www.rsc.org/chemcomm

ChemComm

Automated carbohydrate synthesis to drive chemical glycomics

Peter H. Seeberger*

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA. E-mail: seeberg@mit.edu; Fax: +1 617 253 7929; Tel: +1 617 253 3986

Received (in Cambridge, UK) 17th October 2002, Accepted 27th November 2002 First published as an Advance Article on the web 11th February 2003

This feature article describes the development of the first automated solid-phase oligosaccharide synthesizer. A series of chemical challenges had to be addressed to accomplish this breakthrough and provide rapid access to oligosaccharides of biological significance. Accelerated synthesis of glycoconjugates promises to greatly impact the emerging field of glycobiology. Chemical glycomics uses synthetic carbohydrates and analogs to study their role in recognition, signal transduction pathways and other events of fundamental biomedical significance and shapes up to become the next major wave in biomedical research. The automated synthesis of a novel malaria vaccine candidate is discussed to illustrate the medical potential of chemical glycomics.

Three repeating biopolymers are responsible for most of the signal transduction processes in living organisms-nucleic acids, proteins and glycoconjugates. The role of nucleic acids as well as peptides and proteins has been extensively studied and many tools are available to elucidate their structure, function and interactions with other biomolecules. Genomics and proteomics are areas of much activity in the biochemical, biomedical, biotechnology and pharmaceutical areas. Fundamental breakthroughs in basic science and emerging new technologies have fuelled these developments. A host of new therapeutic targets is vigorously pursued as our understanding of nucleic acids and proteins has significantly improved. In recent years, a fairly well-defined picture of protein-protein interactions, protein-nucleic acid interactions and nucleic acidnucleic acid interactions emerged (Fig. 1) and has been exploited in a series of therapeutic approaches aiming to modify, enhance or disrupt these interactions. Carbohydrates,

Peter H. Seeberger received his B.S. in 1989 from the Universität Erlangen-Nürnberg, where he studied chemistry as a Bavarian government fellow. In 1990 he moved as a Fulbright scholar to the University of Colorado where he earned his Ph.D. in biochemistry under the guidance of Marvin H. Caruthers in 1995. After a postdoctoral fellowship with Samuel J. Danishefsky at the Sloan-Kettering Institute for Cancer Research in New York City he began his independent career at the Massachusetts Institute of Technology in January 1998. He is now the Firmenich Associate Professor of Chemistry at MIT and will assume a position as Professor of Organic Chemistry at the ETH Zuerich in April 2003. His research interests focus on the interface of chemistry and biology and in particular on the role of complex carbohydrates and glycoconjugates in information transfer in biological systems. Other interests include synthetic methodology, total synthesis, immunology, heterocyclic chemistry and microreactor design.



Fig. 1 Interactions of the three main biopolymers.

the third major class of biopolymers, have seen less interest from a drug development perspective as fundamental glycobiology is less well understood.

Two major technological breakthroughs helped genomics and proteomics to blossom. The sequencing of oligonucleotides and proteins is automated and allows for the composition of an unknown sample to be determined quickly and reliably, thereby providing a starting point for structure–function studies and the design of modifications. In addition, the synthesis of defined oligonucleotides¹ and peptides² can now be achieved by nonexperts in an automated fashion. Modified oligonucleotides, peptides and proteins have found use as research tools and therapeutic agents.

Glycomics-a term recently introduced to describe glycobiology and the interaction of carbohydrates with the other two major classes of biopolymers has trailed the explosive growth seen in genomics and proteomics. Carbohydrates are a highly complex and diverse class of biopolymers commonly found in nature as glycoconjugates (glycoproteins and glycolipids).³ Unlike oligonucleotides and peptides, carbohydrates are not just linear oligomers, but are often branched. The nine common monosaccharides found in mammalian cells can be combined in a dazzling variety of ways to form structures more diverse than those accessible with the twenty naturally occurring amino acids or four nucleotides. This structural complexity renders the isolation of pure carbohydrates from natural sources extremely difficult, when at all possible. No amplification methods analogous to the polymerase chain reaction (PCR) for DNA are available for carbohydrates. Until recently the identification of specific carbohydrate sequences responsible for a particular interaction had been hampered greatly by the unavailability of generally applicable sequencing methods. While this shortcoming has been addressed by several groups,⁴ sequencing remains a skill practised by a few expert laboratories.

