glycopeptides [5,9]. A report by St. Hilaire and Meldal demonstrates the use of three different glycosyl amino building blocks to synthesize a library consisting of 300 000 components (**Scheme 15**) [54]. At each step in the library synthesis a small portion of intermediates was capped, producing a ladder of intermediate fragments. An encoded one-bead-one compound heptaglycopeptide library was rapidly synthesized on a PEGA resin containing a photolabile linker by the split and mix technique. The library was screened using an on bead assay against fluorescently labeled lectin from *Lathyrus odoratus*. The active compounds were irradiated on-bead with a MALDI-TOF-MS laser. The resulting mass spectra contained the ladder of fragments capped during the synthesis, allowing for the identification of the active sequence. All the active sequences had a terminal mannose unit signifying the importance of this moiety for lectin recognition.

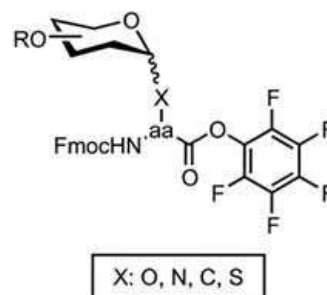
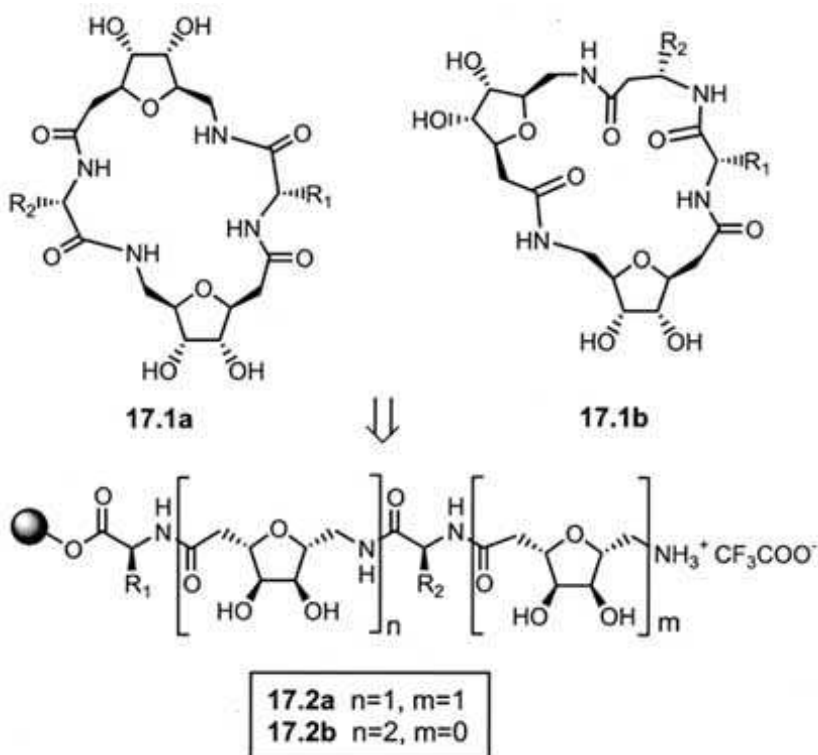


Fig. (2).



