

ACS SYMPOSIUM SERIES **990**

Chemical Glycobiology

Xi Chen, Editor

University of California at Davis

Randall Halcomb, Editor

Gilead Sciences

Peng George Wang, Editor

The Ohio State University

**Sponsored by the
ACS Division of Carbohydrate Chemistry**



American Chemical Society, Washington, DC



ISBN 978-0-8412-7440-2

The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences—Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

Copyright © 2008 American Chemical Society

Distributed by Oxford University Press

All Rights Reserved. Reprographic copying beyond that permitted by Sections 107 or 108 of the U.S. Copyright Act is allowed for internal use only, provided that a per-chapter fee of \$36.50 plus \$0.75 per page is paid to the Copyright Clearance Center, Inc., 222 Rosewood Drive, Danvers, MA 01923, USA. Reproduction or reproduction for sale of pages in this book is permitted only under license from ACS. Direct these and other permission requests to ACS Copyright Office, Publications Division, 1155 16th Street, N.W., Washington, DC 20036.

The citation of trade names and/or names of manufacturers in this publication is not to be construed as an endorsement or as approval by ACS of the commercial products or services referenced herein; nor should the mere reference herein to any drawing, specification, chemical process, or other data be regarded as a license or as a conveyance of any right or permission to the holder, reader, or any other person or corporation, to manufacture, reproduce, use, or sell any patented invention or copyrighted work that may in any way be related thereto. Registered names, trademarks, etc., used in this publication, even without specific indication thereof, are not to be considered unprotected by law.

PRINTED IN THE UNITED STATES OF AMERICA

American Chemical Society Symposium Series Volume 990

Chemical Glycobiology

Editor(s): Xi Chen¹, Randall Halcomb², Peng George Wang³

¹University of California at Davis; ²Gilead Sciences; ³The Ohio State University

Publication Date (Print): September 01, 2008

Copyright 2008 American Chemical Society

ISBN13: 9780841274402

eISBN: 9780841221338

DOI: 10.1021/bk-2008-0990

Sponsoring Divisions: Division of Carbohydrate Chemistry

Table of Contents

<i>Copyright, Foreword</i>	pp i-v
<i>Preface</i> Xi Chen, Randall Halcomb, and Peng George Wang	pp ix-xiii

Chemical Synthetic Methods for Chemical Glycobiology

<i>De Novo Synthesis in Carbohydrate Chemistry: From Furans to Monosaccharides and Oligosaccharides</i> Xiaomei Yu and George O'Doherty	Chapter 1, pp 3-28
<i>Chemical Syntheses of Hyaluronic Acid Oligosaccharides</i> Lijun Huang, Xiaowei Lu, and Xuefei Huang	Chapter 2, pp 29-53
<i>Synthetic Efforts in Preparations of Components of the Bacterial Cell Wall</i> Mijoon Lee, Dusan Heseck, and Shahriar Mobashery	Chapter 3, pp 54-78
<i>New Glycosyl Thio-carboamino Peptides as New Tools for Glycobiology</i> Zbigniew J. Witczak	Chapter 4, pp 79-94

Chemoenzymatic Methods for Chemical Glycobiology

<i>Chemoenzymatic Synthesis of Sialosides and Their Applications</i> Hai Yu, Harshal A. Chokhawala, Shengshu Huang, and Xi Chen	Chapter 5, pp 96-122
<i>Convergent Chemoenzymatic Synthesis of Complex N-Glycopeptides</i> Lai-Xi Wang	Chapter 6, pp 123-50

Glycolipids

<i>Synthesis of Glycolipid Antigens</i> Suvarn S. Kulkarni and Jacquelyn Gervay-Hague	Chapter 7, pp 153-66
<i>Chemical Glycobiology of Glycosphingolipids</i> C. Xia, Y. Zhang, W. Zhang, W. Chen, J. Song, Q. Yao, Y. Liu, D. Zhou, G. De Libero, and P.G. Wang	Chapter 8, pp 167-94

Glycovaccines

Clinical Outcome of Breast and Ovarian Cancer Patients Treated after High-Dose Chemotherapy and Autologous Stem Cell Rescue with Theratope (STn-KLH) Cancer Vaccine Theratope (STn-KLH) Cancer Vaccine following Autologous Transplant

Leona A. Holmberg, Katherine A. Guthrie, and Brenda M. Sandmaier

Chapter 9, pp 197-215

Synthesis and Evaluation of Anticancer Vaccine Candidates, C-Glycoside Analogs of STn and PSA

Michel Weiwer, Fei Huang, Chi-Chang Chen, Xuejun Yuan, Kazuo Tokuzaki, Hiroshi Tomiyama, and Robert J. Linhardt

Chapter 10, pp 216-38

Tools for Chemical Glycobiology

An Automated Method for Determining Glycosylation and Site Diversity in Glycoproteins

Hyun Joo An, John Tillinghast, and Carlito B. Lebrilla

Chapter 11, pp 241-50

Chemical Approaches to Glycobiology

Nicholas J. Agard

Chapter 12, pp 251-71

Automated Solution-Phase Oligosaccharide Synthesis and Carbohydrate Microarrays: Development of Fluorous-Based Tools for Glycomics

Nicola L. Pohl

Chapter 13, pp 272-87

Polypeptide-Based Glycopolymers for the Study of Multivalent Binding Events

R. Maheshwari, S. Liu, B.D. Polizzotti, Y. Wang, and K.L. Kiick

Chapter 14, pp 288-305

Indexes

Author Index

pp 309

Subject Index

pp 311-25

Foreword

The ACS Symposium Series was first published in 1974 to provide a mechanism for publishing symposia quickly in book form. The purpose of the series is to publish timely, comprehensive books developed from ACS sponsored symposia based on current scientific research. Occasionally, books are developed from symposia sponsored by other organizations when the topic is of keen interest to the chemistry audience.

Before agreeing to publish a book, the proposed table of contents is reviewed for appropriate and comprehensive coverage and for interest to the audience. Some papers may be excluded to better focus the book; others may be added to provide comprehensiveness. When appropriate, overview or introductory chapters are added. Drafts of chapters are peer-reviewed prior to final acceptance or rejection, and manuscripts are prepared in camera-ready format.

As a rule, only original research papers and original review papers are included in the volumes. Verbatim reproductions of previously published papers are not accepted.

ACS Books Department

Preface

Chemical Glycobiology is an emerging sub-field of Chemical Biology, a multidisciplinary field that focuses on applying chemical techniques and tools to study and manipulate biology systems. Centered on understanding the biological roles of carbohydrates in nature, Chemical Glycobiology is diverse and involves disciplines ranging from chemical and enzymatic synthesis to assay development, through drug development and glycobiology.

Carbohydrates are biomolecules that are abundant in nature. They play key roles in a broad range of physiologically and pathologically important processes, including cellular recognition and communication, signal transduction, immune responses, bacterial and viral infection, development, and tumor metastasis. With the rapid advancement of basic scientific research in glycochemistry and glycobiology, the importance of carbohydrates in the management and prevention of many diseases caused by human infectious agents has been repeatedly acknowledged. Many promising pharmaceutical applications of carbohydrate-based materials are currently emerging.

An increasing number of chemists and biochemists are developing interests in the Chemical Glycobiology field. Chemical glycobiologists need to fully master the concepts of both glycochemistry and glycobiology in order to answer biological questions efficiently, either by directly probing/targeting living systems at the molecular level with small molecules or by applying more traditional bioorganic/medicinal chemistry or pharmacology/biochemistry approaches by synthesizing compound libraries and developing efficient assay methods (e.g., high-throughput screening approaches) to identify lead compounds. In order

to advance the Chemical Glycobiology field, it is essential for scientists from different disciplines to meet to exchange knowledge and share ideas and research findings. The 2-day Chemical Glycobiology Symposium in the 232nd American Chemical Society (ACS) National Meeting in San Francisco, California provided a symposium platform to stimulate discussion and collaboration. The symposium established a stronger connection of Glycochemistry and Glycobiology.

In order to share information with non-meeting participants and to attract more scientists to this rapidly advancing Chemical Glycobiology field, we assembled this ACS Symposium Series book. The book serves as an introduction to current advances in Chemical Glycobiology for students, scientists, and engineers, including those from industry, academia, and government, who are currently in the fields of chemical glycobiology, carbohydrate chemistry, glycobiology, natural product, and other carbohydrate-related research; it also introduces Chemical Glycobiology to those whose research will soon involve knowledge of this emerging field.

This symposium series book describes the current progress of research in Chemical Glycobiology, including the development of chemical synthetic methods for carbohydrates, recent advances in chemoenzymatic synthesis of complex oligosaccharides and glycoconjugates, synthesis and application of glycolipids, cancer vaccine development and its clinical evaluation, as well as new tools developed for Chemical Glycobiology.

The slow progress in the past towards the understanding of biological roles of carbohydrates was mainly due to the difficulties in obtaining homogeneous compounds and the lack of efficient analyzing tools. The importance of synthetic method development has been increasingly noticed. Novel synthetic methods, including chemical, enzymatic, and chemoenzymatic methods have emerged. The first four chapters report the advanced chemical synthetic methods for producing monosaccharides, oligosaccharides, glycosylated natural products, polysaccharides, peptidoglycans, and glycopeptide analogs. The first chapter by O'Doherty et al. provides a summary on de novo synthesis of monosaccharides, oligosaccharides, and glycosylated natural products. The second chapter by Huang et al. presents a comparison of conventional chemical approaches and the newly developed iterative one-pot strategy in producing hyaluronic acid oligosaccharides for biological studies. The

third chapter by Mobashery and co-workers reports their practice and methods used by others in synthesizing the components of bacterial cell wall including peptidoglycans as well as lipid I, II and IV. In Chapter 4, Witczak reviews the synthetic studies of his laboratory in developing glycosyl thio-carboamino peptides as small glycopeptide derivatives to understand the function of specific glycopeptide moieties in glycoproteins.

Two examples of recent advances in chemoenzymatic synthesis of complex oligosaccharides and glycoproteins are shown in Chapter 5 and Chapter 6. In Chapter 5, Chen and co-workers summarize their effort in using a highly efficient one-pot three-enzyme system in the synthesis of α 2,3- and α 2,6-linked sialoside libraries and their endeavor in developing a high-throughput screening platform to study substrate specificities of sialidases. L.-X. Wang describes in Chapter 6 the recent advances in the synthesis of large, homogeneous *N*-linked glycopeptides by endoglycosidase-catalyzed transglycosylation.

Glycosphingolipids have been found to stimulate immune effector cells and have potential in treating cancer and other diseases. Chapter 7 by Gervay-Hague et al. summarizes the current progress in synthesizing α -linked glycolipids and presents the advantage of applying glycosyl iodides in producing both *O*- and *C*-linked glycolipid analogs. P. G. Wang et al. reported in Chapter 8 their studies in the synthesis of α -Gal ceramide and isoglobotrihexosylceramide derivatives by varying the structures of lipid, modifying the terminal monosaccharide structures, and changing linkers to produce metabolically stable analogs. The difference of the obtained derivatives in stimulating immune effector cells is also reported.

Oligosaccharides, polysaccharides, and glycoproteins have been found as tumor-associated antigens. There are on-going studies on developing glycovaccines. Holmberg et al. report in Chapter 9 a successful clinical study of cancer immunotherapy by treating breast and ovarian cancer patients using a carbohydrate-protein conjugate vaccine Theratope (Sialyl Tn-keyhole limpet hemocyanin). Chapter 10 by Linhardt and co-workers presents the synthesis and evaluation of *C*-glycoside derivatives of sialyl Tn and polysialic acid (PSA) as potential non-hydrolyzeable glycovaccines.

The last four chapters report the currently developed tools and methods for Chemical Glycobiology. Lebrilla et al. in Chapter 11

described an automated Mass Spectrometry-based method for determining the site of glycosylation and the heterogeneous oligosaccharide structures in glycoproteins. Agard summarizes in Chapter 12 the unique metabolic engineering approach to study protein posttranslational modifications by introducing a chemical handle to the carbohydrate moieties in glycoconjugates of live cells and animals. In order to achieve the ultimate automated synthesis of carbohydrates, Pohl focuses on the development of fluororous tag-based strategy. The fluororous tag incorporated into the synthetic scheme also allows the direct transfer of the carbohydrate products to microarrays for high-throughput screening of carbohydrate-binding proteins. The strategy and application are discussed in Chapter 13. The interactions of carbohydrate and proteins are multivalent binding events. Development of polymers with multivalent carbohydrate units is curritical to truly understand these important events. The final chapter by Kiick and co-workers describes the design, synthesis, and evaluation of polypeptides of defined secondary structure and size as scaffolds to present galactose residues. It provies a novel platform for studying the multivalent interactions of carbohydrate and protein.

Enjoy reading and join us to the exciting field of Chemical Glycobiology!

Acknowledgements

We are pleased to acknowledge the financial support of the National Institute of Allergy and Infectious Diseases (NIH R13AI073074), Wyeth Research, Genzyme Corporation, and the ACS Division of Carbohydrate Chemistry for the Chemical Glycobiology Symposium. We are grateful for the dedication and the support from the Office of Research and the Business Office of the Department of Chemistry at the University of California, Davis. The help from the administrative personnels of the ACS Division of Carbohydrate Chemistry in organizing the symposium is greatly appreciated. We would also like to sincerely thank our symposium speakers, meeting participants, and especially the authors and peer reviewers for their contributions. The

publication of this book will not be possible without the dedication of personnels from the ACS Books Department.

Xi Chen

Department of Chemistry
University of California, Davis
One Shields Avenue
Davis, CA 95616

Randall Halcomb

Gilead Sciences
333 Lakeside Dr.
Foster City, CA 94404

Peng George Wang

Departments of Biochemistry and Chemistry
The Ohio State University
876 Biological Sciences Building
484 West 12th Avenue
Columbus, OH 43210

Chapter 1

De Novo Synthesis in Carbohydrate Chemistry: From Furans to Monosaccharides and Oligosaccharides

Xiaomei Yu and George O'Doherty*

Department of Chemistry, West Virginia University,
Morgantown, WV 26506

*Corresponding author: George.ODoherty@mail.wvu.edu

In an effort to address the medicinal chemist's need for new synthetic methods for the preparation of unnatural carbohydrates, new de novo methods for carbohydrate synthesis have been developed. These routes use asymmetric catalysis to set the sugar absolute stereochemistry, a palladium-catalyzed glycosylation reaction to stereoselectively control the anomeric center, and subsequent diastereoselective post-glycosylation to install the remaining sugar stereocenters. The utility of this method has been demonstrated by the syntheses of several classes of mono-, di- and tri-saccharides.

Introduction

Vital to the discovery of new carbohydrate based therapeutic agents is an understanding of the critical role that carbohydrates play in the biological activity of medicinally active natural products. It has long been recognized that

the sugar portions of the natural products play a crucial role in the mechanism of action for many drugs (e.g., target binding, solubility, tissue targeting, membrane transport) (1). For instance, the corresponding aglycons of natural products are often devoid of activity. Since the initial discovery of biologically active glycosylated natural products, medicinal chemists have desired the ability to vary the carbohydrate structures of natural products, to elucidate the role the carbohydrate plays in a given natural product, as well as to prepare analogues with improved activities. While nature uses a diverse array of carbohydrate structures in these natural products only a limited number of sugar isomers are provided in usable quantities for SAR studies, ultimately affecting the ability of medicinal chemists to install rare/unnatural sugars. Clearly, synthetic alternatives are required to address this medicinal chemistry need.

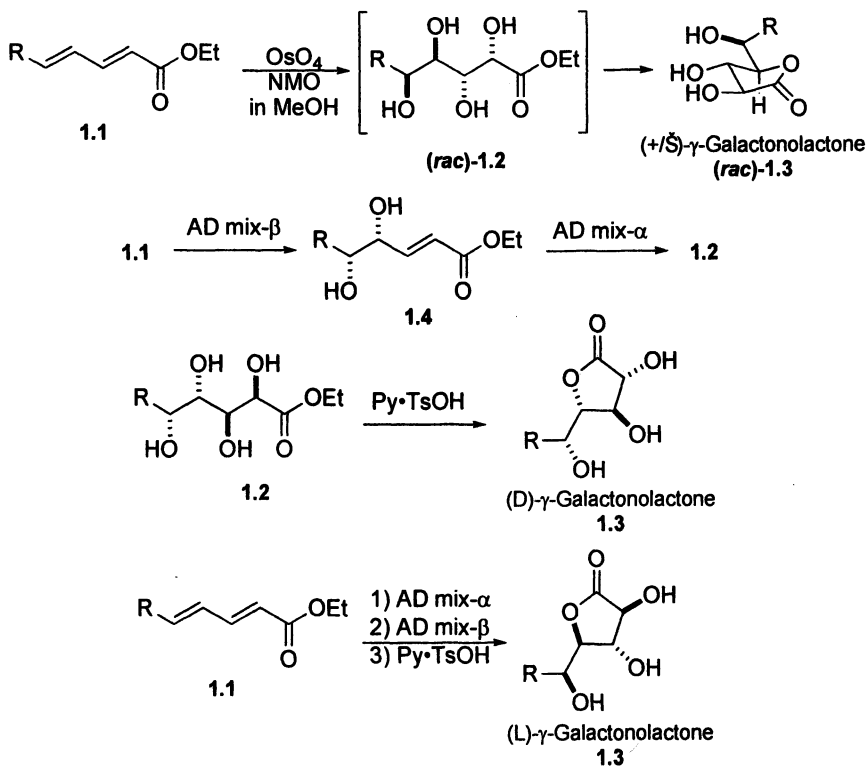
An alternative approach would be to avoid the carbohydrate starting materials and assemble the desired sugar functionality and stereochemistry from simple achiral starting materials (2-4). This *de novo* synthetic approach should enable the medicinal chemist to install a much broader range of carbohydrate structures amenable to SAR-investigation. This *de novo* approach stands in stark contrast to using known sugars as starting materials where variability is limited and protecting group chemistry predominates. The term “*de novo*” in our approach is used to describe the way we install the desired functionality and chirality within each sugar, as opposed to the traditional carbohydrate approach, which starts with known sugars. Herein we describe our development of this methodology and its on going application towards the synthesis of various carbohydrate structural motifs.

De Novo Synthesis of Monosaccharides

Over the years, considerable efforts have been made toward the development of new synthetic routes to monosaccharides. Of particular interest is the *de novo* preparation of these carbohydrates (i.e., from achiral starting materials using asymmetric catalysis). While there are many uses of the term “*de novo*” in carbohydrate chemistry (5), herein the term *de novo* asymmetric synthesis refers to the use of catalysis for the asymmetric synthesis of carbohydrates from achiral compounds (*vide infra*).

The *de novo* enantioselective synthesis of the hexoses stands as a challenge to asymmetric catalysis (3). Despite some germinal efforts toward the hexoses, notably by Masamune/Sharpless (epoxidation) (2), Danishefsky (Diels-Alder) (6), Johnson/Hudlicky (enzymatic desymmetrization) (4) and Wong/Sharpless (osmium/enzyme) (7), there still does not exist a practical, non-enzymatic route to all the hexoses. More recently this challenge has also been taken up by MacMillan (iterative aldol strategy) (8) and White (allylic oxidation) (9).