Access to pure carbohydrates for biological, biochemical and biophysical studies relies on chemical or enzymatic synthesis. Given the structural complexity of carbohydrates, regio- and stereoselectivity of glycosylation reactions is the key challenge for the assembly of oligosaccharides.⁵ Synthetic chemists have developed increasingly powerful and versatile methods that have resulted in the assembly of ever more complex oligosaccharides and glycosaminoglycans. Still, the preparation of such structures is technically difficult, extremely time consuming and is carried out by a few highly specialized laboratories. Advances in enzymatic synthesis and in automated synthesis planning have alleviated some of these challenges,⁶ but still no automated synthesis method was available until recently to fuel the growing need for defined oligosaccharide structures as glycomics efforts gather steam. In this feature article, the development of the first automated oligosaccharide synthesizer is summarized.

Solid phase synthesis has proven extremely efficient for the assembly of peptides and oligonucleotides as it does not require purification after each reaction step, utilizes excess reagent to drive reactions to completion and lends itself to automation.^{1,2} A series of different approaches to solid phase oligosaccharide synthesis had been described and all critical aspects including the choice of synthetic strategy, differentially protected glycosylating agents, solid support materials, and linkers to attach the first monosaccharide to the support matrix were explored.⁷

Much progress concerning different aspects of solid phase oligosaccharide assembly had been made, but no generally applicable approach had evolved when we began our efforts towards an automated solid phase oligosaccharide synthesizer in 1998.⁷ We decided to utilize the acceptor bound approach and to attach the anomeric position of the reducing end sugar to the solid support. This strategy had been selected by many other groups as well since the glycosylating agent, the reactive species, can be used in excess to drive the glycosidic bond forming reaction to completion. Decomposed glycosyl donor can be simply removed by washing the resin.

Although a host of powerful glycosylating agents had been developed over the past century, we focused initially on the establishment of a novel, readily prepared, effective and versatile glycosylating agent. Inspired by nature's use of glycosyl nucleotide diphosphates in glycosyl transferase reactions, we focused our attention on glycosyl phosphate triesters (glycosyl phosphates) (Scheme 1). Glycosyl phosphates had



 α -phosphate

Scheme 1 Synthesis of glycosyl phosphate triesters from glycal precursors. R: protecting groups; R': Bn or butyl, R'': Ac, Bz, or Piv.

been reported previously to function as glycosylating agents but were considered too cumbersome to prepare and too unstable for routine use in glycosylation reactions.⁸ Differentially protected glycals proved convenient starting materials for a three step, one-pot procedure ensuring the rapid procurement of different glycosyl phosphates and dithiophosphates. Epoxidation with dimethyl dioxirane (DMDO) was followed by opening of the 1,2-anhydrosugar with a phosphoric acid diester to result in the installation of the anomeric phosphate. Depending on the nature of the solvent used for the opening of the epoxide, either the more reactive β -phosphates or the very stable somewhat less reactive α -phosphates were preferentially formed.⁹ Acylation of the C2 hydroxyl group produced the desired glycosyl phosphates in good to excellent yield following filtration through a short plug of silica gel. While glycosyl phosphates are relatively labile under acidic conditions, in pure form they can be stored for many months in the refrigerator (β -phosphates) and even at room temperature (α -phosphates).¹⁰

After ready access to glycosyl phosphates had been established, treatment with trimethyl silyltriflate (TMSOTf) or *tert*butyldimethyl silyl triflate (TBSOTf) resulted in high yielding glycosylation reactions, at low temperature and very short reaction times.^{8,10} A host of *cis*- and *trans*-glycosidic linkages was accessible in excellent yield and complete selectivity. In addition to hydroxyl groups, thiols, *C*-alkyl and *C*-aryl groups also functioned as nucleophiles in glycosylation reactions to form thioglycosides and *C*-glycosides (Scheme 2). These methods served well in natural products total synthesis.¹¹



Scheme 2 Glycosyl phosphates as versatile glycosylating reagents.