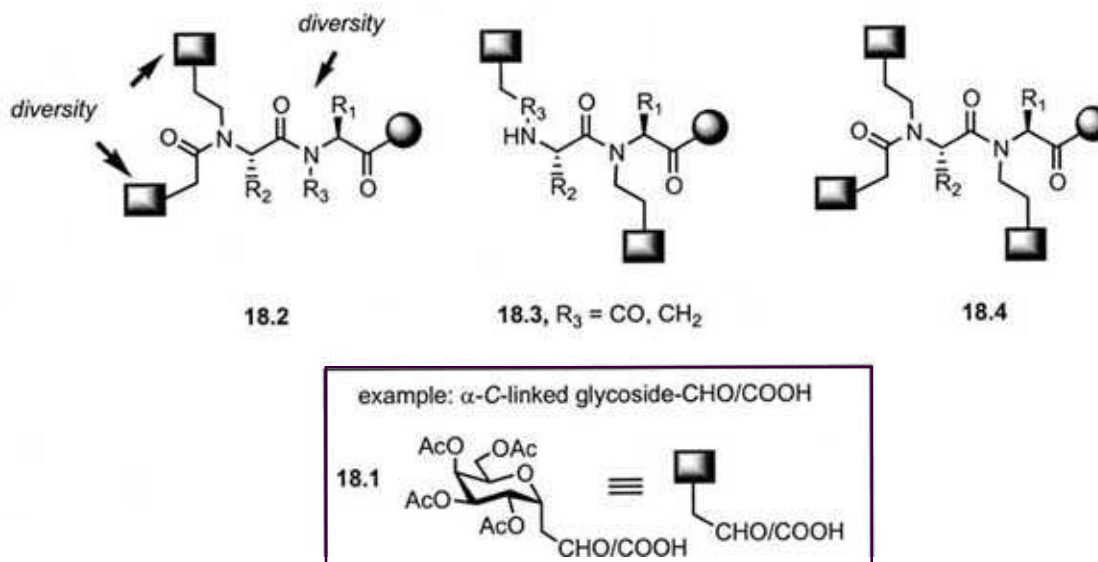
Scheme 17.

molecules **17.1a** and **17.1b**. The structures of the molecules were determined by NMR and preliminary conformational analysis of one product indicates that it adopts a defined conformation.

AUTOMATED, MULTI STEP APPROACH TO ARTIFICIAL GLYPEPTIDE LIBRARIES

Research carried out by Arya, *et al.*, has lead to development of an automated, multi-step, solid-phase strategy for the parallel synthesis of artificial glycopeptide

libraries (**Scheme 18**). Using this approach, different α - or β -carbon linked carbohydrate based aldehyde and carboxylic acid derivatives, protected as acetates **18.1** [58], can be incorporated either at the *N*-terminal moiety or at the internal amide nitrogen of short peptides/pseudopeptides in a highly flexible and controlled manner. The chain length of the *C*-glycoside can be varied and the carbohydrate moiety can be synthesized in either the pyranose or furanose form. The type of carbohydrate building block is not limited to monosaccharide derivatives, it is conceivable to use disaccharides and higher order oligosaccharides. *C*-Glycosides are more stable to enzymatic and acid/base



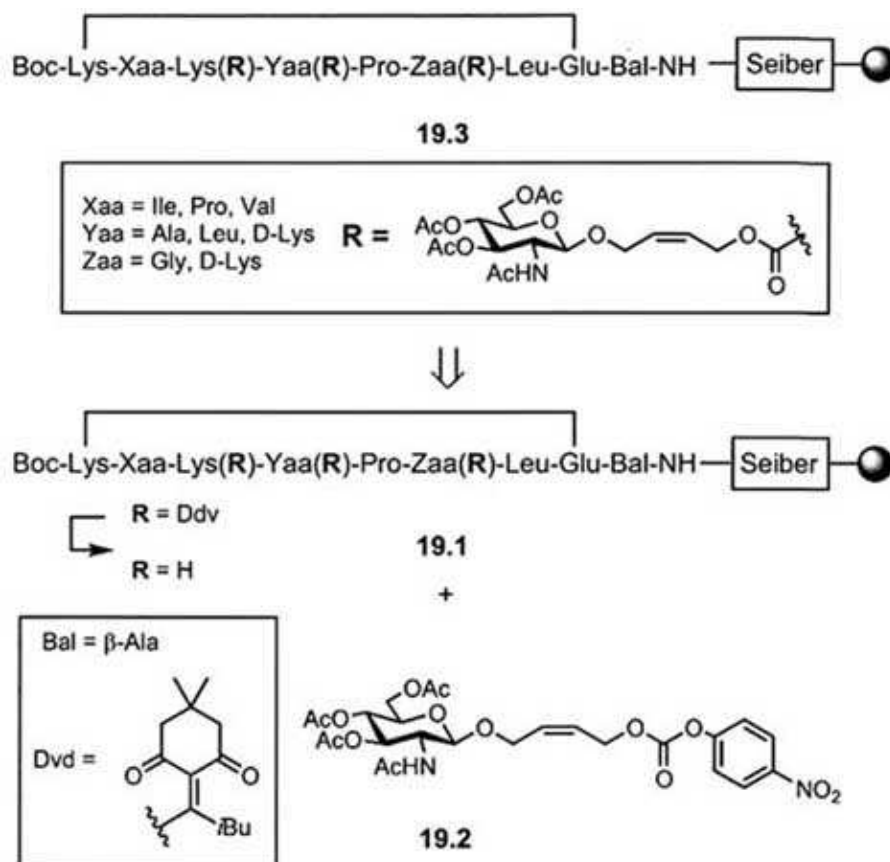
Scheme 18.