For the last ten years we have endeavored to develop practical methods for the de novo synthesis of the hexoses. These efforts have resulted in the discovery of two orthogonal approaches to pyrano-hexoses with variable C-6 substitution. These approaches being an iterative dihydroxylation strategy (10-12) to primarily *galacto*-sugars and an Achmatowicz strategy (3a, 3c, 13) to primarily *manno*-sugars.



Scheme 1. De novo approach to galacto-sugar γ -lactones from dienoates

Galactono-lactones

Of our two approaches, the iterative asymmetric dihydroxylation of dienoates (Scheme 1) is the most efficient in terms of steps (1 to 3 steps). For instance, dienoates, like ethyl sorbate (1.1 R = CH₃), react under the Upjohn conditions (OsO₄/NMO) (14) to give racemic γ -galactono-lactones in only one

pot, via a sequential bis-asymmetric dihydroxylation (i.e., **1.1** to (*rac*)-**1.3** via (*rac*)-**1.2**) (*10b*). Key to this discovery is the recognition that the more electron-rich γ,δ -double of dienoate **1.1** reacts first to form diol **1.4**, which once formed reacts again in a diastereoselective fashion ($> 4:1$) to give the tetrol **1.2**. When tetrol **1.2** is formed in a polar protic solvent like MeOH, it undergoes a base-catalyzed (NMM) lactonization to give the γ -galactono-lactone (*rac*)-**1.3**. Alternatively, the initial dihydroxylation can be run with the Sharpless reagent to give diol **1.4**, which when dihydroxylated with OsO₄/NMO in MeOH gives the galactono-lactone **1.3** in high enantioexcess. By performing the second dihydroxylation with the enantiomeric Sharpless reagent, this allows for a highly stereoselective 3-step synthesis of the γ -galactono-lactones **1.3** with diverse C-6 substitution and near perfect enantio- and diastereocontrol.

By taking advantage of some palladium π -allyl chemistry, this approach can be employed towards the synthesis of pyranoses (Scheme 2). That is to say, the C-4 hydroxyl group can be selectively removed or protected with a *p*-methoxyphenol, which enforces six-membered ring lactonization. Thus, by converting the initial diol **2.1** into a cyclic carbonate **2.3** and treatment with palladium(0) a π -allyl intermediate **2.4** can be generated and trapped with various nucleophiles (e.g., H, OAr). When the C-4 hydroxyl group is replaced with an arylether (X = OAr) increase diastereoselectivity is seen in the dihydroxylation of **2.5** under the Upjohn conditions (OsO₄/NMO). In contrast, when the hydroxyl group is replaced with hydrogen no diastereoselectivity is observed. This loss of stereocontrol can be solved by use of the Sharpless reagent, which can selectively provide either the C-4 deoxysugars (**2.7** and **2.9**, X = H). Similarly, this approach has also been selectively applied towards the synthesis of C-4 fluorodeoxysugars (**2.11**) (*15*).

Achmatowicz Approach

The second and possibly more general approach to the hexopyranoses relies on the use of the Achmatowicz rearrangement (Scheme 3) (*3*, *13*), which is an oxidative rearrangement of furfuryl alcohols to 2*H*-pyran-3(6*H*)-ones (pyranone). These furfuryl alcohols in turn can be produced via asymmetric synthesis (Scheme 4). Using this approach, we have succeeded in developing a short route that is flexible enough for the synthesis of three of the eight possible diastereomeric hexoses as well as several deoxy- (*16*) and 4- and 6-substituted aminosugars (*17*). In practice, these Achmatowicz products can be selectively acylated to give primarily the axial isomer (e.g., 3.2- α -D). With the C-1 and C-5 stereocenters set, the C-4 allylic alcohol stereochemistry can be installed by a highly diastereoselective Luche reduction (**3.2** to **3.3**). In a similarly highly stereoselective

reaction, the *manno*-stereochemistry of **3.4** can be installed by an Upjohn dihydroxylation (*14*). Finally two other diastereomeric sugars (*talo*-**3.6** and *gulo*-**3.8**) can be prepared by Mitsunobu inversion and dihydroxylation of **3.3**.

These furfuryl alcohols can be produced in either enantiomeric form via asymmetric catalysis. Our preferred method for the asymmetric synthesis of these furan alcohols **4.2** is by the highly enantioselective Noyori reduction of achiral acylfurans **4.1** (Scheme 4). Alternatively furfuryl alcohols like **4.4** can be prepared by the Sharpless asymmetric dihydroxylation of vinylfuran **4.3**. Key to this later approach was the recognition that vinylfuran **4.3** could be made by a Petersen olefination reaction.

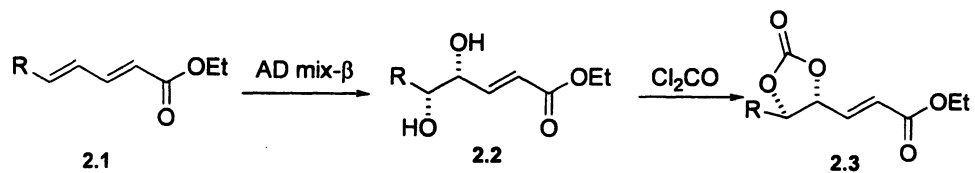
Papulacandins

The papulacandins are an important group of antifungal antibiotics extracted from the fermentation broth of *Papularia sphaerosperma* (*18*), which inhibit 1,3- β -D-glucan synthase. We have successfully used both the Achmatowicz and dienone oxidation approaches to a class of *C*-aryl glycoside natural products (e.g., the papulacandin ring system). These two routes both derive their asymmetry from a Sharpless AD-mix dihydroxylation (Scheme 5) and have resulted in the preparation of four sugar diastereoisomers of the papulacandins (*gluco*-, *manno*-, *allo*- and *galacto*-). The *galacto*-isomer was formed as a mixture both its *pyrano*-**5.4** and *furano*-**5.5** forms.

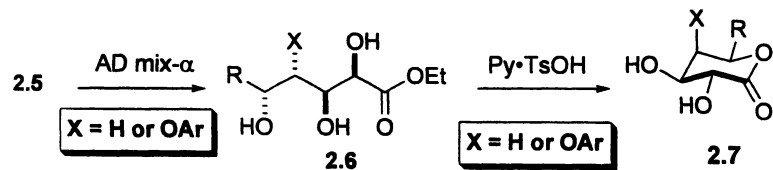
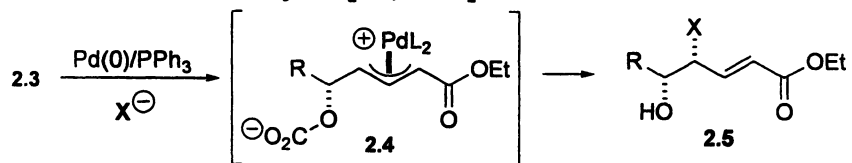
Using the iterative dihydroxylation strategy on dienone **5.1** the *galacto*-stereochemistry can be installed, as in tetraacetate **5.3**. A one-pot deprotection/spiroketalization provides a 2:1 mixture of the *galacto*-papulacandins in both *pyrano*-**5.4** and *furano*-**5.5** forms. Ultimately we hoped that a *galacto*-pyranosyl **5.4** would lead to the formation of disaccharide papulacandins via an S_N2 -type inverse glycosylation at *C*-4.

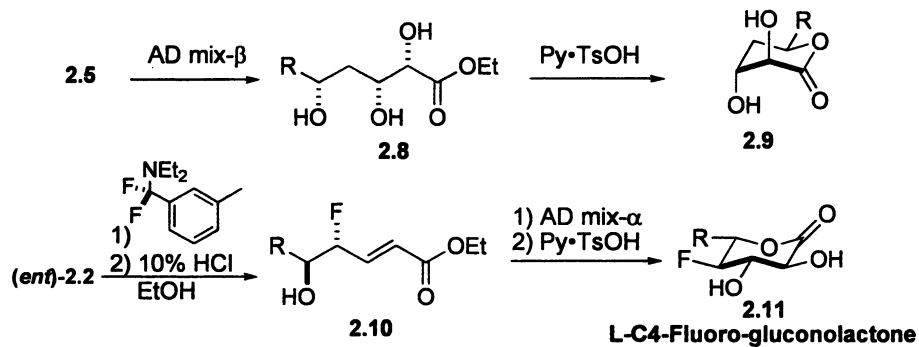
In contrast, using the Achmatowicz approach, only pyranoses were formed. This approach began with the Sharpless dihydroxylation of achiral vinylfuran **5.6** to install the *C*-5 D-absolute stereochemistry as in **5.7**. The furfuryl alcohol **5.7** can be stereoselectively converted into the α -spiroketal **5.8** by Achmatowicz oxidation, spiroketalization, Luche reduction and TBS-protection. Upjohn dihydroxylation of **5.8** was used to prepare both the *manno*-**5.9** and *allo*-**5.11** isomers, with the *manno*-isomer being formed as the major isomer (4:1) (*14*).

Using an oxidation/reduction sequence the *gluco*-papulacandin **5.11** was also produced from **5.8**. The dihydroxylation product of **5.8** was selectively protected at *C*-3 and the *C*-2 axial alcohol was oxidized (Dess-Martin) and reduced (DibalH) to install the *gluco*-stereochemistry, which upon TBS-deprotection give the *gluco*-isomer **5.11**.

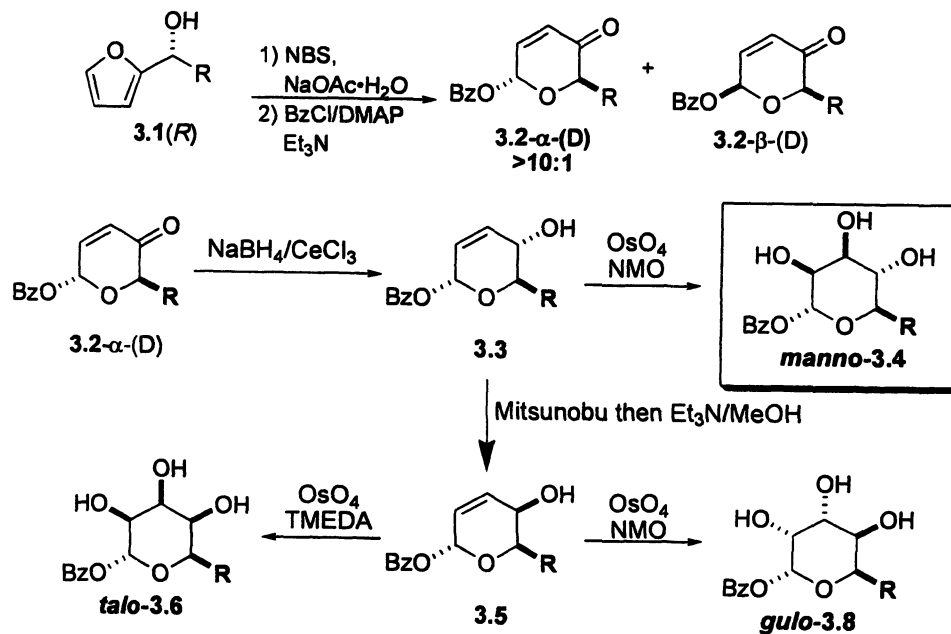


R = Alkyl, CH₂OR, or CH₂NHCbz

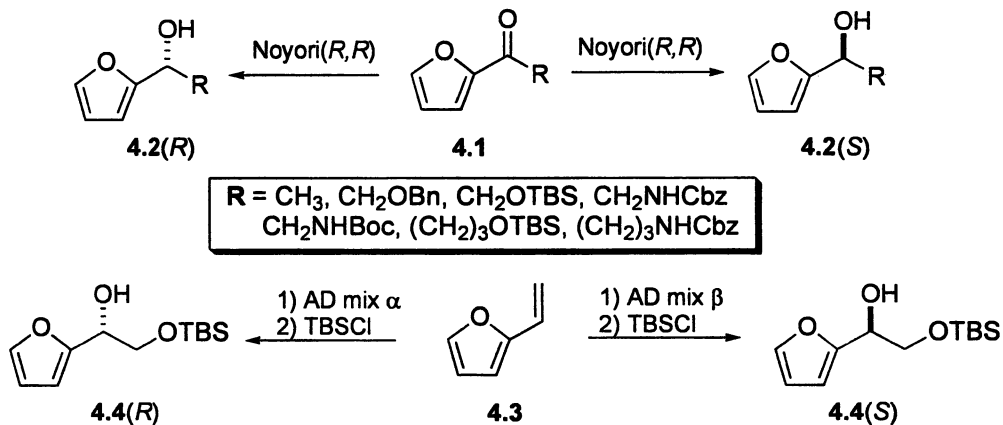




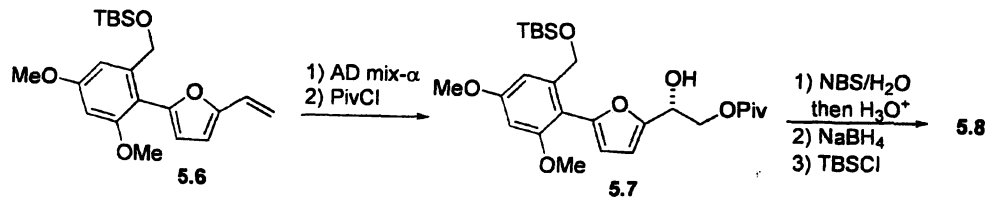
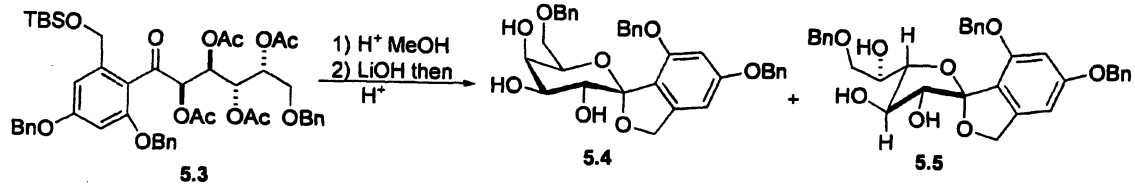
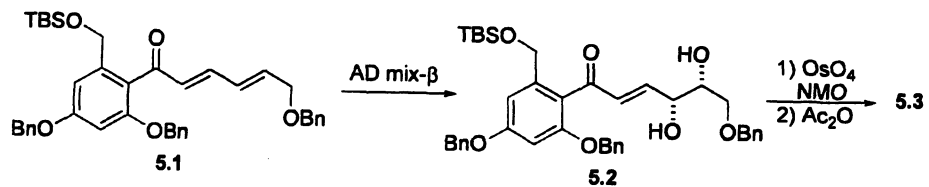
Scheme 2. De novo approach to galacto-sugar δ -lactones from dienoates

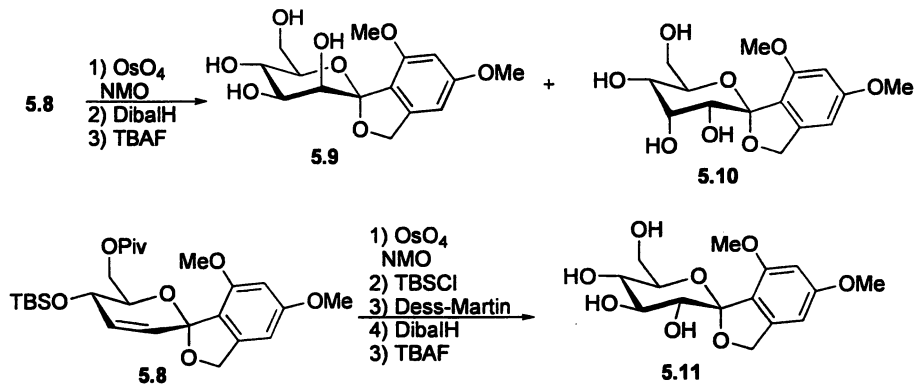


Scheme 3. Achmatowicz approach to carbohydrates



Scheme 4. Asymmetric synthesis of chiral furfuryl alcohols





Scheme 5. Papulacandins from dienones and vinylfurans

De Novo Synthesis of Natural and Unusual Oligosaccharides

This de novo approach to carbohydrates has also been applied to oligosaccharides (Schemes 9 to 14). Key to the success of this approach is the development of a mild palladium-catalyzed glycosylation (Scheme 6) in combination with the use of the previously developed highly stereoselective enone reduction and dihydroxylation reaction (Scheme 3) as post-glycosylation transformation for the installation of *manno* stereochemistry (19).

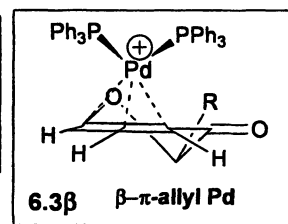
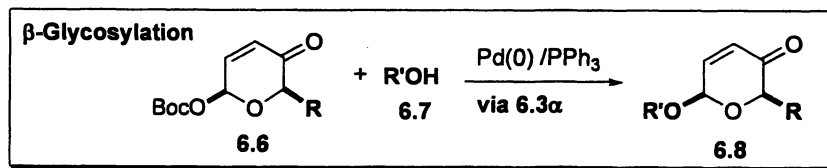
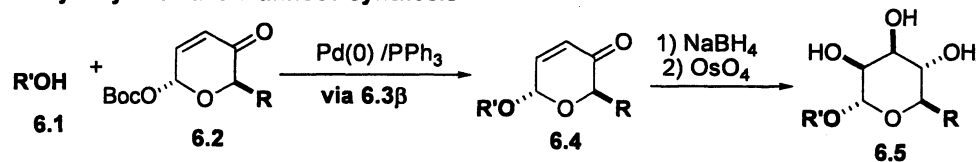
Glycosylation of Alcohols

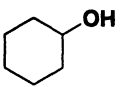
In an effort to develop a palladium-catalyzed glycosylation reaction, we initially tried to generate a Pd π -allyl intermediate from the triacetoxyl glucal, but were unsuccessful at ionizing the C-3 equatorial acetate under typical Pd π -allyl forming conditions. We next investigated pyrans like **3.3** in the reaction hopeful that a leaving group at the C-1 position would be more conducive for Pd π -allyl formation. Moving the leaving group to C-1 proved to be the solution for generating both diastereomeric Pd π -allyl intermediates from either C-1 axial or equatorial carboxylates. Evidence for this ionization can be seen when the Pd intermediates are reacted with the traditional Pd π -allyl nucleophiles (i.e., various anions of malonates and phenols) products are formed with net retention of stereochemistry. Unfortunately, in our hands the same Pd π -allyl intermediates did not react with the simplest of alcohols. Thus, we decided to try to generate the presumably more electrophilic Pd π -allyl intermediate **6.3** from the corresponding pyranones **6.2**. To our delight these two changes resulted in a general glycosylation reaction (Scheme 6).

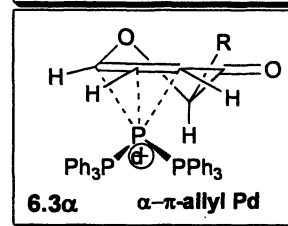
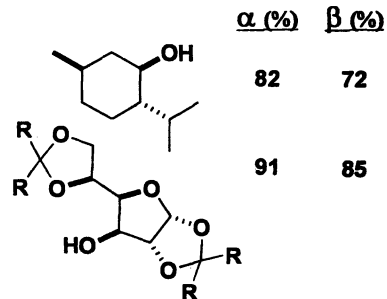
As outlined in Scheme 6, the Pd(0) catalyzed glycosylation reaction is both general and stereospecific (20, 21). The reaction occurs rapidly and in high yields for both the α -(**6.2** to **6.4**) and β -(**6.6** to **6.8**) systems. While carboxylate leaving groups also work, the *t*-butoxycarbonate group is critical for the successful implementation of this reaction with alcohol nucleophiles. When the *t*-butoxycarboxy group is replaced with a benzoyl or pivaloyl group, the palladium-catalyzed glycosylation reaction is significantly slower.