Besides establishing glycosyl phosphates as a novel glycosyl building block, attention was also paid to the development of protective groups that may be selectively removed from a complex carbohydrate in order to form branched oligosaccharides. Halobenzyl ether protective groups that can be removed by palladium catalyzed aryl amination followed by treatment with acid expanded the ether protecting group assortment.¹² The 2-azidomethyl benzoate can be removed *via* Staudinger reduction under neutral conditions and served well in the assembly of larger structures.¹³ A variety of different glycosyl phosphate building blocks equipped with a range of protecting group patterns was prepared for use on solid support.

With building blocks capable of efficient and selective glycosylations at hand, the linker, the crucial protective group connecting the first sugar and the solid support had to be selected. This linker has to withstand all chemistries throughout the assembly process, but has to succumb rapidly and efficiently to cleavage conditions. An alkene containing linker was chosen as a unique functional group otherwise not found in naturally occurring oligosaccharides. We introduced an octenediol linker that is readily accessible from cyclooctadiene, easily installed and allows for the cleavage of the oligosaccharides from solid support in form of *n*-pentenyl glycosides by olefin crossmetathesis (Scheme 3).¹⁴ *n*-Pentenyl glycosides offer much flexibility since they may serve as glycosylating agents and facilitate a host of reactions to connect oligosaccharides released from the solid support to surfaces, proteins or labels.¹⁵



Scheme 3 Installation and versatility of the octane diol linker.

Polystyrene resins such as Merrifield's resin (chloromethyl polystyrene) and macroporous polystyrene (*e.g.* ArgoporeTM) were readily equipped with the octenediol linker and served in the subsequent solid phase assembly of oligosaccharides.

Linear oligosaccharides were the first target of solid phase assembly. Initially, manual solid phase experiments established that the assembly of linear sequences of glucoses and mannoses using glycosyl phosphates and glycosyl trichloroacetimidates was quite straightforward.¹⁴ Still, lengthy washing and drying steps limited oligosaccharide elongation to incorporation of one building block per day.

The repetitive nature of the coupling and deprotection steps involved in oligosaccharide assembly suggested the use of an automated synthesizer to carry out all manipulations. Several practical matters had to be considered in designing an automated oligosaccharide synthesizer. The scale of the reaction should be flexible since anywhere from milligrams of oligosaccharide to grams of product should be accessible. The addition of each building block should be as fast as possible without compromising coupling or deprotection efficiencies. Finally, glycosylation reactions require low temperatures in contrast to amide bond and phosphate diester formation in peptide and nucleic acid synthesis. Rather than designing an entire new synthesizer we adapted an ABI 433 peptide synthesizer (Fig. 2).¹⁶ The saccharide building blocks are stored in cartridges that are advanced toward an intake needle in the order of incorporation. All other reagents and solvents are kept in argon-pressurized bottles and delivered to the reaction vessel via a series of valves. The reactions occur in a custom designed, double walled, coolable reaction vessel that may be vortexed. A chiller allows for the adjustment of the reaction temperature from -25 °C to +40 °C. The reaction vessel can be purged by forcing the solution through the frit at the bottom of the reaction vessel by argon pressure.

Based on the experience from solution phase experiments, we utilized a coupling cycle (Table 1) consisting of a coupling step using five equivalents of glycosylating agent that was repeated to ensure high coupling yields, followed by a deprotection step that was also repeated.¹⁶ After a series of washing steps the next cycle began. Similar coupling cycles were used for the incorporation of glycosyl trichloroacetimidates differing in the amount of activator and the reaction temperature. After the utility of this automated approach had been demonstrated on some linear oligosaccharides, the synthesis of biologically relevant carbohydrates was undertaken.

Initially we focused on the automated assembly of the phytoalexin elicitor β -glucan (Scheme 4). This class of molecules is important for signalling in plants and has served as a benchmark to evaluate new synthetic methodologies for carbohydrate synthesis.¹⁷ To facilitate the synthetic process we decided to address the challenge of constructing a branched structure by initially incorporating a disaccharide into a linear assembly scheme. Monosaccharide **2** and disaccharide building



Fig. 2 The first automated oligosaccharide synthesizr based on an ABI 433 peptide synthesizer.