hydrolysis than their oxygen counterparts. This method is more versatile than the glycosylated amino acid building block in which the choice of amino acids is limited.

Using this approach, libraries of artificial glycopeptides could be readily synthesized for probing carbohydrate-protein interactions. Several "working models" that display multiple copies of carbohydrates have been developed **18.2**, **18.3** and **18.4** (Scheme 18) while the dipeptide scaffold may contribute to secondary interactions with the biological target [59]. Initially, the artificial glycopeptides were synthesized by a convergent strategy on a peptide synthesizer [60]. The synthesis of these artificial glycopeptide libraries has been successfully transferred to a fully automated multiple organic synthesizer and each step in the synthesis has been optimized [61]. This methodology involves coupling an amino acid to a solid-support resin such as Rink amide MBHA resin or TentaGel derivatized Rink amide resin. After removal of the protecting group on the amino acid, the sugar aldehyde undergoes reductive amination (models **18.3** and **18.4**) with the resin bound amino group followed by amino acid coupling of the second amino acid. After deprotection of the amino acid, a second reductive amination can occur and/or a sugar acid can be coupled. The sugar moieties are then deacetylated, and the compounds are cleaved from the resin. The synthesis of a 96 compound library can be obtained from just 24 dipeptides and two sugar aldehydes [62]. The choice of the dipeptide and sugar

moieties are reflected in the biological mechanism under investigation. The compounds in this library were used as chemical probes for studying protein folding and trafficking, primarily of *N*-linked glycoproteins [63]. The same library was tested in enzyme systems that convert a glucose moiety to rhamnose prior to incorporation of the rhamnose unit during the biosynthesis of the mycobacterium cell wall [64]. The dipeptide scaffold consisted of negatively charged amino acids, polar amino acids as well as hydrophobic residues, while two different monosaccharide units were used. A number of potential glycoside-based inhibitors containing at least one negatively charged amino acid residue were identified and detailed biological studies are in progress. The synthesis of several new libraries is in progress, including those with a nucleoside moiety at the *C*-terminal end of the model compounds (**18.2-18.4**), as chemical probes for investigating various glycosyl transferase-based reactions.

A recent communication described the synthesis of multivalent cyclic neoglycopeptides [65]. A new urethane-type linker based on the Alloc protecting group has been developed for the glycosylation reaction which proceeds virtually quantitatively. A library of cyclic peptides was synthesized using the split and mix method on TentaGel resin linked via the Sieber linker **19.1** (Scheme 19). The synthesis was monitored by withdrawal of a small amount of resin from the well and analysis by HPLC in combination with electrospray mass spectrometry. The *p*-nitrophenyl



Scheme 19.

carbonate derivative of the sugar moiety **19.2** was attached to three points of the cyclic peptide in one step using a five-fold excess (per attachment point) of the sugar in the presence of DIPEA. A library of eighteen cyclic neoglycopeptides **19.3** was efficiently synthesized. This methodology can be applied to the synthesis of many different libraries by varying the distances between the carbohydrate moieties as well as the carbohydrate moiety itself.