This reaction has great potential for preparing various D- and L-sugars because the starting 6-*t*-butoxycarboxy-2*H*-pyran-3(6*H*)-ones (**7.3** and **7.6**) can easily be prepared from optically pure furfuryl alcohols (either (*R*) or (*S*) form) (**22**) by a two-step procedure related to the synthesis of **3.2** (Scheme 7). Depending on the reaction temperature of the second step, the acylation reaction can selectively give the α -Boc pyranones **7.3 α** and **7.7 α** at -78 °C; whereas at room temperature, a 1:1 ratio of the α - and β -Boc protected enones were produced. Thus, this procedure can be used to prepare multigram quantities of both α - and β -pyranones in either enantiomeric form.

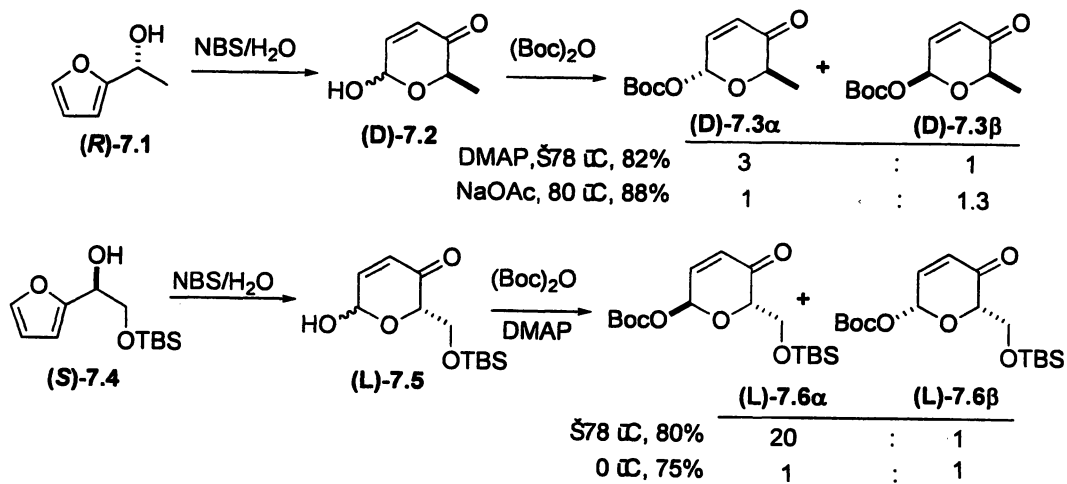
α -Glycosylation and Mannose synthesis



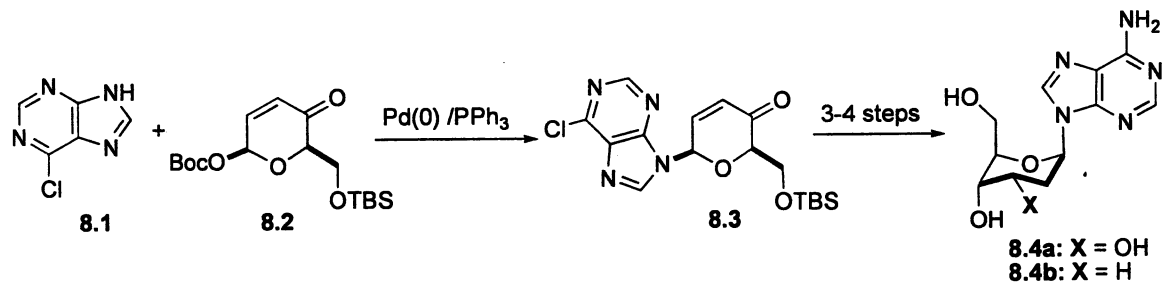
<u>ROH =</u>	<u>α (%)</u>	<u>β (%)</u>
MeOH	87	85
BnOH	89	85
<i>t</i> -BuOH	70	75
	88	80



Scheme 6. Palladium-catalyzed α / β -selective glycosylation



Scheme 7. Enantio- and diastereoselective pyranone synthesis



Scheme 8. Synthesis of β -N-glycosides

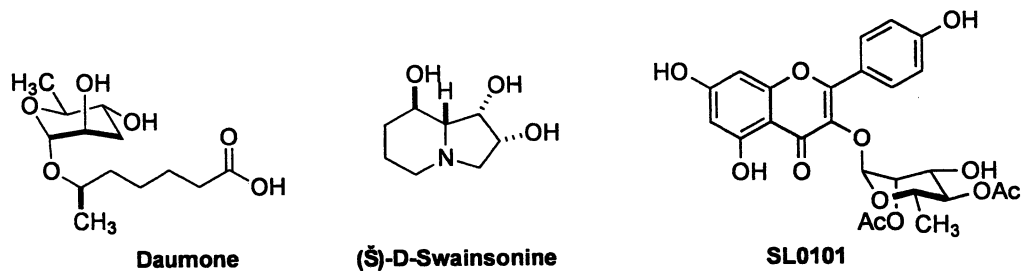


Figure 1. De Novo Synthesis of Carbohydrate-Based Natural Products

Synthesis of β -linked *N*-glycosides

This reaction has great potential for the preparation of various D- and L-sugars in either the α - or β -configuration. For instance, we have found that nitrogen nucleophiles like 6-chloro-9*H*-purine **8.1** also work well in this palladium-catalyzed glycosylation (**8.1** + **8.2** to **8.3**). This can be seen in our syntheses of homo-adenosine and homo-deoxy adenosine (**8.4a/b**, which have been sent to NCI for testing) (23).

De Novo Synthesis of Monosaccharide Natural Products

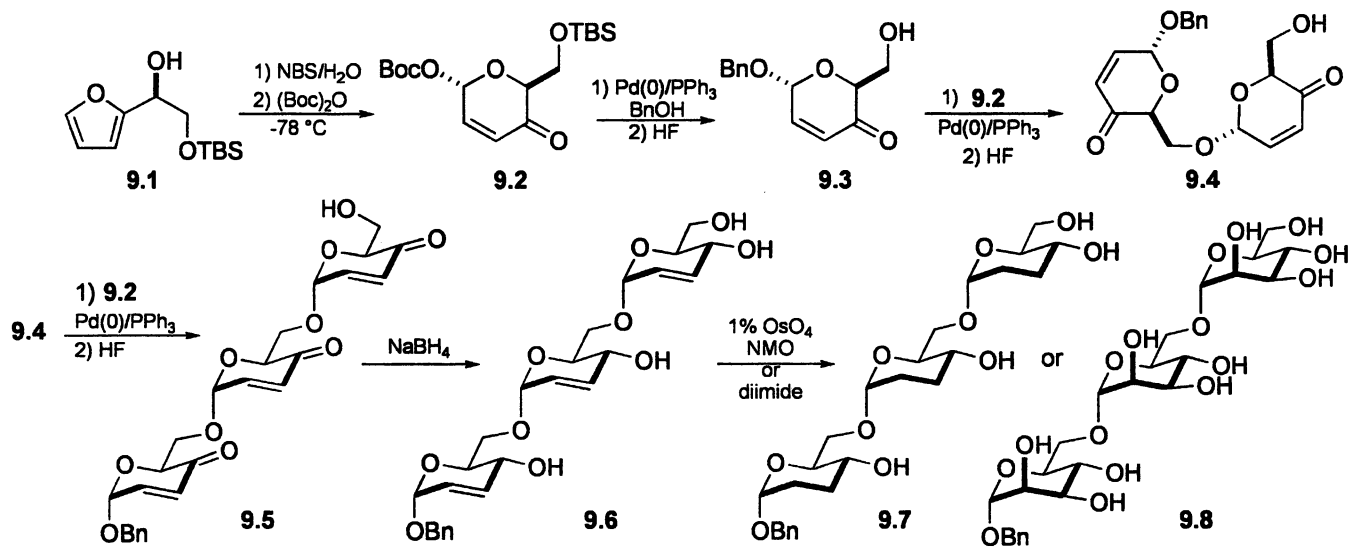
We have used the same de novo strategy to prepare other carbohydrate-based natural products. This list includes the pheromone daumone, anticancer/antiviral agent swainsonine, the Rsk inhibitor SL0101 (Figure 1), as well as the trisaccharide digitoxin. These routes have provided significant amounts of material for biological testing and SAR studies, allowing access to analogues that are not readily available by traditional carbohydrate methods (e.g., regioisomers, diastereomers and enantiomers).

Oligosaccharide Synthesis

Possibly the most powerful application of this π -allyl Pd-catalyzed glycosylation reaction will be in the generation of biologically important oligosaccharides. The most interesting of these will be the unnatural forms. It is worth noting that in terms of total steps the trisaccharides we so far have prepared are invariably assembled in fewer steps than more traditional carbohydrate routes, primarily due to the extensive use of protection/deprotection strategies in the latter. Needless to say, a traditional carbohydrate approach would also not allow for the incorporation of non-6-deoxy-L-sugars or other unnatural possibilities.

1,6-Oligosugars

As outlined in Scheme 9, the 1,6-*manno*-trisaccharide **9.8** was prepared from enone **9.2** with a C-6 OTBS group. The synthesis was accomplished by a simple glycosylation/TBS-deprotection strategy to prepare first a disaccharide **9.4** and then trisaccharide **9.5** by repeating the two steps. The *manno* stereochemistry of **9.8** was installed by a one pot NaBH₄ reduction and one-pot dihydroxylation of **9.5**. Thus, in only these two steps, nine stereocenters were



Scheme 9. Synthesis of 1,6-manno-trisaccharides

set. Similarly the 1,6-tri-dideoxytrisaccharide **9.7** was prepared by switching a diimide reduction for the dihydroxylation reaction (**24**).

1,4-Oligosaccharides

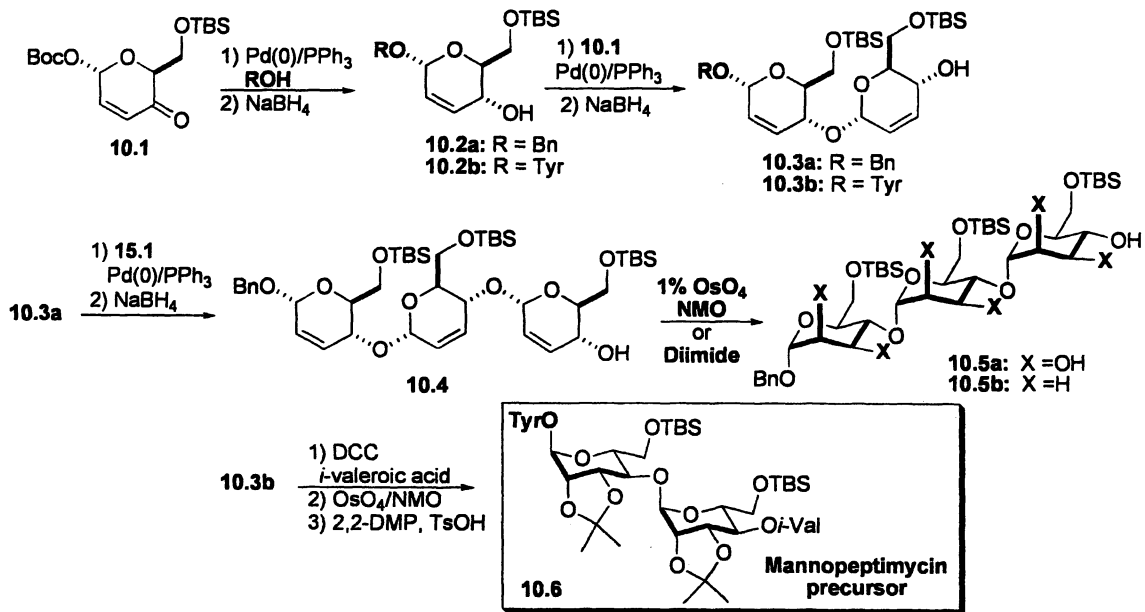
Outlined in Scheme 10 is how we used a similar strategy for the preparation of both the 1,4-*manno* trisaccharide **10.5a** and the 2,3-deoxy trisaccharide **10.5b**, as well as one of the glycosylated amino acid **10.6**, which is required for Mannopeptimycin, disaccharide (**25**). Our route to the 1,4-linked trisaccharides started with enone **10.1**, which was prepared from the appropriate furfuryl alcohol in 2 steps. Glycosylation and NaBH₄ reduction of **10.1** provides **10.2** with a C-4 alcohol ready for a glycosylation. Repeating this process converts **10.2** to **10.3** and then again trisaccharide **10.4**. Finally, a tris-dihydroxylation reaction installed the remaining C-2/C-3 hydroxyl groups with *manno* stereochemistry. Key to the success of this sequence were the diimide reduction and highly stereoselective dihydroxylation reactions, which in the case of the dihydroxylation, installed six stereocenters in one transformation (**26**).

A de novo approach to the cardiac glycosides

Recently several cardiac glycosides, like the trisaccharide digitoxin and the monosaccharide oleandrin, were found to be potential anticancer agents (Figure 2) (**27**). In vitro studies have shown that a sub-cardiotoxic dose of digitoxin has quite a strong anticancer effect (**28**). Detailed mechanism of action studies have shown that pronounced apoptoses were induced by digitoxin in several different cell lines. Similar antitumor activity was seen for the related cardiac glycosides digoxin and gitoxin as for digitoxin (**29**). As with digitoxin, all of these cardiac glycosides were found to be more active than their corresponding aglycons, establishing the importance of the carbohydrate to cytotoxic activity.

To better elucidate the relationship between carbohydrate structure and anticancer activity the Thorson group prepared a neoglycoside library of digitoxin (**30**, **31**). This study discovered several monosaccharide neoglycoside analogues of digitoxin with increased activity against several cancer cell lines (**30**). To delineate the role of the sugar versus the methoxyamine linkage used by Thorson, we decided to use our glycosylation methodology for the synthesis of unnatural glycosylated digitoxin analogues (Scheme 11).

At the outset we felt that this de novo approach could identify novel agents with improved selectivity toward tumor cells and believed the best way to find these new structures was to initiate structure-activity studies by manipulating functionality at the sugar portion of the molecules. The carbohydrate portion of the



Scheme 10. Synthesis of 1,4-linked manno-trisaccharides

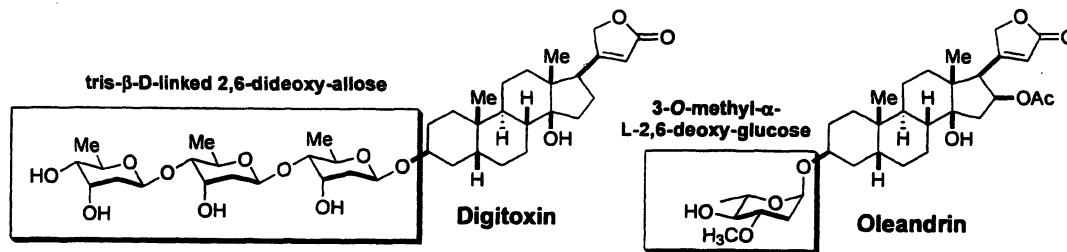
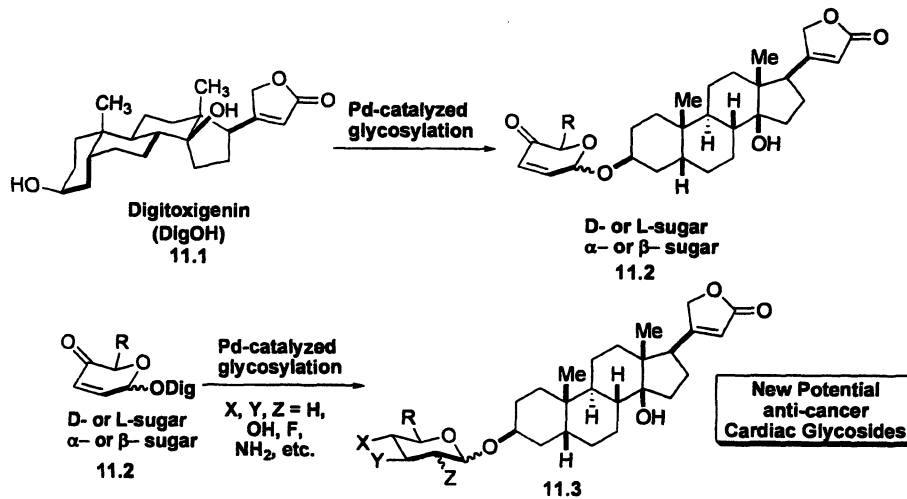


Figure 2. D- and L-Cardiac Glycosides with anti-tumor



Scheme 11. From natural products to aglycons to new bioactive leads

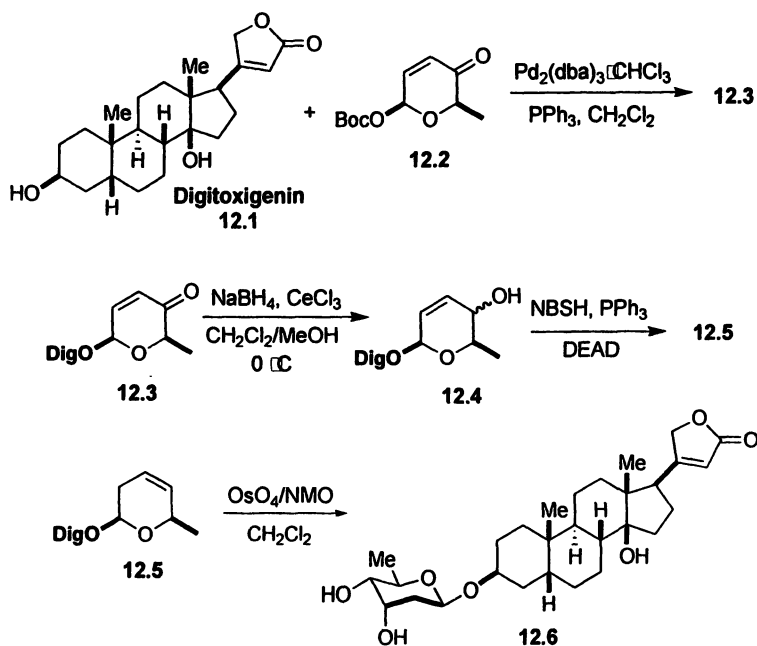
cardiac glycosides can have many biological roles, like solubilizing aglycons, facilitating membrane transport and targeting specific tissue. These sorts of efforts have been inhibited in the past the belief that *O*-glycosides have limited in vivo lifetimes. We believe that our de novo approach to install unnatural sugars is ideally suited for the generation of new bio-stable drug candidates. Thus, we decided to investigate the application of this approach for the synthesis of digitoxin from digitoxigenin as well as its mono- and disaccharide analogues (32).