Table 1 Coupling cycles used with phosphate donors

Step	Function	Reagent	Time/min
1	Couple	5 Equiv. donor and 5 equiv. TMSOTf	30
2	Wash	Dichloromethane	6
3	Couple	5 Equiv. donor and 5 equiv. TMSOTf	30
4	Wash	1:9 Methanol : dichloromethane	4
5	Wash	Tetrahydrofuran	4
6	Wash	3:2 Pyridine : acetic acid	3
7	Deprotection	2×20 Equiv. hydrazine (3 : 2 pyridine : acetic acid)	80
8	Wash	3:2 Pyridine : acetic acid	3
9	Wash	1:9 Methanol : dichloromethane	4
10	Wash	0.2 M Acetic acid in tetrahydrofuran	4
11	Wash	Tetrahydrofuran	4
12	Wash	Dichloromethane	6

block **3** were incorporated in an alternating fashion. Using our standard coupling cycle, hexasaccharide **4** was obtained after 9 h, followed by cleavage from the solid support. Synthesis of the dodecasaccharide **5** proved equally facile and was accomplished in 16 h after purification by HPLC.¹⁶

As is the case for the other repeating biopolymers, nucleic acids and peptides, the formation of internal deletion sequences,



Scheme 4 Automated solid phase synthesis of β -phytoalexin elicitor glucans using glycosyl phosphate building blocks.

caused by incomplete couplings during solid phase assembly, are the most difficult impurities to remove at the end of the synthesis. Inspired by the capping steps used routinely in oligonucleotide assembly, we explored a 'cap-tag' approach.¹⁸ In addition to placing a cap on all hydroxyl groups that did not react during the coupling step, the chemical moiety introduced to render the hydroxyl group unreactive, also served as a tag that destined all oligosaccharides marked for ready removal at the end of the synthesis. Two cap-tags were introduced (Scheme 5).



Scheme 5 Cap-tags to facilitate purification of the products of automated oligosaccharide syntheses.

An azide containing tag (A-Tag) was reduced after assembly was complete and could be removed by exposure to a scavenger resin. Alternatively, fluorous technology was employed by placement of a fluorinated silyl ether (F-Tag). Deletion sequences marked in this manner were readily removed by passing the reaction mixture after cleavage from the solid support through a column of fluorous silica gel. The purification of oligosaccharides assembled by using this cap-tag step were significantly simplified as demonstrated on the synthesis of a trisaccharide using short coupling times and suboptimal amounts of building blocks in order to result in couplings of low efficiency (75%). This example illustrated the power of the captag approach to massively simplify the purification of oligo-saccharide syntheses.¹⁸

Following a series of methodological advances and the initial proof of principal of the automated oligosaccharide synthesizer that accelerates carbohydrate assembly 50-100 fold, we intended to utilize this instrument to address biochemical and biomedical questions. In the context of specific target structures required for biological studies, the scope of the automated synthesizer was extended. Branched oligosaccharides had previously been accessed via incorporation of disaccharides (Scheme 6).16 A branched cap tetrasaccharide, found exclusively on the cell surface of the protozoan parasite Leishmania served as the initial synthetic challenge in accessing branched structures.¹⁹ The central mannose building block 9 was equipped with orthogonal acetate and levulinate ester protecting groups. After attachment to the solid support, the levulinate was selectively removed, followed by installation of glactose phosphate 10. Addition of two additional mannose trichloroacetimidates 11 furnished the desired tetrasaccharide 12 in just 9 h (Scheme 7).¹⁹ This synthesis demonstrated that branched structures are accessible by automation, using different temporary protecting groups while also showing that different glycosylating reagents can be incorporated into one automated synthesis cycle. The tetrasaccharide obtained from this synthesis was conjugated to a carrier protein and has shown



Scheme 6 Synthesis of a trisaccharide using a cap-tag step to facilitate purification.