SUMMARY AND OUTLOOK

One of the important factors in understanding the biological role and functions of glycoconjugates is the availability of efficient synthetic tools that affords chemically well defined complex carbohydrates and glycoconjugates. During the past few decades, several groups have developed elegant solution and solid-phase methods to achieve this goal. However, developing solid-phase approaches that could be subjected to automation for complex carbohydrates and glycoconjugates is not a trivial undertaking. Few groups have been successful in bringing "automated technologies" to the synthesis of complex carbohydrates and glycoconjugates, but these are early stages. The continued development of highly stereoselective glycoside coupling reactions on solid-phase is of prime importance in establishing automated synthetic processes.

Over the years, several solid-phase methods for the synthesis of nucleic acid and peptide mimics have emerged. These methods are crucial for developing combinatorial, high-throughput approaches for obtaining large sets of compounds over a short period. In several cases, the successful outcome has resulted in several promising compounds with useful biological properties. Similarly, there is a need for developing efficient, solid-phase methods for the synthesis of structural and functional mimics of complex carbohydrates and glycoconjugates. It is hoped that some of these solid-phase methods could be subjected to automation for developing combinatorial, high-throughput syntheses. These will be powerful methods in providing chemically well defined carbohydrate-derived compounds, a much needed effort in understanding the biological roles and functions of complex carbohydrates and glycoconjugates.