Our efforts to prepare the digitoxin monosaccharide began with the investigation of the glycosylation reaction of the digitoxin aglycon, digitoxigenin (12.1, Scheme 12). At the outset, we were concerned about the compatibility of both the tertiary alcohol and butenolide in digitoxigenin (12.1) with our four-step sequence (12.1 to 12.6). To our surprise, this quickly turned into a non-issue. That is to say, the glycosylation occurred without interference from the tertiary alcohol. We next turned to the post-glycosylation transformation for the elaboration of the 2,6-dideoxy-allose sugar.

Ominously, these efforts began with an unselective ketone reduction of pyranone 12.1 with NaBH₄ (33) to give a mixture of allylic alcohols 12.2. Fortunately, both diastereomers could be used in the next reaction (12.2 to 12.3). Thus, applying the Myers reductive rearrangement conditions (NBSH, PPh₃/DEAD, NMM, -30 °C to rt) to 12.2 cleanly provided olefin 12.3 as a single diastereomer. Finally exposing olefin 12.3 to the Upjohn conditions (14) (OsO₄/NMO) gave exclusively the *allo*-diol 12.4 in good yield and as a single diastereomer. To our delight, the butenolide survived both the Luche reduction and Myers reductive rearrangement with no deleterious effects. Finally, the pyran double bond was selectively dihydroxylated under the Upjohn conditions with no sign of reaction at the butenolide double bond, thus providing the desired digitoxin monosaccharide 12.6 in 4 steps and good overall yield.

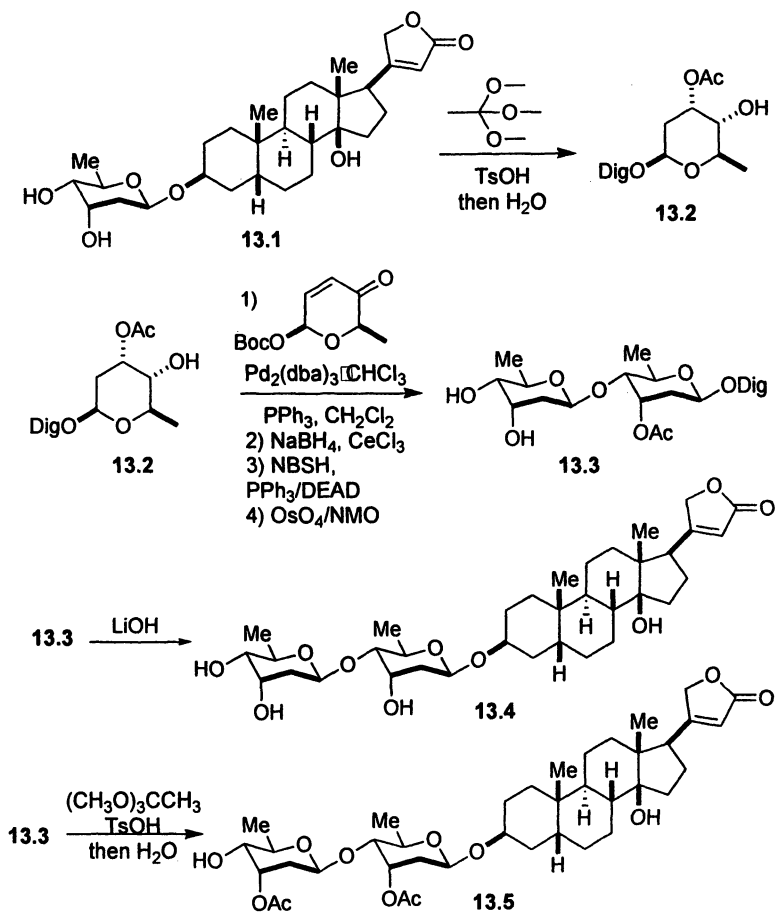
Our success with the installation of one sugar on the digitoxigenin boded well for the subsequent sequence (Scheme 13). The installation of the second digitoxin sugar required the regioselective acylation of the syn-3/4-diol in 13.1 (also 12.6 in Scheme 12). We decided to incorporate a selective protection step, which was achieved by the formation of an orthoester and regioselective ring opening (Scheme 13) (34). Specifically, the diol 13.1 was treated with trimethyl orthoacetate and catalytic *p*-toluenesulfonic acid to form an orthoester intermediate, which upon regiospecific acid hydrolysis (TsOH/H₂O) opened to the kinetically preferred axial acetate 13.2 in excellent yield. To our delight the equatorial alcohol in 13.2 smoothly underwent the Pd-catalyzed glycosylation, which was analogously reduced, rearranged, dihydroxylated and acylated with complete stereo- and regiocontrol to form the disaccharide 13.3, which after ester hydrolysis (LiOH) gave the desired digitoxin disaccharide 13.4 for testing. Similarly, disaccharide 13.3 could be selectively acylated using the same orthoester chemistry to give the disaccharide 13.5.

From the precedent established in Schemes 12 and 13, it came as no surprise that the same six-step sequence could be used to prepare digitoxin (14.3). Thus, in an identical sequence and with commensurate yields disaccharide 14.1 could be selectively glycosylated and deprotected to give digitoxin, which was prepared from its aglycon in only 15 steps.



Scheme 12. Synthesis of Digitoxin Monosaccharide

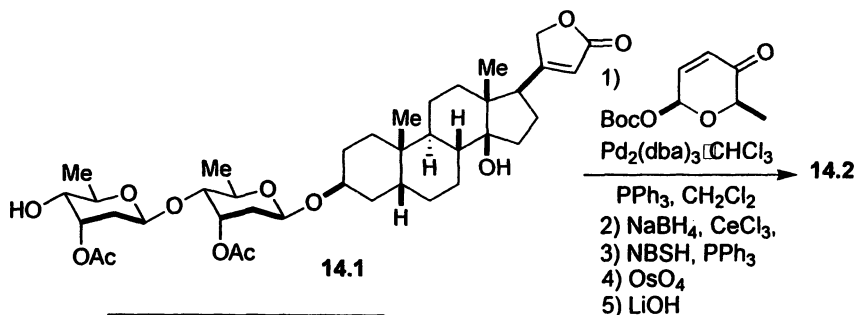
With the three cardiac glycosides (12.6, 13.4 and 14.3) in hand we have started to evaluate them in cytotoxicity assays. Emboldened by these earlier results we plan to continue these de novo synthetic efforts to prepare other unnatural mono-, di- and tri-saccharide digitoxin analogues for further studies. Of particular interest are the diastereomeric L-sugar digitoxin analogues as well as the α -linked diastereomeric analogues. Once the optimal stereochemistry has been established further studies would include the de novo synthesis of digitoxin glycosides with various C-6 substitution. We feel this approach should also be amenable for the medicinal chemistry study of other sugar containing natural products.



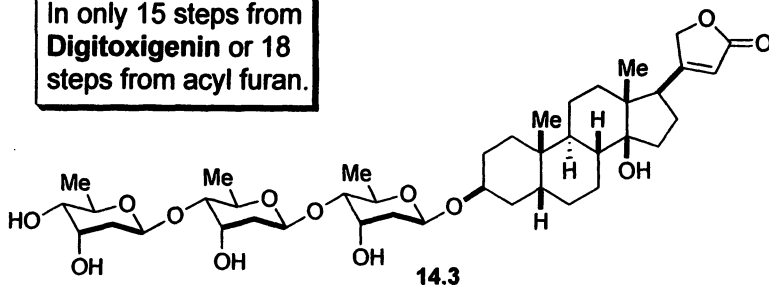
Scheme 13. Synthesis of Digitoxin Disaccharide

Conclusion

Reviewed was the use of our newly developed Pd-catalyzed glycosylation reaction for the incorporation of rare and unnatural sugars on unique biologically relevant structures. This de novo approach for the introduction of carbohydrate structures, under non-Lewis acidic conditions, into preexisting structural motifs is unique. To date, there simply does not exist an equivalent chemical or biochemical method for the stereoselective glycosylation of under-functionalized D- and L-sugars, which can be elaborated into a myriad of carbohydrate analogues. Not to mention this method also allows for complete α/β -stereocontrol. The power and uniqueness of this approach is probably best



Digitoxin:
In only 15 steps from
Digitoxigenin or 18
steps from acyl furan.



Scheme 14. Synthesis of digitoxin

exemplified in our synthesis and study of various sugar stereoisomers, and in turn the medicinal chemistry SAR study of these structures, which until now were not obtainable by traditional carbohydrate routes.

Acknowledgments

We are grateful to WVU-ARTS program, NIH (GM63150) and NSF (CHE-0415469) for the support of our research program and NSF-EPSCoR (0314742) for a 600 MHz NMR spectrometer at WVU.

References

1. For a review of carbohydrates in medicinal chemistry see: Gijzen, H. J. M.; Qiao, L.; Fitz, W.; Wong, C.-H. *Chem. Rev.* **1996**, *96*, 443-473.
2. For the first use of catalysis towards the hexoses, see: Ko, S. Y.; Lee, A. W.

- M.; Masamune, S.; Reed, L. A.; Sharpless, K. B. *Science* **1983**, *220*, 949-951.
- For a good review: a) Zamoiski, A., Banaszek, A.; Grynkiewicz, G. *Advances in Carbohydrate Chemistry and Biochemistry* **1982**, *40*, 61-72. b) Achmatowicz, O.; Bielski, R. *Carbohydr. Res.* **1977**, *55*, 165-176. c) Grapsas, I., K.; Couladouros, E. A.; Georgiadis, M. P. *Pol. J. Chem.* **1990**, *64*, 823-826.
 - (a) Johnson, C. R.; Golebiowski, A.; Steensma, D. H.; Scialdone, M. A. *J. Org. Chem.* **1993**, *58*, 7185-7194. (b) Hudlicky, T.; Pitzer, K. K.; Stabile, M. R.; Thorpe, A. J.; Whited, G. M. *J. Org. Chem.* **1996**, *61*, 4151-4153.
 - For alternative view of de novo carbohydrate synthesis see: a) Timmer, M. S. M.; Adibekian, A.; Seeberger, P. H. *Angew. Chem.* **2005**, *117*, 7777-7780; *Angew. Chem. Int. Ed.* **2005**, *44*, 7605-7607. b) Schmidt, R. R. *Pure Appl. Chem.* **1987**, *59*, 415-424. c) Kirschning, A.; Jesberger, M.; Schöningh, K.-U. *Synthesis* **2001**, 507-540.
 - For a review: a) Danishefsky, S. J. *Chemtracts* **1989**, 273. For improved catalysis see: b) Schaus, S. E.; Branalt, J.; Jacobsen, E. N. *J. Org. Chem.* **1998**, *63*, 403-405.
 - Henderson, I.; Sharpless, K. B.; Wong, C.-H. *J. Am. Chem. Soc.* **1994**, *116*, 558-561.
 - Northrup, A. B.; MacMillan, D. W. C. *Science* **2004**, *305*, 1752-1755.
 - Covell, D. J.; Vermeulen, N. A.; Labenz, N. A.; White, M. C. *Angew. Chem. Int. Ed.* **2006**, *45*, 8217-8220.
 - (a) Ahmed, Md. M.; O'Doherty, G. A. *Tetrahedron Lett.* **2005**, *46*, 3015-3019. (b) Ahmed, Md. M.; Berry, B. P.; Hunter, T. J.; Tomcik, D. J.; O'Doherty, G. A. *Org. Lett.* **2005**, *7*, 745-748.
 - For its application in synthesis, see: (a) Ahmed, Md. M.; O'Doherty, G. A. *Tetrahedron Lett.* **2005**, *46*, 4151-4155. (b) Gao, D.; O'Doherty, G. A. *Org. Lett.* **2005**, *7*, 1069-1072.
 - Ahmed, Md. M.; O'Doherty, G. A. *J. Org. Chem.* **2005**, *67*, 10576-10578.
 - (a) Harris, J. M.; Keranen, M. D.; O'Doherty, G. A. *J. Org. Chem.* **1999**, *64*, 2982-2983. (b) Harris, J. M.; Keranen, M. D.; Nguyen, H.; Young, V. G.; O'Doherty, G. A. *Carbohydr. Res.* **2000**, *328*, 17-36. For its use in oligosaccharide synthesis see: (c) Babu, R. S.; Zhou, M.; O'Doherty, G. A. *J. Am. Chem. Soc.* **2004**, *126*, 3428-3429.
 - VanRheenen, V.; Kelly, R. C.; Cha, D. Y. *Tetrahedron Lett.* **1976**, *17*, 1973-1976.
 - Ahmed, Md. M.; O'Doherty, G. A. *Carbohydr. Res.* **2006**, *341*, 1505-1521.
 - Haukaas, M. H.; O'Doherty, G. A. *Org. Lett.* **2002**, *4*, 1771-1774.
 - Haukaas, M. H.; O'Doherty, G. A. *Org. Lett.* **2001**, 3899-3992.
 - Traxler, P.; Gruner, J.; Auden, J. A. L. *J. Antibiot.* **1977**, *30*, 289-296.
 - Concurrent with these studies was the similar discovery by Feringa and Lee: (a) Comely, A. C.; Eelkema, R.; Minnaard, A. J.; Feringa, B. L. *J. Am.*

- Chem. Soc.* **2003**, *125*, 8714-8715. (b) Kim, H.; Men, H.; Lee, C. *J. Am. Chem. Soc.* **2004**, *126*, 1336-1337.
20. Recently, both the poor reactivity in Pd-catalyzed allylation reaction of alcohols as well as a nice solution to this problem was reported, see: Kim, H.; Lee, C. *Org. Lett.* **2002**, *4*, 4369-4372.
 21. For a related Rh system, see: Evans, P. A.; Kennedy, L. *J. Org. Lett.* **2000**, *2*, 2213-2215.
 22. We have had great success at the preparation of optically pure furan alcohols from the Noyori reduction of the corresponding acylfuran, see: refs. (a) Li, M.; Scott, J. G.; O'Doherty, G. A. *Tetrahedron Lett.* **2004**, *45*, 1005-1009. (b) Li, M.; Scott, J. G.; O'Doherty, G. A. *Tetrahedron Lett.* **2004**, *45*, 6407-6411. (c) Guo, H.; O'Doherty, G. A. *Org. Lett.* **2006**, *8*, 1609-1612.
 23. Guppi, S. R.; Zhou, M.; O'Doherty, G. A. *Org. Lett.* **2006**, *8*, 293-296.
 24. Babu, R. S.; Zhou, M.; O'Doherty, G. A. *J. Am. Chem. Soc.* **2004**, *126*, 3428-3429.
 25. (a) Babu, R. S.; Guppi, S. R.; O'Doherty, G. A. *Org. Lett.* **2006**, *8*(8), 1605-1608. (b) Guppi, S. R.; O'Doherty, G. A. *J. Org. Chem.* **2007**, *72*, 4966-4969.
 26. Zhou, M.; O'Doherty, G. A. *Org. Lett.* **2006**, *8*, 4339-4342.
 27. (a) Haux, J.; Klepp, O.; Spigset, O.; Tretli, S. *BMC Cancer* **2001**, *1*, 11; (b) Manna, S. K.; Sah, N. K.; Newman, R. A.; Cisneros, A.; Aggarwal, B. B. *Cancer Res.*, **2000**, *60*, 3838-3847.
 28. Haux, J. *Med. Hypotheses* **1999**, *53*, 543-548.
 29. Lopez-Lazaro, M.; Pastor, N.; Azrak, S. S.; Ayuso, M. J.; Austin, C. A.; Cortes, F. *J. Nat. Prod.* **2005**, *68*, 1642-1645
 30. Langenhan, J. L.; Peters, N. R.; Guzei, I. A.; Hoffman, F. M.; Thorson, J. S. *Proc. Natl. Acad. Sci.* **2005**, *102*, 12305-12310.
 31. For neoglycoside formation see: (a) Peri, F.; Dumy, P.; Mutter, M. *Tetrahedron* **1998**, *54*, 12269-12278. (b) Peri, F.; Jimenez-Barbero, J.; Garcia-Aparico, V.; Tvaroka, I.; Nicotra, F. *Chem. Eur. J.* **2004**, *10*, 1433-1444.
 32. For other approaches to the sugar portion of digitoxin see: (a) Wiesner, K.; Tsai, T. Y. R.; Jin, H. *Helv. Chim. Acta* **1985**, *68*, 300-314. (b) Wiesner, K.; Tsai, T. Y. R. *Pure Appl. Chem.* **1986**, *58*, 799-810. and (c) McDonald, F. E.; Reddy, K. S.; Diaz, Y. *J. Am. Chem. Soc.* **2000**, *122*, 4304-4309. (d) McDonald, F. E.; Reddy, K. S. *Angew. Chem. Int. Ed.* **2001**, *40*, 3653-3655.
 33. For reduction of β -pyanones, the $CeCl_3$ is necessary to avoid 1,4-reduction products. For Luche reduction, see: (a) Luche, J. L. *J. Am. Chem. Soc.* **1978**, *100*, 2226-2227. and ref. 13, 16, and 17.
 34. (a) King, J. F.; Allbutt, A. D. *Can. J. Chem.* **1970**, *48*, 1754-1769. (b) Lowary, T. L.; Hindsgaul, O. *Carbohydr. Res.* **1994**, *251*, 33-67.

Chapter 2

Chemical Syntheses of Hyaluronic Acid Oligosaccharides

Lijun Huang^{1,2} Xiaowei Lu,¹ and Xuefei Huang^{1,*}

¹Department of Chemistry, The University of Toledo, 2801 West Bancroft
Street, MS 602, Toledo, OH 43606

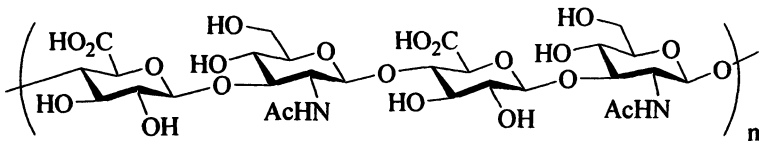
²Current address: Anatrace Inc., 434 West Dussel Drive,
Maumee, OH 43537

*Corresponding author: xuefei.huang@utoledo.edu

The increasing recognition of important biological functions of hyaluronic acid oligosaccharides (sHA) has led to developments of new methodologies for sHA syntheses. Herein, we provide an overview of chemical approaches for assembling these oligosaccharides. Compared with traditional methods, our iterative one-pot strategy provided a new platform for rapid assembly of sHA without resorting to multiple protective group manipulations, anomeric adjustments or intermediate purifications. We expect that the continual development of chemical synthetic methodologies of sHA will provide a facile access to structurally well defined sHA derivatives that will greatly facilitate structure-activity relationship studies.

Introduction

Hyaluronic acid (HA), first discovered by Meyer and Palmer in 1934 (1), belongs to the glycosaminoglycan (GAG) superfamily. As the only unsulfated GAG, HA contains linear tandem disaccharide repeating units of D-glucuronic acid and 2-deoxy-2-N-acetyl-D-glucose [β -D-GlcpA-(1 \rightarrow 3)- β -D-GlcpNAc-(1 \rightarrow 4)-]. HA is broadly distributed in the extracellular space of vertebrates with the highest concentrations in soft connective tissues (2,3). As a high molecular weight polyanion, an essential role of HA is to participate in a hydrated network with collagen fibers, where it acts as an organizing core to form complex intercellular aggregates. Due to its involvement in a variety of cellular events such as cell adhesion, cell migration, atherosclerosis and wound healing as well as its relatively low toxicity and immunogenicity (4-11), HA has been widely used in clinical applications including in eye surgery and as an intra-articular matrix supplement (6,12).