Scheme 7 Automated synthesis of the Leishmania cap tetrasacharide.

promising results in initial animal experiments targeted at the development of a synthetic leishmaniasis vaccine.²⁰

Automated assembly of an anti-toxin malaria vaccine

Malaria has been detrimental to the development of many poor countries. This infectious disease afflicts currently 5% of the world's population, resulting in 100 million clinical cases and 3 million deaths per year. A vaccine against this disease would be of great importance as current treatments are facing increasingly resistant parasites. To illustrate the potential of chemical glycomics to fundamentally impact medicine, the automated oligosaccharide synthesizer was applied to the development of a synthetic anti-toxin malaria vaccine. Much of malaria's mortality is due to an inflammatory cascade initiated by a glycosylphosphatidylinositol (GPI) malarial toxin, released when parasites rupture the host's red blood cells. We demonstrated that anti-GPI vaccination can prevent malarial pathology in an animal model. Mice immunized with chemically synthesized GPI **13** (Fig. 3) bound to a carrier protein were



Fig. 3 The anti-toxin malaria GPI vaccine candidate.

substantially protected from death caused by malaria parasites.²¹ Between 60 and 75% of vaccinated mice survived, compared to a 0 to 9% survival rate for unvaccinated mice. While the solution-phase synthesis of **13** allowed for the procurement of much larger amounts of GPI than through isolation of natural GPI, faster access to **13** was important for the further development of anti-toxin malaria vaccines.

Ideally, the entire carbohydrate skeleton would be prepared on solid phase, but the α linkage between inositol and glucosamine presented too great a challenge to a fully automated approach at this time. Thus, GPI **13** was to be derived from disaccharide **15** prepared in solution and tetra-mannosyl fragment **14** rapidly prepared using automated solid-phase methodology (Scheme 8).²² The two fragments were to be joined to fashion a hexasaccharide for further elaboration to vaccine **13**. Tetrasaccharide **14a** was accessed by automated solid phase synthesis using the readily available trichloroacetimidate building blocks **11**, **16–18** (Scheme 9).

Cleavage of the octenediol linker using Grubbs' catalyst in an atmosphere of ethylene provided *n*-pentenyl tetrasaccharide **14a** in just 9 h. Conversion of **14a** into the corresponding tetrasaccharide trichloroacetimidate **14b** was followed by union with disaccharide **15** to afford a fully protected **20**. Further elaboration and removal of all protective groups furnished malarial toxin **13** in the manner previously outlined (Scheme 9).²²

Current work

Currently, much effort is directed at identifying a defined set of building blocks that will allow access to most naturally occurring structures. More efficient syntheses for the approximately 50 building blocks are being developed to enable access to the monomers employed in excess on solid support. The automated oligosaccharide synthesizer is used to assemble examples of all major classes of glycoconjugates including structures representative of *N*-linked and *O*-linked glycoproteins, glycolipids and glycosaminoglycans including heparin. The applications of these synthetic structures are manifold; rapid access to a host of oligosaccharides facilitates the creation of carbohydrate arrays to identify carbohydrate-protein, carbo-



Scheme 8 Synthetic plan for the semi-automated assembly of a anti malaria vaccine candidate.

hydrate-nucleic acid and carbohydrate-carbohydrate interactions in a highthroughput format. Signaling processes involving glycolipids and glycoproteins are being elucidated and we are beginning to unravel the role carbohydrates play in the fundamental processes of the immune response. Carbohydrate based vaccines are being pursued not only against tropical diseases but also against cancer and bacterial infections.

Conclusion

Our laboratory has developed the first automated oligosaccharide synthesizer by using glycosyl phosphates as readily accessible, efficient and selective building blocks, introducing a versatile octene diol linker and coupling protocols that allow for a simple two step coupling-deprotection cycle to be used for oligosaccharide assembly. A peptide synthesizer was reengineered to serve as automated oligosaccharide synthesizer and provides access to structures as large as dodecasaccharides about twenty fold faster than previous methods. Even branched structures are now accessible and a series of oligosaccharides of biological relevance have been prepared. The synthesis of an anti-toxin malaria vaccine and many other interesting carbohydrates has come within reach. While this synthesizer has greatly simplified access to a wide variety of naturally occurring carbohydrates some linkages are not yet accessible on this instrument and further development work remains to be done. Instruments for non-specialists are still several years away, but those skilled in organic synthesis may find the use of an automated oligosaccharide synthesizer to greatly speed up oligosaccharide assembly. Faster access to defined glycoconjugates has already begun to affect chemical glycomics. Further exciting advances in this rapidly evolving field are just around the corner.