REFERENCES

- (a) NRC publication no. 43876
- (b) Post-doctoral fellow, 2000-2001.
- (c) Post-doctoral fellow, 1998-2001.
- [1] Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683. Varki, A. *Glycobiology* **1993**, *3*, 97; Branza-Nichita, N.; Petrescu, A. J.; Negroiu, G.; Dwek, R. A.; Petrescu, S. M. *Chem. Rev.* **2000**, *100*, 4697. Butters, T. D.; Dwek, R. A.; Platt, F. M. *Chem. Rev.* **2000**, *100*, 4683. Duus, J. Ø.; St. Hilaire, P. M.; Meldal, M.; Bock, K. *Pure Appl. Chem.* **1999**, *71*, 755.
- [2] Danishefsky, S. J.; Roberge, J. Y. *Pure Appl. Chem.* **1995**, *67*, 1647. Gege, C.; Vogel, J.; Bendas, G.; Rothe, U.; Schmidt, R. R. *Chem. Eur. J.* **2000**, *6*, 111. Seeberger, P. H.; Danishefsky, S. J. *Acc. Chem. Res.* **1998**, *31*, 685. Meinjohanns, E.; Meldal, M.; Paulsen, H.; Dwek, R. A.; Bock, K. *J. Chem. Soc. Perkin Trans 1*, **1998**, 549. Sames, D.; Chen, X.-T.; Danishefsky, S. J. *Nature*, **1997**, 389, 587. Seitz, O.; Wong, C.-H. *J. Am. Chem. Soc.* **1997**, *119*, 8766. Danishefsky, S. J.; Bilodeau, M. T. *Angew. Chem. Int. Ed. Engl.* **1996**, *35*, 1380. Wang, Z.-G.; Zhang, X.-F.; Y. Ito; Nakahara, Y.; Ogawa, T. *Carbohydrate Res.* **1996**, *295*, 25. Danishefsky, S. J.; Allen, J. A. *Angew. Chem. Int. Ed. Engl.* **2000**, *39*, 836 and references therein.
- [3] See, Preparative Carbohydrate Chemistry (Ed.: S. Hanessian), Marcel Dekker, **1997**. Benjamin, G.; Davis, J. *J. Chem. Soc. Perkin Trans. 1*, **2000**, 2137. Toshima, K.; Tatsuta, K. *Chem. Rev.* **1993**, *93*, 1503. Hanessian, S.; Lou, B. *Chem. Rev.* **2000**, *100*, 4443. Boons, G.-J.; Demchenko, A. V. *Chem. Rev.* **2000**, *100*, 4539. Jung, K.-H.; Muller, M.; Schmidt, R. R. *Chem. Rev.* **2000**, *100*, 4423.
- [4] For recent reviews on solid-phase oligosaccharide synthesis and programmed synthetic approaches to complex carbohydrates and carbohydrate conjugates, see: Seeberger, P. H.; Haase, W.-C. *Chem. Rev.* **2000**, *100*, 4349. Koeller, K. M.; Wong, C.-H. *Chem. Rev.* **2000**, *100*, 4465. Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495.
- [5] Marcaurelle, L. A.; Bertozzi, C. R. *Chem. Eur. J.* **1999**, *5*, 1384. Simanek, E. E.; McGarvey, G. J.; Jablonowski, J. A.; Wong, C.-H. *Chem. Rev.* **1998**, *98*, 833. Sears, P.; Wong, C.-H. *Angew. Chem. Int. Ed.* **1999**, *38*, 2300. St. Hilaire, P. M.; Meldal, M. *Angew. Chem. Int. Ed.* **2000**, *39*, 1162. Dondoni, A.; Marra, A. *Chem. Rev.* **2000**, *100*, 4395.
- [6] Thompson, L. A.; Ellman, J. A. *Chem. Rev.* **1996**, *95*, 555. Gordon, E. M.; Gallop, M. A.; Patel, D. V. *Acc. Chem. Res.* **1996**, *29*, 144. Lam, K. S.; Lebel, M.; Krchnak, V. *Chem. Rev.* **1997**, *97*, 411. Dolle, R. E.; Nelson Jr., K. H. *J. Comb. Chem.* **1999**, *1*, 235. Balkenhohl, F.; von dem Bussche-Hunnefeld, C.; Lansky, A.; Zechel, C. *Angew. Chem. Int. Ed.* **1996**, *35*, 2288. Guillier, F.; Orain, D.; Bradley, M.; *Chem. Rev.* **2000**, *100*, 2091.
- [7] Merrifield R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149. Furka, A.; Sebestyen, F.; Asgedom, M.; Dibo, G. *Int. J. Peptide Protein Res.* **1991**, *37*, 487. Houghten, R. A.; Pinilla, C.; Blondelle, S. E.; Appel, J. R.; Dooley, C. T.; Cuervo, J. H. *Nature*, **1991**, *354*, 84.
- [8] Combinatorial Chemistry and Molecular Diversity in Drug Discovery (Eds.: Gordon, E. M.; Kerwin, Jr, J. F.), Wiley, **1998**; Molecular Diversity and Combinatorial Chemistry- Libraries and Drug Discovery (Eds.: Chaiken, I. M.; Janda, K. D.), ACS Series, **1996**; Solid-Supported Combinatorial and Parallel Synthesis of Small-Molecular-Weight Compound Libraries by Obrecht, D.; Villalgorido, J. M. Pergamon, **1998**; Solid-phase Organic Synthesis (Ed. Burgess, K.), Wiley-Interscience, **2000**. See, a special issue on combinatorial chemistry, *Chem. Rev.* **1997**, *2*.
- [9] For early reviews on combinatorial efforts in complex carbohydrates, carbohydrate conjugates and carbohydrate mimics, see: Sofia, M. J. *Mol. Diversity*