HA can be transformed to hyaluronic acid oligosaccharides (sHA) ($\sim 2 \times 10^3$ D) *in vivo*. Biological studies indicated that sHA can have completely distinct novel properties from high molecular weight HA polymers, including stimulation of endothelial cell proliferation and migration (13), stimulation of angiogenesis (7), potentiation of the innate immune system (14-16), inhibition of tumor growth (17) and regulation of multi-drug resistance in cancer cells (18). Several cellular receptors for sHA have been identified. For example, Toll like receptor 4, which is a membrane bound receptor involved in initiation of innate immunity and prevention of early spread of pathogens, has been shown to be activated by sHA tetrasaccharides and hexasaccharides, rather than by the high molecular weight HA (14,15). CD44, a membrane glycoprotein, is involved in many biological events such as cell migration during morphogenesis, angiogenesis, and tumor invasion and metastasis (19-21). HA is the primary ligand for CD44, with the minimum length of a sHA hexasaccharide required for binding (22). The binding affinity between sHA and CD44 increases with longer sHA (23,24).

The biological activities of sHA can be sequence specific, an example of which is the report that sHA tetrasaccharides up-regulate the expression of heat-shock protein 72 as well as the corresponding mRNA and suppress cell death under stress condition, while the corresponding di-, hexa- and octa-saccharides are inactive (25). To date, the majority of biological studies is performed using

a mixture of sHA with various lengths and structures derived from enzymatic degradation of HA isolated from natural sources (26,27). The possibility of highly active contaminants notwithstanding, it is difficult to firmly establish structure-activity relationship (SAR) with a mixture of sHA structures. Access to synthetic sHA with distinct length and sequence can greatly facilitate a systematic investigation of its SAR, providing exciting opportunities for the development of novel therapeutics.

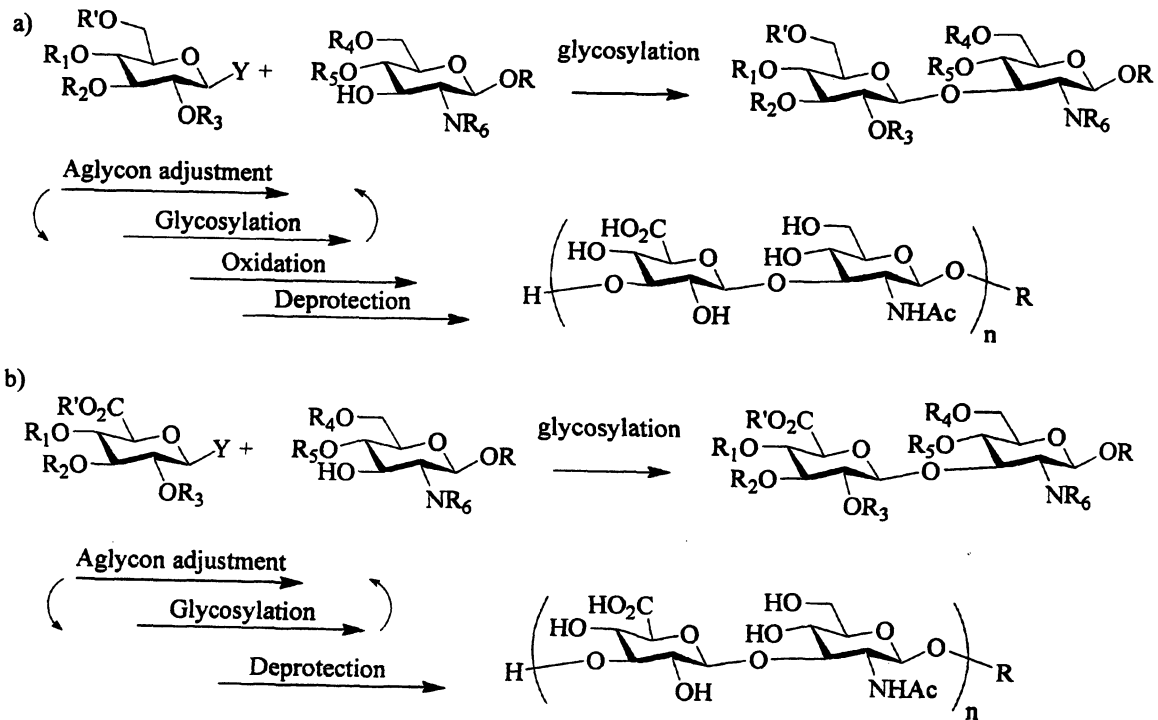
sHA can be assembled by either chemoenzymatic or chemical synthesis. Early chemoenzymatic approaches utilized biosynthetic pathways through HA synthases and associated accessory enzymes (28) or transglycosylation reactions catalyzed by hyaluronidases (29). Recently, oxazolidine containing building blocks, which are transition state analogs for hyaluronidases, have been polymerized to yield HA (30-32). In order to control the length of HA from enzymatic reactions, a HA synthase was converted by mutagenesis into two single-action glycosyltransferases (glucuronic acid transferase and *N*-acetylglucosamine transferase) (33). The alternating stepwise usage of these two novel enzymes led to the construction of a series of monodispersed synthetic sHA. Despite these successes, the inherent substrate specificities of enzymes limit the structural diversity of sHA analogs that can be generated. Chemical synthesis can complement the chemoenzymatic approaches to create greater varieties of sHA structures, facilitating SAR studies. In this review, we will focus on the chemical synthesis of sHA and recent progress on iterative one-pot sHA synthesis.

Conventional Syntheses of sHA

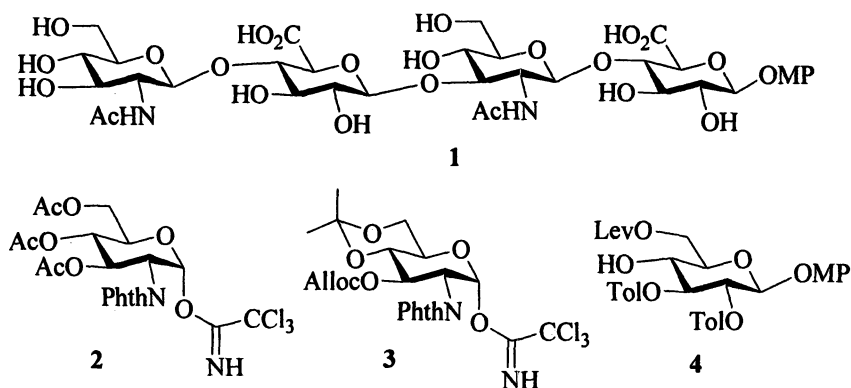
There are three central factors need to be taken into consideration in designing a synthetic route for sHA: 1) stereocontrolled assembly of the oligosaccharide backbone; 2) introduction of glucuronic acid; and 3) installation of acetamido groups. Two general strategies for sHA assembly are typically adapted depending upon the order of transformations (Scheme 1) (34,35). In the first method, the more reactive glucose is used as a building block with post glycosylation conversion to glucuronic acid via oxidation (Scheme 1a) (36-42), while the second method utilizes glucuronic acid directly as the glycosyl donor (Scheme 1b) (43). Currently, both approaches have been successfully applied in preparation of sHA.

For the first approach, a selectively removable protective group must be installed on the 6-hydroxyl group of glucose to allow for downstream oxidation state adjustment. As an example, Vliegthart and coworkers have reported the chemical synthesis of sHA tetrasaccharide **1** with glucuronic acid at the reducing end (Scheme 2) (39). The desired 1,2-*trans* linkages were controlled by the presence of participating neighboring groups, i.e., *N*-phthaloyl (Phth) on

Scheme 1.

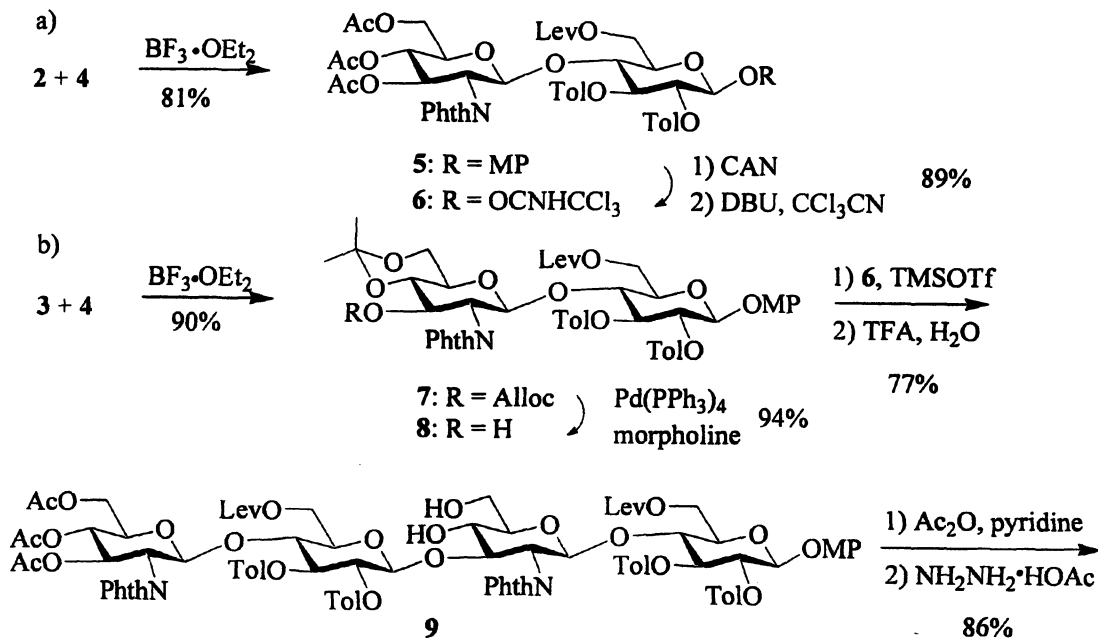


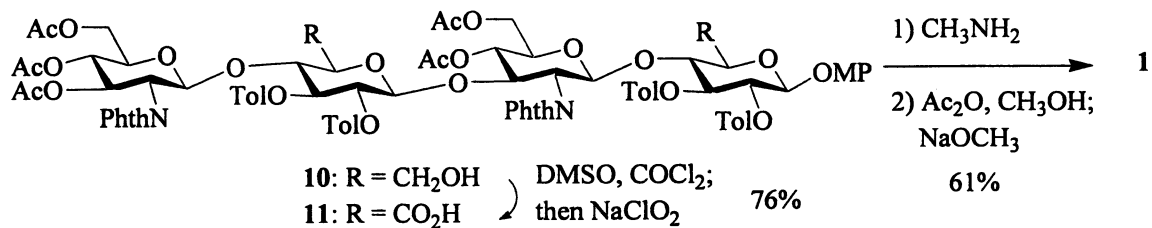
glucosamines **2** and **3** and *p*-toluoyl (Tol) on glucoside **4**. The levulinoyl moiety was installed on 6-*O* of glucoside **4** as the selective protective group. Following glycosylation of **4** by donor **2**, adjustment of the aglycon group at the reducing end by removal of the anomeric methoxyphenyl (MP) and imidation of the resulting hemiacetal yielded disaccharide donor **6** (Scheme 2a). Reaction of **4** with **3** and subsequent deprotection gave disaccharide acceptor **8**, which was followed by glycosylation by donor **6** and isopropylidene removal producing sHA tetrasaccharide core **9** (Scheme 2b). Acetylation and selective cleavage of the two Lev groups by hydrazine acetate yielded diol **10**, which was oxidized to the corresponding di-carboxylic acid **11** by tandem Swern and sodium chlorite oxidations. Treatment of **11** by methyl amine and subsequent *N*-acetylation produced the target sHA tetrasaccharide **1**.



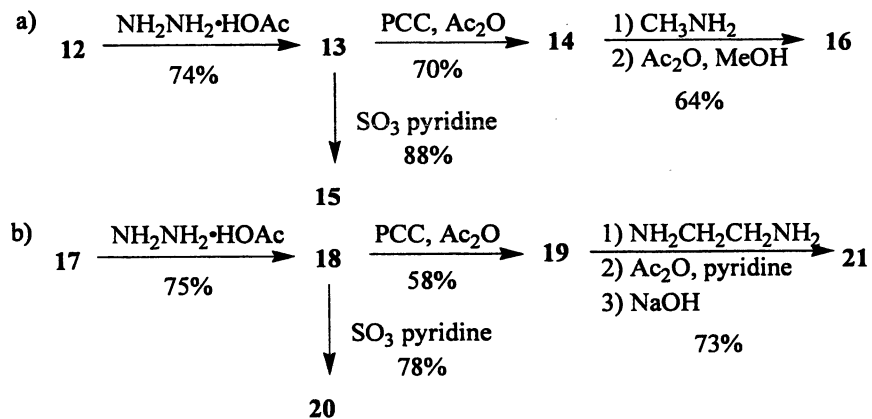
This strategy has been expanded to synthesize sHA pentasaccharide **12** and sHA hexasaccharide **17** with glucosamine at the reducing end (38). However, after delevulinoylation of **12** with hydrazine acetate, oxidation of the diol **13** by the tandem Swern and sodium chlorite oxidations did not give satisfactory results. Instead, the usage of pyridinium chlorochromate (PCC) and acetic anhydride was necessary producing the desired diacid **14** in good yield (Scheme 3a). Following the same oxidation method with PCC, the conversion of triol **18** to triacid **19** was achieved (Scheme 3b). Dephthaloylation/deacetylation of triacid **19** using methylamine in ethanol gave a complex mixture. Ultimately, the removal of *N*-phthaloyl moieties was carried out with treatment of ethylenediamine in 1-butanol followed by *N*-acetylation to afford the fully deprotected sHA hexasaccharide **21**. Besides oxidation, the free hydroxyl groups in **13** and **18** can be sulfated generating analogs **15** and **20**, which are difficult to obtain through chemoenzymatic synthesis, thus highlighting the advantage of chemical synthesis.

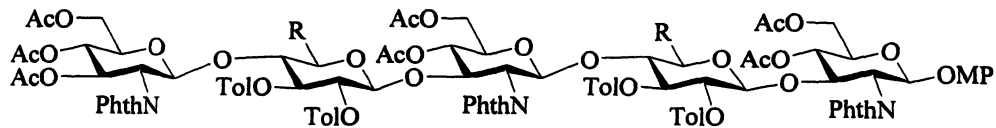
Scheme 2.





Scheme 3.



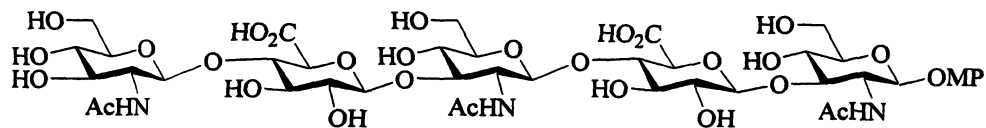


12: R = CH₂OLev

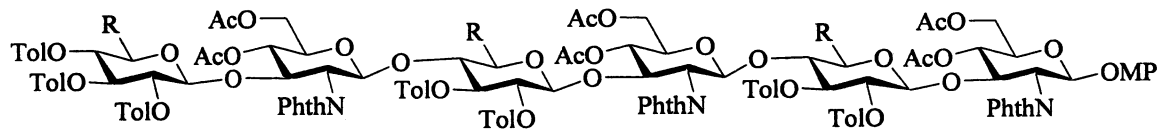
13: R = CH₂OH

14: R = CO₂H

15: R = CH₂OSO₃⁻ Na⁺



16

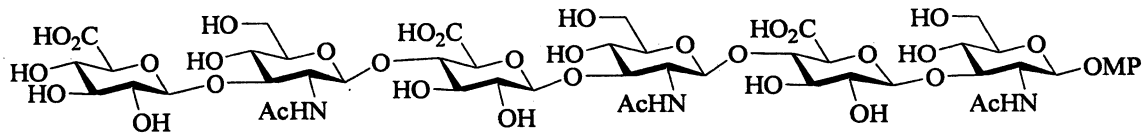


17: R = CH₂OLev

18: R = CH₂OH

19: R = CO₂H

20: R = CH₂OSO₃⁻ Na⁺



21

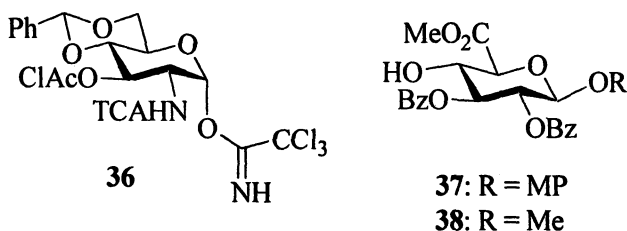
Petillo and coworkers have reported syntheses of sHA trisaccharides with the combination of the sulfoxide and trichloroacetimidate glycosylation methodologies (Scheme 4) (37). Glycosylation of acceptor **22** by glycosyl sulfoxide **23** produced disaccharide **25**. The protective groups on **25** were adjusted leading to acceptor **26** with a *p*-methoxybenzyl (PMB) group masking the 6-hydroxyl of the glucose unit. 2,2,2-Trichloroethoxy carbonyl (Troc) was employed as the 1,2-*trans* directing *N*-protective group for glucosamine trichloroacetimidate donor **24**, which glycosylated alcohol **26** yielding sHA trisaccharide core **27**. Conversion of the *N*-Troc and azido into acetamido moieties gave trisaccharide **28**. The PMB group was selectively removed by cerium ammonium nitrate (CAN) exposing the 6-hydroxyl group of the glucose, which was directly converted to carboxylic acid by a TEMPO mediated NaOCl oxidation in a modest 57% yield. Subsequent deacetylation and hydrogenolysis produced sHA trisaccharide **31**.

Recently, Seeberger and coworkers investigated the conversion of glucoside into glucuronic acid for preparation of sHA building blocks (44). Pyridinium dichromate mediated oxidation of glucoside **32** was found to require large excess of the chromium oxidant and the difficulty in purification led to small amount of the desired glucuronic acid. Attempts to use Swern oxidation yielded elimination of C4-*O*-acetate due to the basic reaction condition. Other oxidants such as Pt/O₂, tetrapropylammonium perruthenate/O₂, RuCl₂(PPh)₃ and NaNO₂/Ac₂O failed to react. Finally, periodic acid (H₅IO₆) combined with a catalytic amount of CrO₃ was found to be suitable. Although 70% yield was obtained for converting glucoside **32** to glucuronic ester **33** by H₅IO₆/CrO₃ oxidation followed by methylation, the yield for transforming disaccharide **34** to **35** was only a modest 39%. These results underscore the challenges in sHA synthesis, accentuating the need to continually develop methodologies for not only glyco-assembly but also associated protective group removal and oxidation.

As an alternative to glucosyl donors, glucuronic acid building blocks can be directly utilized for sHA assembly. Due to the presence of an electron withdrawing carboxylic acid on the pyranose ring, glucuronic acid building blocks are known to be less reactive than glucoside both as glycosyl donors and acceptors. Jacquinet and coworkers developed a nice strategy using the reactive glucuronic acid trichloroacetimidate donors (43). Trichloroacetamide (TCA) was employed as the *N*-protective group of glucosamine, which can be readily converted into the corresponding acetamides under the neutral Bu₃SnH reduction condition without going through amine intermediates.