Acknowledgements

I would like to thank all of my coworkers who, with their dedication and tremendous effort, made the development of the automated synthesizer possible. Their names can be found in the references to the original work but I would like to thank particularly O. J. Plante, R. B. Andrade, L. G. Melean, M. C. Hewitt, E. R. Palmacci, K. R. Love and D. A. Snyder. This work was supported by the Alfred P. Sloan Foundation, Glaxo-



Scheme 9 Semi-automated synthesis of the GPI malaria vaccine candidate.

SmithKline, the Mizutani Foundation for Glycoscience and the Petroleum Research Fund of the ACS.

Notes and references

- 1 M. H. Caruthers, Science, 1985, 230, 281.
- 2 R. B. Merrifield, Angew. Chem., Int. Ed. Engl., 1985, 24, 799.
- 3 Essentials of Glycobiology, ed. A. Varki, R. Cummins, J. Esko, H. Freeze, G. Hart and J. Marth, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1999.
- 4 For a review see: P. M. Rudd, G. R. Guile, B. Küster, D. J. Harvey, G. Oppendakker and R. A. Dwek, *Nature*, 1997, **388**, 205.
- 5 Modern Methods in Carbohydrate Synthesis, ed. S. H. Khan and R. A. O'Neill, Harwood Academic, Amsterdam, 1996.
- 6 P. Sears and C.-H. Wong, Science, 2001, 291, 2344.
- 7 Solid Support Oligosaccharide Synthesis and Combinatorial Carbohydrate Libraries, ed. P. H. Seeberger, Wiley, New York, 2001.
- 8 S. Hashimoto, T. Honda and S. Ikegami, J. Chem. Soc., Chem. Commun., 1989, 685.
- 9 O. J. Plante, R. B. Andrade and P. H. Seeberger, *Org. Lett.*, 1999, 1, 211.
- 10 O. J. Plante, E. R. Palmacci, R. B. Andrade and P. H. Seeberger, J. Am. Chem. Soc., 2001, 123, 9545.
- 11 H. Herzner, E. R. Palmacci and P. H. Seeberger, Org. Lett., 2002, 4, 2965.

- 12 O. J. Plante, S. L. Buchwald and P. H. Seeberger, J. Am. Chem. Soc., 2000, 122, 7148.
- 13 K. R. Love, R. B. Andrade and P. H. Seeberger, J. Org. Chem., 2001, 66, 8165.
- 14 R. B. Andrade, O. J. Plante, L. G. Melean and P. H. Seeberger, Org. Lett., 1999, 1, 1811.
- 15 B. Fraser-Reid, U. E. Udodong, Z. Wu, H. Ottosson, J. R. Merritt, S. Rao, C. Roberts and R. Madsen, *Synlett*, 1992, 927.
- 16 O. J. Plante, E. R. Palmacci and P. H. Seeberger, *Science*, 2001, 291, 1523.
- (a) R. Verduyn, P. A. M. van der Klein, M. Dowes, G. A. van der Marel and J. H. van Boom, *Recl. Trav. Chim. Pays-Bas*, 1993, **112**, 464; (b) B. P. Ossowski, A. Pilotti, P. J. Garegg and B. Lindberg, *Angew. Chem., Int. Ed. Engl.*, 1983, **22**, 793; (c) K. C. Nicolaou, N. Winssinger, J. Pastor and F. DeRoose, *J. Am. Chem. Soc.*, 1997, **119**, 449; (d) K. C. Nicolaou, N. Watanabe, J. Li, J. Pastor and N. Winssinger, *Angew. Chem., Int. Ed. Engl.*, 1998, **37**, 1559.
- 18 E. R. Palmacci, M. C. Hewitt and P. H. Seeberger, *Angew. Chem., Int. Ed.*, 2001, **40**, 4433.
- 19 M. C. Hewitt and P. H. Seeberger, Org. Lett., 2001, 3, 3699.
- 20 L. Schofield and P. H. Seeberger, unpublished results.
- 21 L. Schofield, M. C. Hewitt, K. Evans, M. A. Siomos and P. H. Seeberger, *Nature*, 2002, **418**, 785.
- 22 M. C. Hewitt, D. A. Snyder and P. H. Seebrger, J. Am. Chem. Soc., 2002, 124, 13434.