- 1998, 3, 75. Wang, Z.-G.; Hindsgaul, O. *Glycoimmunology* **1998**, 2, 219. Schweizer, F.; Hindsgaul, O. *Curr. Opin. Chem. Bio.* **1999**, 3, 291. Arya, P.; Ben, R. N. *Angew. Chem. Int. Ed.* **1997**, 36, 1280. Arya, P.; Ben, R. N.; Kutterer, K. M. K. in *Organic Synthesis Highlights, vol 4*, **2000**, 337. Barkley, A.; Arya, P. *Chem. Eur. J.* **2001**, 7, 0000.
- [10] Kanie, O.; Berresi, F.; Ding, Y.; Labbe, J.; Otter, A.; Forsberg, L. S.; Ernst, B.; Hindsgaul, O. *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 2720. Ding, Y.; Labbe, J.; Kanie, O.; Hindsgaul, O. *Bioorg. Med. Chem.* **1996**, 4, 683.
- [11] Boons, G.-J.; Isles, S. *J. Org. Chem.* **1996**, 61, 4262. Johnson, M.; Arles, C.; Boons, G.-J. *Tetrahedron Lett.* **1998**, 39, 9801.
- [12] Izumi, M.; Ichikawa, Y. *Tetrahedron Lett.* **1998**, 39, 2079.
- [13] Wong, C.-H.; Ye, X.-S.; Zhang, Z. *J. Am. Chem. Soc.* **1998**, 120, 7137.
- [14] Zhang, Z.; Ollmann, I. R.; Ye, X.-S.; Wischnat, R.; Bassov, T.; Wong, C.-H. *J. Am. Chem. Soc.* **1999**, 121, 734.
- [15] Ye, X.-S.; Wong, C.-H. *J. Org. Chem.* **2000**, 65, 2410.
- [16] Takahashi, T.; Adachi, M.; Matsuda, A.; Doi, T. *Tetrahedron Lett.* **2000**, 41, 2599.
- [17] Liang, R.; Yan, L.; Loebach, J.; Ge, M.; Uozumi, Y.; Sekanina, K.; Horan, N.; Glidersleeve, J.; Thompson, C.; Smith, A.; Biswas, K.; Still, W. C.; Kahne, D. *Science* **1996**, 274, 1520. Horan, N.; Yan, L.; Isobe, H.; Whitesides, G. W.; Kahne, D. *PNAS* **1999**, 96, 11782.
- [18] Zhu, T.; Boons, G.-J. *Angew. Chem. Int. Ed. Engl.* **1998**, 37, 1898.
- [19] Douglas, S. P.; Whitfield, D. M.; Krepinsky, J. J. *J. Am. Chem. Soc.* **1995**, 117, 2116. Ito, Y.; Kanie, O.; Ogawa, T. *Angew. Chem. Int. Ed. Engl.* **1996**, 35, 2510.
- [20] Rademann, J.; Schmidt, R. R. *J. Org. Chem.* **1997**, 62, 3650.
- [21] Knerr, L.; Schmidt, R. R. *Synlett.* **1999**, 1802-1804.
- [22] Rodebaugh, R.; Joshi, S.; Fraser-Reid, B. Geysen, H. M. *J. Org. Chem.* **1997**, 62, 5660. Anilkumar, G.; Nair, L. G.; Fraser-Reid, B. *Org. Lett.* **2000**, 2, 2587.
- [23] Andrade, R. B.; Plante, O. J.; Melean, L. G.; Seeberger, P. H. *Org. Lett.*, **1999**, 1, 1811.
- [24] Zheng, C.; Seeberger, P. H.; Danishefsky, S. J. *J. Org. Chem.* **1998**, 63, 1126.
- [25] Hummel, G.; Hindsgaul, O. *Angew. Chem. Int. Ed. Engl.* **1999**, 38, 1782.
- [26] Nicolaou, K. C.; Watanabe, N.; Li, J.; Pastor, J.; Winssinger, N. *Angew. Chem. Int. Ed. Engl.* **1998**, 37, 1559.
- [27] Ramstrom, O.; Lehn, J.-M. *Chem. BioChem.* **2000**, 1, 41. Lehn, J.-M. *Chem. Eur. J.* **1999**, 5, 2455.
- [28] Cousins, G. R. L.; Poulsen, S.-A.; Sanders, J. K. M. *Curr. Opin. Chem. Biol.* **2000**, 4, 270.