The two key disaccharides **39** and **41** in Jacquinet synthesis were prepared from the coupling of glycosyl donor **36** with acceptors **37** and **38** respectively (Scheme 5a,b). Disaccharide **39** was transformed into trichloroacetimidate donor **40** over two steps, while disaccharide **41** was subjected to dechloroacetylation to afford acceptor **42**. Coupling of **42** with **40** followed by dechloroacetylation gave tetrasaccharide acceptor **43** (Scheme 5). Repetition of

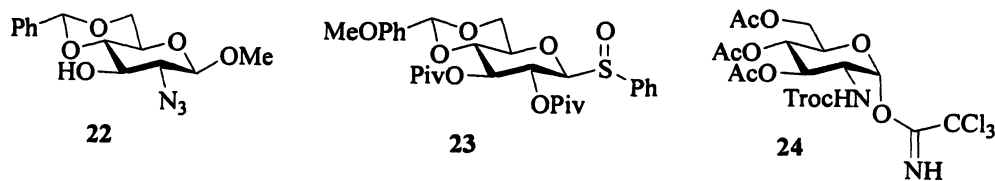
glycosylation by **40**, deprotection of the chloroacetyl group and glycosylation afforded octasaccharide **44**. Subsequent deprotections were performed in the following order: removal of the chloroacetyl group by thiourea, transformation of NH-TCA into NHAc via Bu_3SnH reduction, debenzylidene and saponification to provide sHA octasaccharide **45** (Scheme 5c), which is the longest chemically constructed sHA to date.



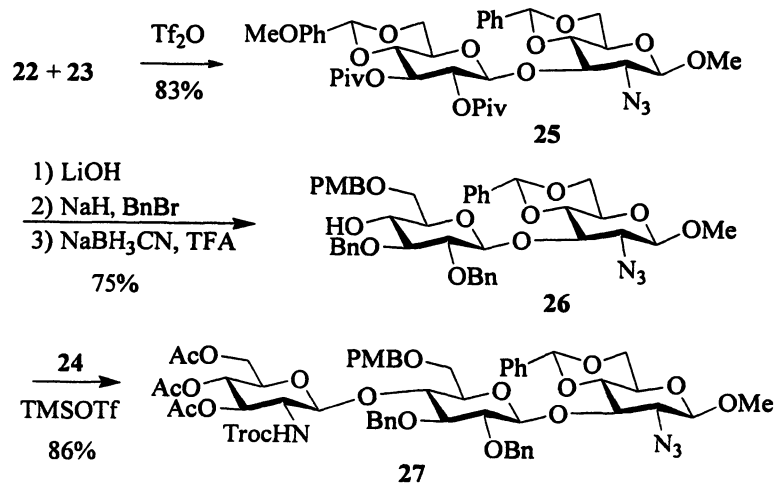
Iterative One-pot Synthesis of sHA

As described above, traditional oligosaccharide synthesis required multiple modifications of aglycons, adjustments of protective groups on advanced oligosaccharide intermediates as well as separations of these intermediates. To expedite the synthesis and circumvent some of the limitations associated with conventional approaches, we have recently developed a novel iterative one-pot oligosaccharide assembly strategy (45), in which a thioglycosyl donor is pre-activated by a thiophilic promoter in the absence of an acceptor, generating a reactive intermediate. Upon addition of a thioglycosyl acceptor to the reaction, the free hydroxyl group of the acceptor will attack on the reactive intermediate leading to a disaccharide. With its anomeric thioacetal moiety, the disaccharide can be subjected to another round of pre-activation and nucleophilic substitution extending the oligosaccharide chain. Multiple glycosylations can be sequentially carried out in the same flask without intermediate separations, thus greatly facilitating glyco-assembly (Scheme 6). Furthermore, because donor activation and glycosylation are carried out in two distinct steps, anomeric reactivities of thioglycosyl donor and acceptor do not need to be differentiated (46), granting greater flexibilities in selecting of protective groups to achieve high yields in glycosylation.

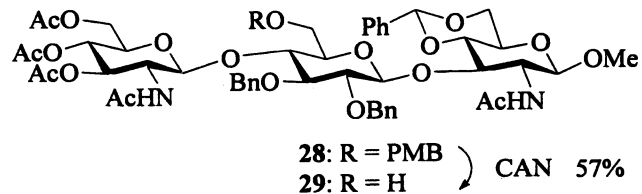
In order to apply the iterative one-pot method to sHA synthesis, we first evaluated several derivatives of glucuronic acid, glucose and glucosamine as building blocks (47). Because glucuronic acid donor **46** led to low glycosylation yield, glucosides were used for further investigation and the PMB moiety was installed to mask 6-hydroxyl groups for future selective deprotection. Glucoside **47** was found to be a suitable donor, as its TBS group is important for high



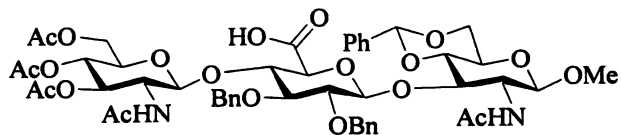
Scheme 4.



1) Cd, DMF/AcOH
 2) Ac₂O
 3) 1,3-propanedithiol
 4) Ac₂O
 90%

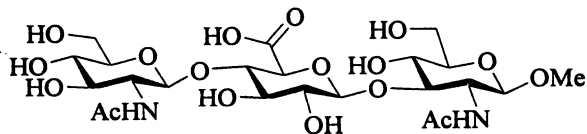


TEMPO
 NaOCl

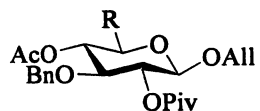


30

1) NaOMe
 2) H₂, Pd/C
 100%

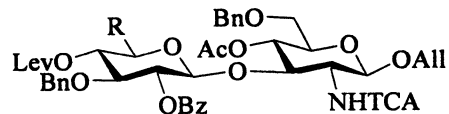


31



32: R = CH₂OH

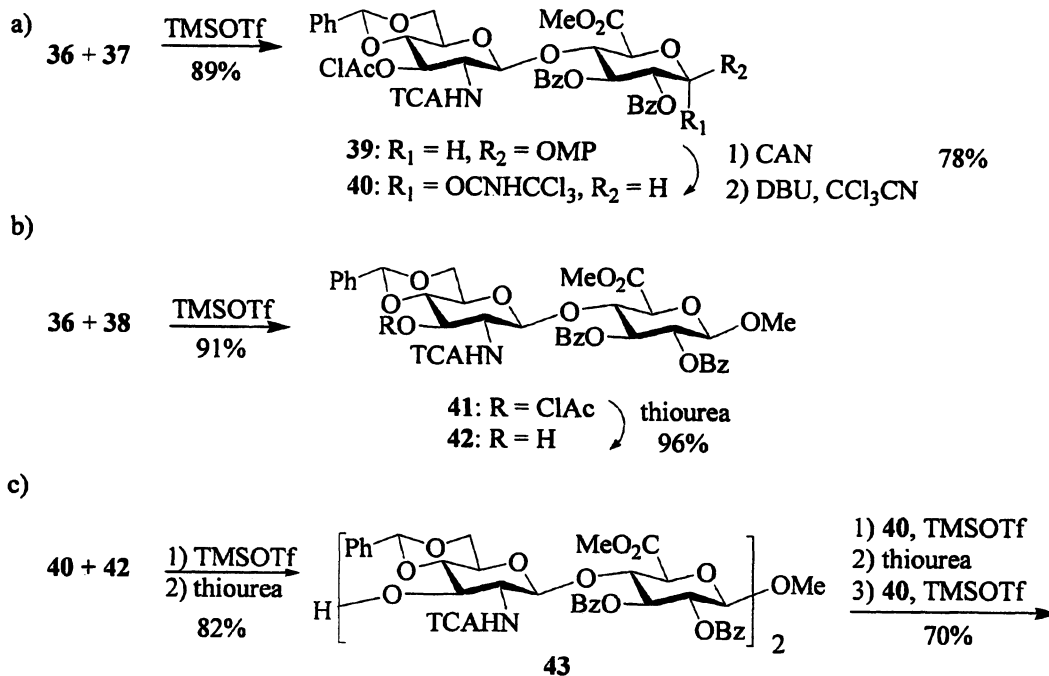
33: R = CO₂Me

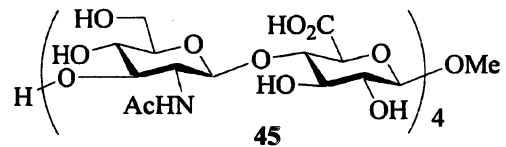
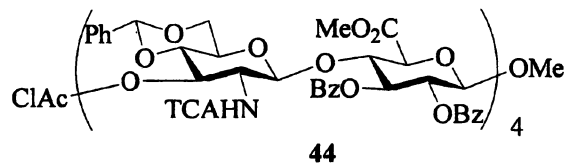


34: R = CH₂OH

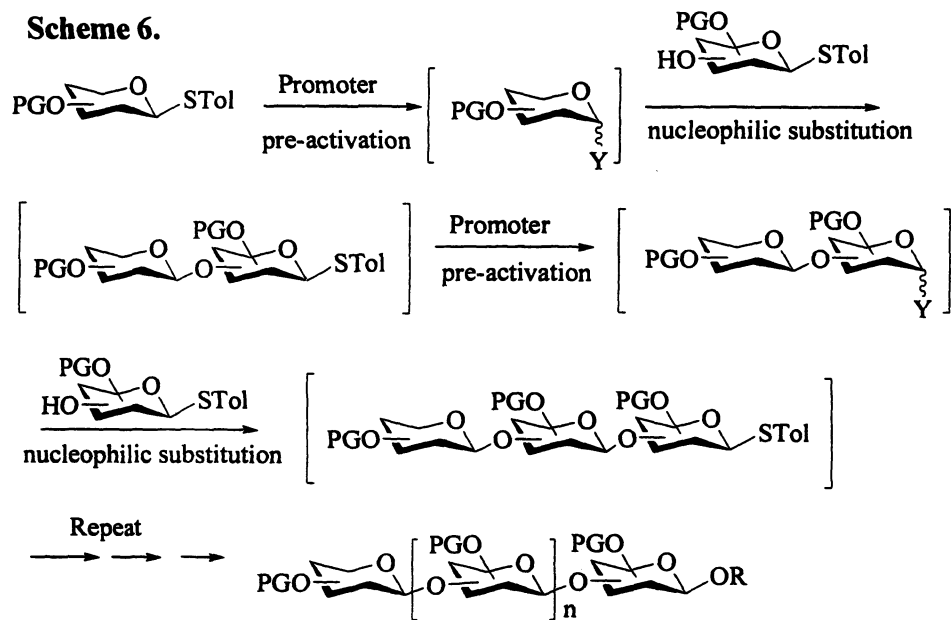
35: R = CO₂Me

Scheme 5.

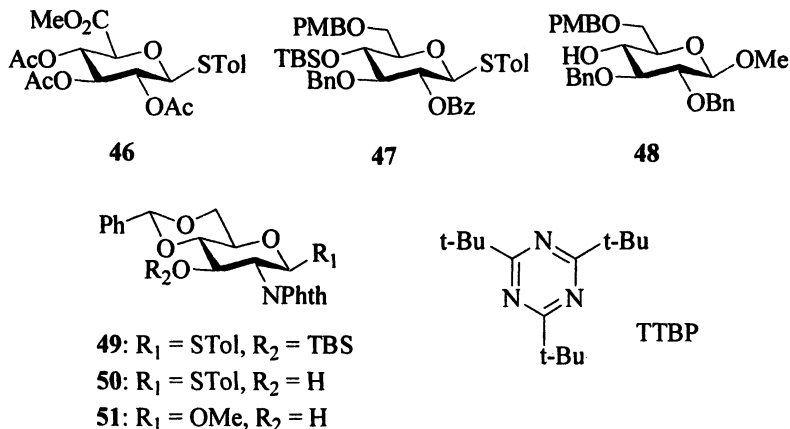




1) thiourea
 2) Bu_3SnH , AIBN
 3) 80% AcOH
 4) NaOH
 63%

Scheme 6.

yielding glycosylation, presumably due to solubility enhancement of the activated donor at low temperature in the reaction solvent diethyl ether. The *N*-protective group of glucosamines did not significantly influence the glycosylation and *N*-Phth was selected for its high 1,2-*trans* directing property.



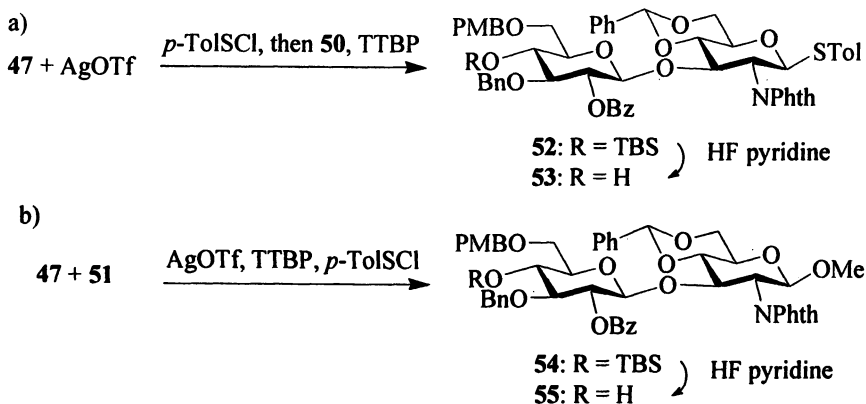
Donor **47** was pre-activated by the promoter *p*-TolSOTf, formed *in situ* through the reaction of AgOTf with *p*-TolSCl (**Scheme 7a**). Subsequent addition of acceptor **50** and a sterically hindered non-nucleophilic base 2,4,6-tri-*t*-butyl pyrimidine (TTBP) to the reaction formed disaccharide **52**, which was deprotected to give disaccharide **53**. Disaccharide acceptor **55** with a methoxy group at the reducing end was prepared by reacting **47** with acceptor **51**, followed by removal of TBS (**Scheme 7b**).

With all necessary building blocks in hand, one pot syntheses were performed. Pre-activation of donor **47** by *p*-TolSOTf was followed by addition of acceptor **50** and TTBP. Upon completion of the reaction, addition of acceptor **55**, TTBP and promoter *p*-TolSOTf to the same reaction flask produced sHA tetrasaccharide core **56** in excellent overall yield in just three hours (Table 1, entry 1), which was the only compound that needed purification in this three component one-pot synthesis.

sHA core sequences containing an odd number of monosaccharide units can also be synthesized. Two pentasaccharide **57** and **58** with different sequences were constructed through three and four component one-pot reactions in excellent yields (Table 1, entries 2, 3). A hexasaccharide **59** was easily accessed as well in high yield following the pre-activation protocol (Table 1, entry 4).

With a near stoichiometric amount of building blocks used for each glycosylation and no oligosaccharide intermediate purification involved, these syntheses can be easily scaled up. Gram quantity of sHA hexasaccharide **59** was obtained within just a few hours in similar yield as the smaller scale reaction. The scalability coupled with the speed of glyco-assembly and high overall yields

Scheme 7.



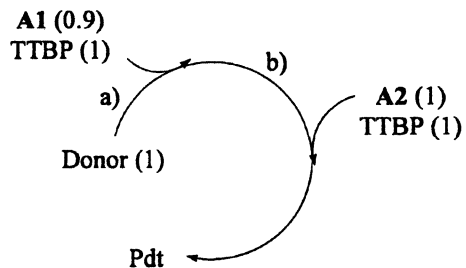
highlights the advantages of using the iterative one-pot approach for complex oligosaccharide synthesis.

After the establishment of sHA core, the next stage of our work was focused on the deprotections and oxidation state adjustments. Careful planning of the deprotection order for various functional groups and strategy for the oxidation of C-6 OH is critical for the success of sHA syntheses.

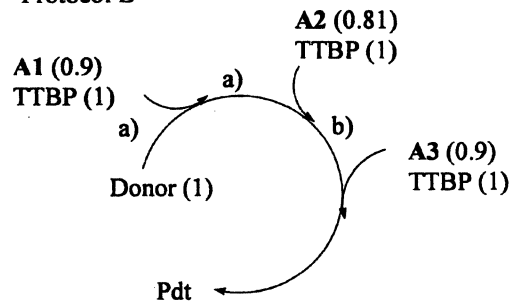
Our initial efforts of removing the three PMB groups from hexasaccharides **60** and **63** failed using either CAN or 2,3-dichloro-5,6-dicyano-1,4-benzoquinone (DDQ), although similar reactions have been reported such as CAN oxidation of sHA trisaccharide **28** (Scheme 4) (37). After repeated trials, we found that the presence of TBS moiety was beneficial for deprotection of PMB, as CAN oxidation of hexasaccharide **59** produced triol **61** in good yield (Scheme 8). Conversion of **61** into tricarboxylic acids turned out to be challenging. Several methods such as TEMPO/NaOCl, NaOCl/NaClO₂ catalyzed by TEMPO, TEMPO/iodobenzenediacetate, and Dess-Martin oxidation followed by NaClO₂ did not give the desired product presumably due to the need to oxidize multiple hydroxyl groups in the same molecule. Instead, multiple partially oxidized products were often obtained from these reactions. Finally, we discovered that a convenient two step one-pot protocol using TEMPO/NaOCl followed by treatment of NaClO₂ afforded the desired tri-carboxylic acid, which was isolated as benzyl ester **62** by subsequent treatment with phenyl diazomethane in high yield and good purity (Scheme 8) (48). This new protocol was also found to be compatible with a variety of sensitive functional groups, such as allyl, thioacetal, PMB and isopropylidene. Final deprotections of sHA **62** was carried out by the reaction

Table 1.

Protocol A

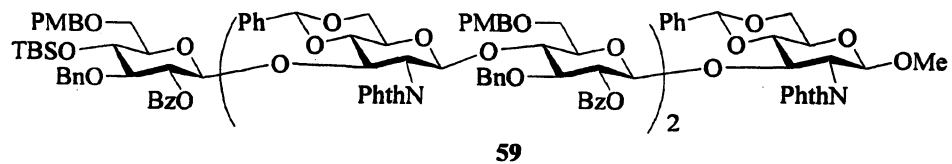
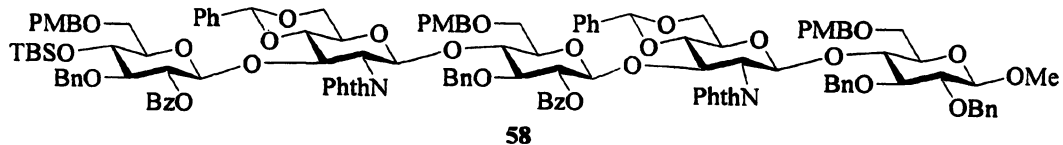
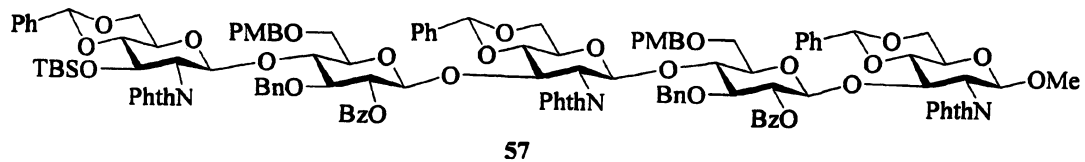
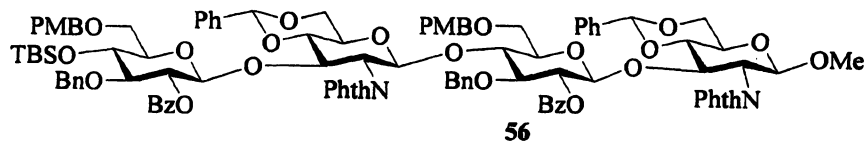


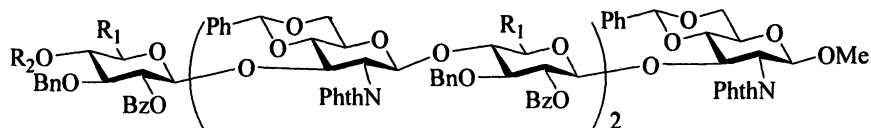
Protocol B



Reagents and conditions: a) AgOTf, *p*-TolSCl, - 65 °C, 10 min; then acceptor, TTBP, 90 min to 0 °C; 15 min, 0 °C; b) acceptor, TTBP, AgOTf, *p*-TolSCl, - 65 °C - 0 °C in 90 min.

Entry	Protocol	Donor	Acceptor 1 (A1)	Acceptor 2 (A2)	Acceptor 3 (A3)	Product	Yield
1	A	47	50	55		56	64 - 75%
2	A	49	53	55		57	65%
3	B	47	50	53	48	58	55%
4	B	47	50	53	55	59	54 - 60%

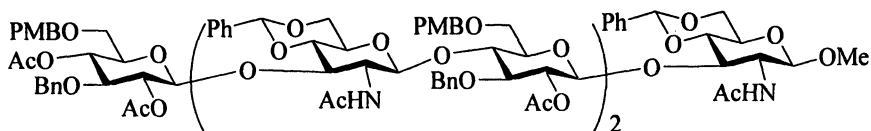




60: R₁ = CH₂OPMB, R₂ = H

61: R₁ = CH₂OH, R₂ = TBS

62: R₁ = CO₂Bn, R₂ = TBS



63

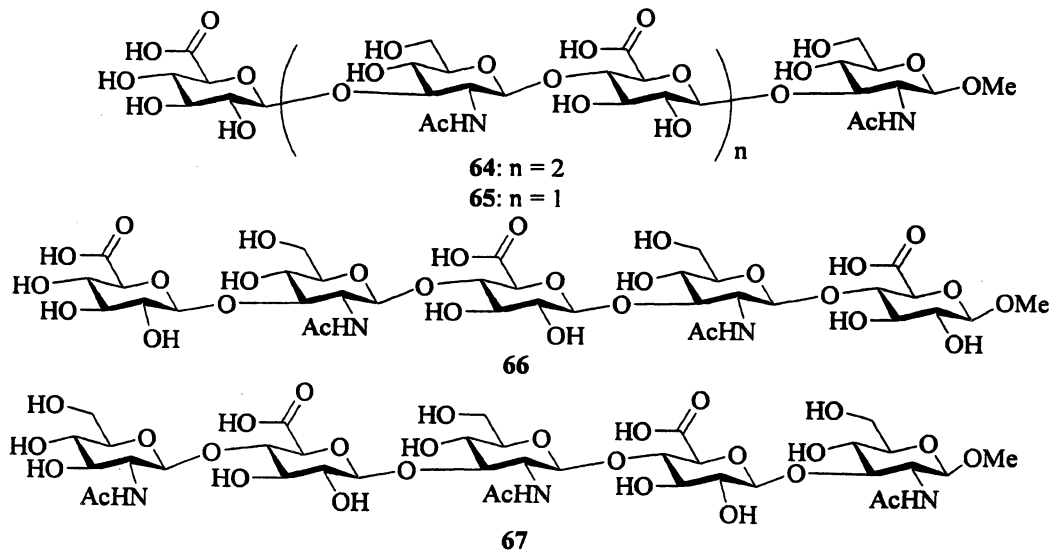
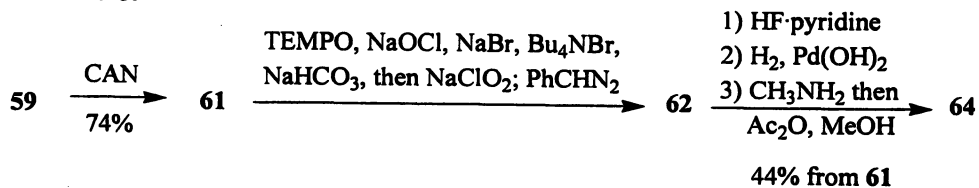
sequence of desilylation, hydrogenation, saponification and N-acetylation to produce the sHA hexasaccharide **64**. This procedure for deprotection and oxidation state adjustment has also been extended to generate sHA tetrasaccharide **65** and pentasaccharides **66** and **67**.

Conclusion

It is evident that sHA play important roles in many interesting biological events. Compared with intensive SAR studies on heparin, the importance of sHA is mostly underappreciated in the chemistry community. The developments of new methodologies for the assemblies of sHA would accelerate the exploration of its biological functions. In contrast to traditional syntheses of sHA, the advantages of iterative one-pot synthesis of sHA are obvious with its speed of glycoassembly and scalability. Another important feature of our strategy was the usage of thioglycosides as the glycosyl donors, which are one of the most stable glycosyl donors (49). Furthermore, the development of a two-step one-pot oxidation method via the combination TEMPO/NaOCl and NaClO₂ has been demonstrated to facilitate the oxidations of multiple hydroxyl groups, which will also benefit synthesis of other GAGs. With chemical approaches, various positions of sHA can be differentiated via protective groups, which can allow for access to structurally varied, well defined sHA derivatives facilitating SAR studies.

Nevertheless, sHA synthesis is yet to become routine practice. It is still non-trivial to assemble sHA longer than an octasaccharide unit, which is necessary to study the interaction of sHA with CD44. With the increase in sHA chain length, the solubility and reactivity of the oligosaccharide decrease, which

Scheme 8.



can affect the glycosylation as well as deprotection. Further studies are necessary for the continual development of sHA glyco-assembly strategies as well as protecting group chemistry.

References

1. Meyer, K.; Palmer, J. W. *J. Biol. Chem.* **1934**, *107*, 629-634.
2. Lapcik, L., Jr. ; Lapcik, L.; De Smedt, S.; Demeester, J.; Chabreck, P. *Chem. Rev.* **1998**, *98*, 2663-2684.
3. Laurent, T. C.; Fraser, J. R. E. *FASEB J.* **1992**, *6*, 2397-2404.
4. Wright, A. J.; Day, A. J. In *Glycobiology and Medicine: Proceedings of the 7th Jenner Glycobiology and Medicine Symposium* Axford, J. S., Ed.; Springer: Dordrecht, 2005; pp 57-69.
5. Hascall, V. C.; Majors, A. K.; de la Motte, C. A.; Evanko, S. P.; Wang, A.; Drazba, J. A.; Strong, S. A.; Wight, T. N. *Biochim. Biophys. Acta* **2004**, *1673*, 3-12.
6. In *Chemistry and Biology of Hyaluronan*; Garg, H. G., Hales, C. A., Eds.; Elsevier: Oxford, 2004.
7. West, D. C.; Fan, T.-P. D. In *New Angiotherapy*; Fan, T.-P. D., Kohn, E. C., Eds.; Humana Press: Totowa, N.J., 2002; pp 177-188.
8. Toole, B. P.; Wight, T. N.; Tammi, M. I. *J. Biol. Chem.* **2002**, *277*, 4593-4596.
9. Lee, J. Y.; Spicer, A. P. *Curr. Opin. Cell Biol.* **2000**, *12*, 581-586.
10. Toole, B. P. In *Carbohydrates in Chemistry and Biology*; Ernst, B., Hart, G. W., Siney, P., Eds.; Wiley-VCH: Weinheim, 2000; Vol. 4, pp 685-700.
11. Wight, T. N. In *Carbohydrates in Chemistry and Biology*; Ernst, B., Hart, G. W., Siney, P., Eds.; Wiley-VCH: Weinheim, 2000; Vol. 4, pp 743-756.
12. *Glycoforum. Hyaluronan Today*, <http://www.glycoforum.gr.jp/science/hyaluronan/hyaluronanE.html>.
13. Slevin, M.; Kumar, S.; Gaffney, J. *J. Biol. Chem.* **2002**, *277*, 41046-41059.
14. Taylor, K. R.; Trowbridge, J. M.; Rudisill, J. A.; Termeer, C. C.; Simon, J. C.; Gallo, R. L. *J. Biol. Chem.* **2004**, *279*, 17079-17084.
15. Termeer, C.; Benedix, F.; Sleeman, J.; Fieber, C.; Voith, U.; Ahrens, T.; Miyake, K.; Freudenberg, M.; Galanos, C.; Simon, J. C. *J. Exp. Med.* **2002**, *195*, 99-111.
16. Termeer, C. C.; Hennies, J.; Voith, U.; Ahrens, T.; Weiss, J. M.; Prehm, P.; Simon, J. C. *J. Immunol.* **2000**, *165*, 1863-1870.
17. Zeng, C.; Toole, B. P.; Kinney, S. D.; Kuo, J.; Stamenkovic, I. *Int. J. Cancer* **1998**, *77*, 396-401.
18. Ghatak, S.; Misra, S.; Toole, B. P. *J. Biol. Chem.* **2002**, *277*, 38013-38020.
19. Ponta, H.; Sherman, L.; Herrlich, P. A. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 33-45.

20. Naor, D.; Nedvetzki, S. *Arthritis Res. Therapy* **2003**, 105-115.
21. Bajorath, J. *Proteins: Struct., Funct., Genetics* **2000**, 39, 103-111.
22. Teriete, P.; Banerji, S.; Noble, M.; Blundell, C. D.; Wright, A. J.; Pickford, A. R.; Lowe, E.; Mahoney, D. J.; Tammi, M. I.; Kahmann, J. D.; Campbell, I. D.; Day, A. J.; Jackson, D. G. *Mol. Cell* **2004**, 13, 483-496.
23. Banerji, S.; Wright, A. J.; Noble, M.; Mahoney, D. J.; Campbell, I. D.; Day, A. J.; Jackson, D. G. *Nat. Struc. Mol. Biol.* **2007**, 14, 234-239.
24. Lesley, J.; Hascall, V. C.; Tammi, M.; Hyman, R. *J. Biol. Chem.* **2000**, 257, 26967-26975.
25. Xu, H.; Ito, T.; Tawada, A.; Maeda, H.; Yamanokuchi, H.; Isahara, K.; Yoshida, K.; Uchiyama, Y.; Asari, A. *J. Biol. Chem.* **2002**, 277, 17308-17314.
26. Mahoney, D. J.; Aplin, R. T.; Calabro, A.; Hascall, V. C.; Day, A. J. *Glycobiology* **2001**, 11, 1025-1033.
27. Tawada, A.; Masa, T.; Oonuki, Y.; Watanabe, A.; Matsuzaki, Y.; Asari, A. *Glycobiology* **2002**, 12, 421-426.
28. Luca, C. D.; Lansing, M.; Martini, I.; Crescenzi, F.; Shen, G.-J.; O'Regan, M.; Wong, C.-H. *J. Am. Chem. Soc.* **1995**, 117, 5869-5870.
29. Saitoh, H.; Takagaki, K.; Majima, M.; Nakamura, T.; Matsuki, A.; Kasai, M.; Narita, H.; Endo, M. *J. Biol. Chem.* **1995**, 270, 3741-3747.
30. Kobayashi, S.; Ohmae, M.; Ochiai, H.; Fujikawa, S. *Chem. Eur. J.* **2006**, 12, 5962-5971.
31. Ochiai, H.; Ohmae, M.; Mori, T.; Kobayashi, S. *Biomacromolecules* **2005**, 6, 1068-1084.
32. Kobayashi, S.; Morii, H.; Itoh, R.; Kimura, S.; Ohmae, M. *J. Am. Chem. Soc.* **2001**, 123, 11825-11826.
33. DeAngelis, P. L.; Oatman, L. C.; Gay, D. F. *J. Biol. Chem.* **2003**, 278, 35199-35203.
34. Karst, N. A.; Linhardt, R. J. *Curr. Med. Chem.* **2003**, 10, 1993-2031.
35. Yeung, B. K. S.; Chong, P. Y. C.; Petillo, P. A. In *Glycochemistry. Principles, Synthesis, and Applications*; Wang, P. G., Bertozzi, C. R., Eds.; Marcel Dekker, Inc.: New York City, 2001; pp 425-492.
36. Codee, J. D. C.; van den Bos, L. J.; Litjens, R. E. J. N.; den Heeten, R.; Overkleeft, H. S.; van Boom, J. H.; van der Marel, G. A. *Org. Lett.* **2003**, 5, 1947-1950.
37. Yeung, B. K. S.; Hill, D. C.; Jankcka, M.; Petillo, P. A. *Org. Lett.* **2000**, 2, 1279-1282.
38. Halkes, K. M.; Slaghek, T.; Hypponen, T. K.; Kruiskamp, P. H.; Ogawa, T.; Kamerling, J. P.; Vliegthart, J. F. G. *Carbohydr. Res.* **1998**, 309, 161-174.
39. Slaghek, T.; Hypponen, T. K.; Ogawa, T.; Kamerling, J. P.; Vliegthart, J. F. G. *Tetrahedron Assym.* **1994**, 5, 2291-2301.

40. Slaghek, T.; Nakahara, Y.; Ogawa, T.; Kamerling, J. P.; Vliegthart, J. F. G. *Carbohydr. Res.* **1994**, *255*, 61-85.
41. Slaghek, T.; Hypponen, T. K.; Ogawa, T.; Kamerling, J. P.; Vliegthart, J. F. G. *Tetrahedron Lett.* **1993**, *34*, 7939-7942.
42. Slaghek, T.; Nakahara, Y.; Ogawa, T. *Tetrahedron Lett.* **1992**, *33*, 4971-4974.
43. Blatter, G.; Jacquinet, J.-C. *Carbohydr. Res.* **1996**, *288*, 109-125.
44. Palmacci, E. R.; Seeberger, P. H. *Tetrahedron* **2004**, *60*, 7755-7766.
45. Huang, X.; Huang, L.; Wang, H.; Ye, X.-S. *Angew. Chem. Int. Ed.* **2004**, *43*, 5221-5224.
46. Zhang, Z.; Ollmann, I. R.; Ye, X.-S.; Wischnat, R.; Baasov, T.; Wong, C.-H. *J. Am. Chem. Soc.* **1999**, *121*, 734-753.
47. Huang, L.; Huang, X. *Chem. Eur. J.* **2007**, *13*, 529-540.
48. Huang, L.; Teumelsan, N.; Huang, X. *Chem. Eur. J.* **2006**, *12*, 5246-5252.
49. Garegg, P. J. *Adv. Carbohydr. Chem. Biochem.* **1997**, *52*, 179-205.

Chapter 3

Synthetic Efforts in Preparations of Components of the Bacterial Cell Wall

Mijoon Lee, Dusan Heseck, and Shahriar Mobashery*

Department of Chemistry and Biochemistry, University of Notre Dame,
Notre Dame, IN 46556

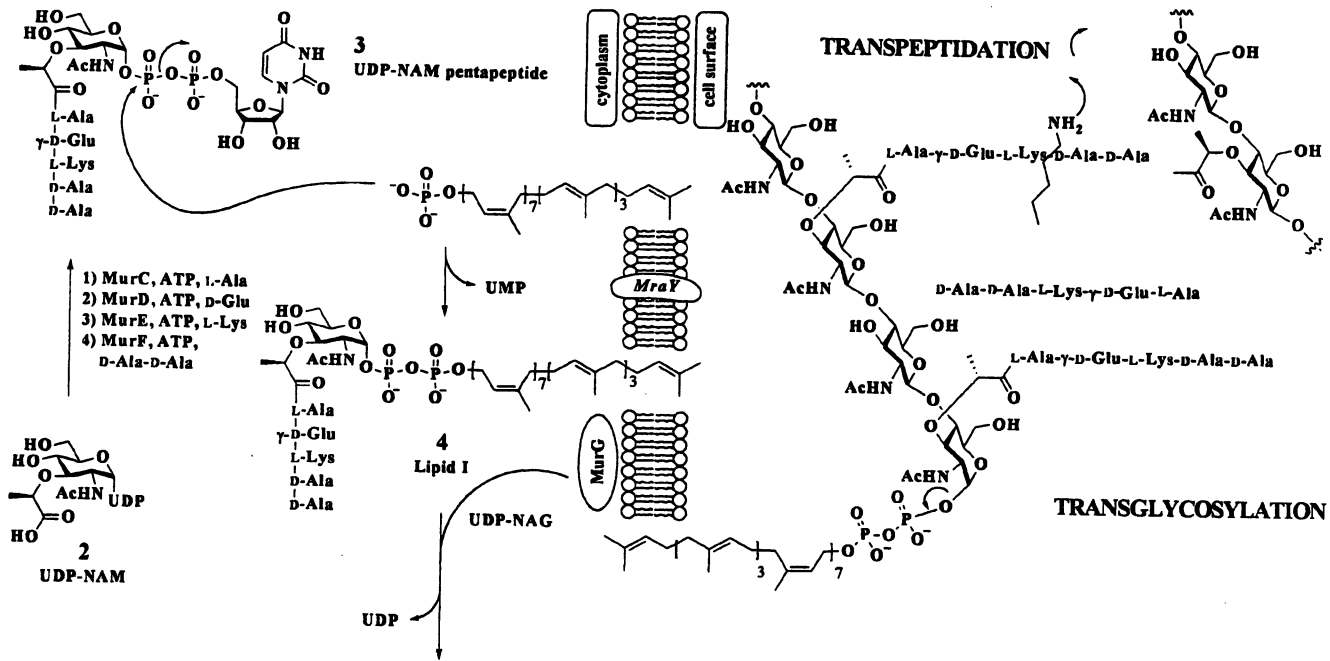
Bacterial cell wall is a marvel of nature. It is a single molecule that encases the entire organism. Its size in *Escherichia coli*—3300 kDa—is larger than that of the chromosome. The complexity of this polymeric entity has impeded progress in understanding its properties, despite the fact that it is a critical structural component of bacteria. A healthy cell wall is vital for bacterial survival and since the cell wall is a unique biological structure, the enzymes of its biosynthesis and the cell wall itself are targets for inhibition by antibiotics. Despite a clear realization of the importance of cell wall to bacteria in the past 50 years, its structure has not been amenable to elucidation. This is in part due to the large structure of the cell wall. Previous studies relied on growing bacteria, fragmenting the cell wall, and attempting purification of the fragments for analysis. This has not been fruitful as the amounts of the sample were invariably small and the purity often was problematic. The more recent efforts in the past 10 years have centered on preparing synthetic samples for analysis. These efforts are bearing fruit only recently.