- [29] Sofia, M. J. *Med. Chem. Res.* **1998**, 8, 362.
- [30] Hirschmann, R.; Yao, W.; Cascieri, M. A.; Strader, C. D.; Maechler, L.; Cichy-Knight, M. A.; Hynes Jr., J.; van Rijn, R. D.; Sprengeler, P. A.; Smith III, A. B. *J. Med. Chem.* **1996**, 39, 2441. Hirschmann, R.; Hynes, Jr., J.; Cichy-Knight, M. A.; van Rijn, R. D.; Sprengeler, P. A.; Spoors, P. G.; Shakespeare, W. C.; Pietranico-Cole, S.; Barbosa, J.; Liu, J.; Yao, W.; Rohrer, S.; Smith, III, A. B. *J. Med. Chem.* **1998**, 41, 1382.
- [31] Sofia, M. J.; Hunter, R.; Chan, T. Y.; Vaughan, A.; Dulina, R.; Wang, H.; Gange, D. *J. Org. Chem.* **1998**, 63, 2802.
- [32] Nicolaou, K. C.; Xiao, X. Y.; Parandoosh, Z.; Senyei, A.; Nova, M. P. *Angew. Chem. Int. Ed. Engl.* **1995**, 34, 2289.
- [33] Silva, D. J.; Wang, H.; Allanson, N. M.; Jain, R. K.; Sofia, M. J. *J. Org. Chem.* **1999**, 64, 5926.
- [34] Sofia, M. J.; Allanson, N.; Hatzenbuehler, N. T.; Jain, R.; Kakarla, R. Kogan, N.; Liang, R.; Liu, D.; Silva, D. J.; Wang, H.; Gange, D.; Anderson, J.; Chen, A.; Chi, F.; Dulina, R.; Huang, B.; Kamau, M.; Wang, C.; Baizman, E.; Branstrom, A.; Bristol, N.; Goldman, R.; Han, K.; Longley, C.; Midha, S.; Axelrod, H. R. *J. Med. Chem.* **1999**, 42, 3193.
- [35] Wunberg, T.; Kallus, C.; Optaz, T.; Henke, S.; Schmidt, W.; Kunz, H. *Angew. Chem. Int. Ed.* **1998**, 37, 2503.
- [36] Kallus, C.; Opatz, T.; Wunberg, T.; Schmidt, W.; Henke, S.; Kunz, H. *Tetrahedron Lett.* **1999**, 40, 7783.
- [37] Silva, D. J.; Sofia, M. J. *Tetrahedron Lett.* **2000**, 41, 855.
- [38] Ghosh, M.; Dulina, R. G.; Kakarla, R.; Sofia, M. J. *J. Org. Chem.* **2000**, 65, 8387.
- [39] Wong, C.-H.; Hendrix, M.; Manning, D. D.; Rosenbohm, C.; Greenberg, W. A. *J. Am. Chem. Soc.* **1998**, 120, 8319.
- [40] Brill, W. K.-D.; De Mesmaeker, A.; Wendeborn, S. *Synlett* **1998**, 1085.
- [41] Brill, W. K.-D.; Tirefort, D. *Tetrahedron Lett.* **1998**, 39, 787.
- [42] Wendeborn, S.; De Mesmaeker, A.; Brill, W. K.-D.; Berteina, S. *Acc. Chem Res.* **2000**, 33, 215.
- [43] Ugi, I. *Angew. Chem. Int. Ed.* **1982**, 21, 810. Domling, A. *Curr. Opin. Chem. Biol.* **2000**, 4, 318. Domling, A.; Ugi, I. *Angew. Chem. Int. Ed.* **2000**, 39, 3168.
- [44] Sutherlin, D. P.; Stark, T. M.; Hughes, R.; Armstrong, R. W. *J. Org. Chem.* **1996**, 61, 8350.
- [45] Park, W. K. C.; Auer, M.; Jaksche, H.; Wong, C.-H. *J. Am. Chem. Soc.* **1996**, 118, 10150.
- [46] Lockhoff, O. *Angew. Chem. Int. Ed.* **1998**, 37, 3436.
- [47] Oertel, K.; Zech, G.; Kunz, H. *Angew. Chem. Int. Ed.* **2000**, 39, 1431.
- [48] Nilsson, U. J.; Fournier, E. J.-L.; Hindsgaul, O. *Bioorg. Med. Chem.* **1998**, 6, 1563.

- [49] Hummel, G.; Hindsgaul, O. *Angew. Chem. Int. Ed.* **1999**, 38, 1782.
- [50] Xu, W.; Springfield, S. A.; Koh, J. T. *Carbohydr. Res.* **2000**, 325, 169.
- [51] Bols, M. *Acc. Chem. Res.* **1998**, 31, 1, and references cited therein.
- [52] Byrgesen, E.; Nielsen, J.; Willert, M.; Bols, M. *Tetrahedron Lett.* **1997**, 38, 5697.
- [53] Lohse, A.; Jensen, K. B.; Lundgren, K.; Bols, M. *Bioorg. Med. Chem.* **1999**, 7, 1965.
- [54] Meldal, M.; St. Hilaire, P. M. *Curr. Opin. Chem. Biol.* **1997**, 1, 552. St. Hilaire, P. M.; Lowary, T. L.; Meldal, M.; Bock, K. *J. Am. Chem. Soc.* **1998**, 120, 13312.
- [55] Lampe, T. F. J.; Weitz-Schmidt, G.; Wong, C.-H. *Angew. Chem. Int. Ed.* **1998**, 37, 1707.
- [56] Kurokawa, K.; Kumihara, H.; Kondo, H. *Bioorg. Med. Chem. Lett.* **2000**, 10, 1827.
- [57] van Well, R. M.; Overkleeft, H. S.; Overhand, M.; Vang Carstenen, E.; van der Marel, G. A.; van Boom, J. H. *Tetrahedron Lett.* **2000**, 41, 9331.
- [58] Arya, P.; Dion, S.; Shimizu, G. K. H. *Bioorg. Med. Chem. Lett.* **1997**, 7, 1537.
- [59] Arya, P.; Kutterer, K. M. K.; Qin, H.; Roby, J.; Barnes, M. L.; Kim, J. M.; Roy, R. *Bioorg. Med. Chem. Lett.* **1998**, 8, 1127. Arya, P.; Kutterer, K. M. K.; Qin, H.; Roby, J.; Barnes, M. L.; Lin, S.; Lingwood, C. A.; Peter, M. G. *Bioorg. Med. Chem.* **1999**, 7, 2823.
- [60] Kutterer, K. M. K.; Barnes, M. L. Arya, P. *J. Comb. Chem.* **1999**, 1, 28.
- [61] Arya, P.; Kutterer, K. M. K.; Barkley, A. *J. Comb. Chem.* **2000**, 2, 120.
- [62] Barkley, A.; Randell, K. D.; Arya, P. *unpublished results*.
- [63] Ellgarrd, L.; Molinari, M.; Helenius, A. *Science*, **1999**, 286, 1882. A. Zapun, A.; Jakob, C. A.; Thomas, D. Y.; Bergeron, J. J. M. *Structure*, **1999**, R173.
- [64] McNeil, M. In *Genetics of Bacterial Polysaccharides*, Goldberg, J. B., Ed.; CRC Press; **1999**, 207-223.
- [65] Wittmann, V.; Seeberger, S. *Angew. Chem. Int. Ed.* **2000**, 39, 4348.