The basic chemical structure of the cell wall is comprised of cross-linked strands of peptidoglycan. Peptidoglycan itself, as the name indicates, is built of a saccharide backbone to which peptide stems are attached (Figure 1). The saccharide backbone is comprised of alternating *N*-acetylglucosamine (NAG) and *N*-acetylmuramic acid units (NAM). The NAG-NAM repeating units are strung together to varying degrees, with nine repeats being typical for *E. coli*. The peptide stem, typically a pentapeptide, is attached to the carboxylate of the NAM unit. The peptide has unusual amino acids such as two D-Ala residues and a D-Glu, which incidentally participates in bonding via its side chain carboxylate. There are approximately 3.5 million peptidoglycan strands in *E. coli*. Transpeptidases cross-link the peptides of the neighboring strands of the peptidoglycan, whereby creating a network of covalency among all strands of the peptidoglycan, which provides the structural support needed by bacterial.

The biosynthesis of the cell wall has been the subject of excellent reviews and will not be discussed here in any great lengths (1, 2). However, the biosynthetic process starts with UDP-*N*-acetylglucosamine (UDP-NAG, 1), which experiences the incorporation of the D-lactyl moiety at C₃ by the action of the enzymes MurA and MurB to give UDP-*N*-acetylmuramic acid (UDP-NAM, 2). In five sequential ATP-dependent steps catalyzed by enzymes MurC to MurF the amino acids of the peptide stem are incorporated onto the lactate moiety of NAM to result in the key intermediate 3. All these steps progress in the cytoplasm. The function of the integral membrane protein MraY on 3 incorporates undecaprenyl phosphate, such that the resultant product (lipid I, 4) is now anchored on the inner leaflet of the cytoplasmic membrane. Another membrane enzyme, MurG, incorporates NAG onto lipid I, giving rise to the key building block for the cell wall, namely lipid II (5).

Lipid II is translocated to the outer leaflet of the cytoplasmic membrane by a process that is not understood. It has been argued that a protein might facilitate this process. However, once it is translocated, it serves as the substrate for transglycosylases, which is a processive event incorporate several NAG-NAM units into the nascent peptidoglycan, in the course of which the undecaprenyl pyrophosphate is released and recycled. As the polymerization process catalyzed by the transglycosylases proceed, transpeptidases cross-link the peptide stems from two neighboring peptidoglycan strands to result in the network of the growing cell wall that is critical for the survival of bacteria.

The complexity of the problem in hand is immediately obvious. The subject of study, the bacterial cell wall, is a large polymer with a number of distinct constituents, which organize themselves into a three-dimensional entity. Despite the many attempts at elucidation of the three-dimensional structure by many methods, none had yielded definitive results. In order to make inroads into this complex problem, several groups have attempted to prepare synthetic samples of the building units and of fragments of the larger cell wall itself.



- 1) MurC, ATP, L-Ala
- 2) MurD, ATP, D-Glu
- 3) MurE, ATP, L-Lys
- 4) MurF, ATP, D-Ala-D-Ala

TRANSEPTIDATION

TRANSGLYCOSYLATION

cytoplasm
cell surface

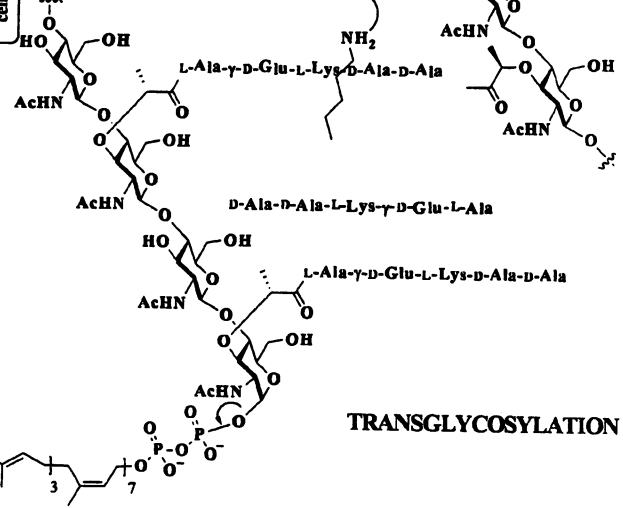
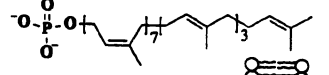
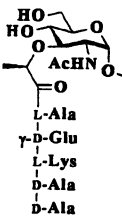
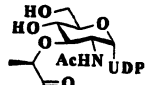
3
UDP-NAM pentapeptide

4
Lipid I

2
UDP-NAM

MraY

MurG



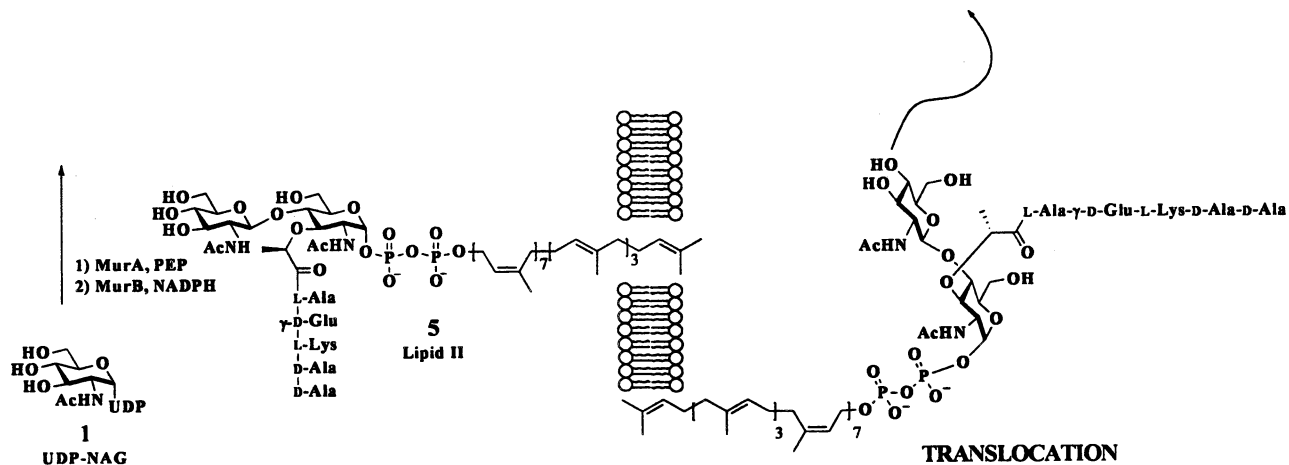


Figure 1. The bacterial cell wall biosynthesis pathway.

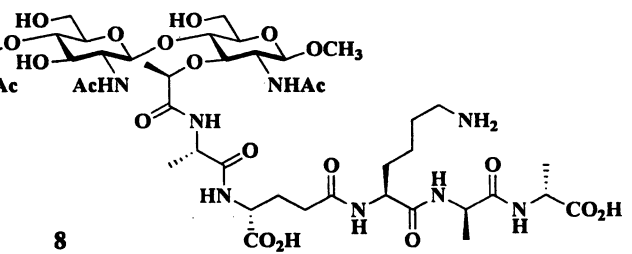
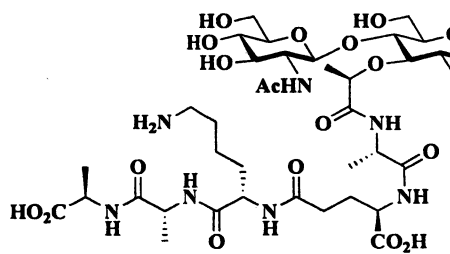
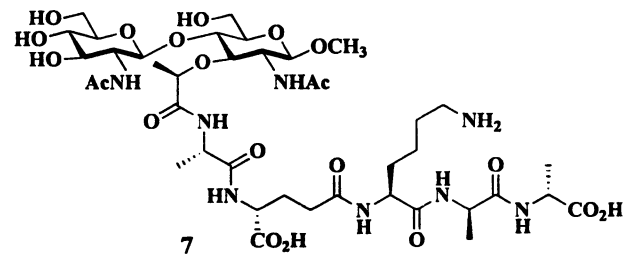
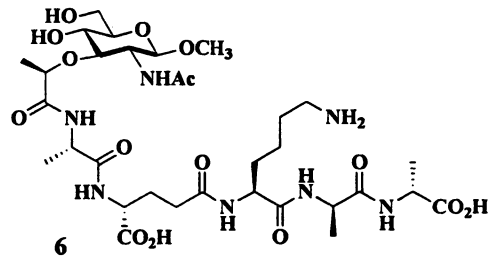
Efforts in Preparation of Fragments of the Cell Wall

The length of the chemically synthesized polysaccharide portion of the peptidoglycan ranges from 1 to 16 (as many as 8 repeats of NAG-NAM unit). For the peptides portion, syntheses range from a simple dipeptide (L-Ala- γ -D-Glu or L-Ala- γ -D-Gln) to the full pentapeptide (L-Ala- γ -D-Glu/Gln-L-Lys-D-Ala-D-Ala, plus the diaminopimelate type in place of L-Lys), as well as the cross-linked peptide through the pentaglycine-bridge by standard peptide coupling methods (3).

In contrast to the enzymic chemistry that assembles units without the need for protective groups, chemical syntheses of the peptidoglycan cannot avoid using orthogonal protective groups. There are a number of issues in construction of polymeric NAG-NAM repeats. First, it is critical to construct selectively β -1,4-linkage—found exclusively in the peptidoglycan—between monosaccharides using suitable coupling methods, with or without neighboring group participation of the amino protecting group at C₂ of the glucosamine unit. The formation of the 1,4-glycosidic bond is often challenging, since the hydroxyl at C₄ in glucosamine ring is sterically hindered and often the least reactive among all the hydroxyl groups in the ring. Second, it is important to settle on the suitable point in the course of synthesis for the introduction of the lactyl moiety at the C₃ in the NAM ring, which will affect the nature of the orthogonal protecting group strategy. Third, the choice of amino group protection is key, since it can facilitate the formation of selective β -1,4-glycoside bond via neighboring group participation, as well as affecting the solubility of the synthetic intermediates. Commonly used protective groups are DMM (dimethylmaleoyl), Troc (2,2,2-trichloroethoxycarbonyl), acetyl, azido, Phth (phthalimido), or TCP (tetrachlorophthalimido), etc.

We completed recently the synthesis of NAG-MPP-NAG-MPP (4) (MPP, muramyl pentapeptide, NAM-L-Ala- γ -D-Glu-L-Lys-D-Ala-D-Ala) and NAG-MPP and MPP (compounds 6-8) (5, 6). These samples were intended as fragments of the larger cell wall, hence the reducing end of the saccharides had to be tied up as methyl glycosides, also as β -(1-4) linkage. This capping of the sugars would retain the structural integrity of the backbone as all β -(1-4) glycosides.

A key strategy used in our syntheses is the introduction of the lactyl ether at 3-O position late in the synthesis to generate NAM. This step takes place after the construction of the glycoside bond, in contrast to most of the other syntheses in the literature. More typical has been the construction of the glycoside bond by coupling of protected glucosamine derivative and muramyl acid derivative, which is already incorporated with the lactate moiety at the 3-O-position of the sugar ring. The advantage of the late introduction of the lactate moiety is to prevent synthetic complications such as the intramolecular lactonization at the 4-hydroxyl position (7), intramolecular lactamization at the free amino group at C₂



(8) and more seriously racemization of the lactate residue (these side reactions will be discussed below with examples). This strategy prompted us to develop synthetic routes involving orthogonal protective groups that differentiate the hydroxyl groups at C₃ in NAM from those in the NAG unit.

The trichloroacetimidate method was used in the presence of trifluoromethanesulfonic acid (TfOH) as the main promoter to ensure selective β -(1-4)-glycoside bond formation. The *N*-dimethylmaleoyl (DMM) protective group was chosen to mask the amino group at C₂ of glucosamine ring, which participates in glycoside bond formation, as well as improves the solubility of polymeric sugar intermediates. The issue of solubility is not mentioned often in the literature, but this can be problematic when it comes to synthesis of polymeric glucosamine or to muramic acid derivatives and their scale-up synthesis. Phthalimido group, for example, which gives neighboring group participation but imparts poor solubility to the polymeric glucosamine/muramic acid intermediates, was not useful. Oxazoline glycosyl donor, though it has good solubility and is already integrated within it with the *N*-acetyl group at C₂ position after glycoside bond formation, was excluded for poor reactivity with the 4-hydroxyl group in glycosyl acceptor, resulting in poor glycosylation yields. These protective groups or intermediates may be useful for synthesis of mono or bis sugar units, but become problematic for synthesis of larger pieces.

Common protecting groups such as *O*-benzylidene-, *O*-acetyl-, and *O*-*tert*-butyldimethylsilyl were chosen to mask the free hydroxyl groups in glucosamine derivatives and were built to enable selective unmasking of the functional groups at the right time in the synthesis. All peripheral hydroxyl groups in glucosamines and all side chains in the pentapeptide were protected with the benzyl or the Cbz groups, which could be globally removed by catalytic hydrogenolysis in the final step of synthesis.

After a systematic analysis of assembly of suitable libraries of glycosyl donors and acceptors derived from D-glucosamine, construction of the disaccharide core of the tetrasaccharide appeared to be the most reasonable, when considering the need for minimal purification and improved coupling yield (for the strategic reasoning, consult citation (4)). The glucosamine-derived glycosyl donor **9** and the acceptor **10** were allowed to react in the presence of TfOH to give the disaccharide **11**. Compound **11** was converted to the glycosyl donor **12** after removal of the silyl group and the reaction with trichloroacetonitrile (Figure 2). The disaccharide **12** was then subjected to a second glycosylation reaction with the glycosyl acceptor **13** to give the trisaccharide **14**. Solvent-driven regioselective reductive ring opening of the 4,6-*O*-benzylidene group in **14** was carried out in the presence of the borane-trimethylamine complex (BH₃·NM₃) and BF₃·OEt₂, similarly to a reported method (9). This reaction resulted in the formation of the trisaccharide acceptor **15** with an unmasked 4-hydroxyl functionality (and not the 6-hydroxyl). Selectivity in this particular step plays an important role in constructing the last

β -(1-4)-glycoside bond in the tetrasaccharide backbone. If the reaction product were to be contaminated with the 6-hydroxyl compound, which incidentally is very difficult to separate from the desired 4-hydroxyl compound, this contaminant will be carried over to form an undesired β -(1-6)-glycoside bond. The purification in this case would have been both tricky and critically necessary for good product purity. The last glycosylation was carried out between the glycosyl donor 16 and acceptor 15 to afford the fully protected tetrasaccharide 17.

The complete NAG-NAM-NAG-NAM sequence (20) was in place after the removal of the DMM groups and the acetate groups at the C₃ position of the NAM rings, followed by incorporation of the lactate functionality at the C₃ position (Figure 3). *p*-Nitrophenyl activation of compound 20 was used for the formation of the amide bond between the lactate moiety and L-Ala of the protected pentapeptide. A final global deprotection of the fully protected tetrasaccharide pentapeptide 21 by hydrogenolysis over palladium gave the desired compound 8.

The monosaccharide pentapeptide 6 and disaccharide pentapeptide 7 were prepared in a similar manner, starting from compound 22 and 23, respectively (5). The D-isoGln incorporated pentapeptide (instead of D-Glu) and pentapeptide incorporated with mono-, bis-, and tetrasaccharide derivatives were also prepared (6). This variation in the structure is seen in some bacteria.

These synthetic cell wall fragments have been utilized extensively in various studies, including those with penicillin-binding proteins (5), studies of allostery by cell wall (6, 10), studies with the cell-wall binding antibiotic vancomycin (11), and in investigations of interactions with the peptidoglycan-binding proteins (12), to name a few. The solution structure of compound 8 was also determined by NMR to address structural aspect of the bacterial cell walls (13). The extended saccharide backbone conforms to a right-handed helical structure with three repeats of the NAG-NAM unit per turn of the helix. The peptides are largely unstructured and extend to the milieu.

The peptidoglycan and fragments of it exhibit pronounced immunostimulatory properties. The minimal structure required for the immunostimulation is a muramyl dipeptide (MDP, *N*-acetylmuramyl-L-Ala- γ -D-Gln). Many investigations were directed towards the synthesis of derivatives of MDP. Fukase *et al.* have synthesized several MDP derivatives, as well as polymeric MPP derivative (compounds 24-29) (14, 15) in these efforts.

Fukase and colleagues constructed distinct repeating NAG-NAM units effectively by the coupling of the key disaccharide 32, which was transformed to the glycosyl donor 33 and the acceptor 34, to form the tetrasaccharide 35. Octasaccharide 36 was readily formed when the tetrasaccharide 35 was subjected to the same sequence of reactions. Stereoselective glycosylation of the disaccharide units was achieved by neighboring group participation of the *N*-Troc group and appropriate reactivity of *N*-Troc-glucosaminyltrichloro-

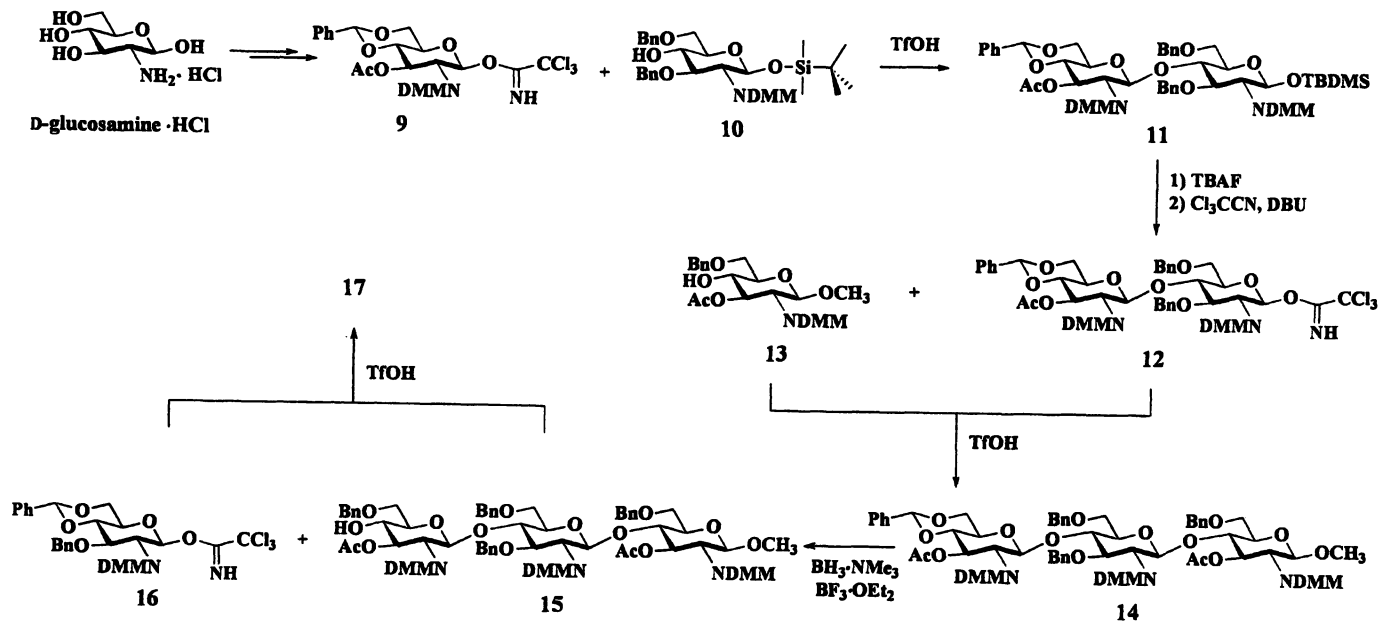


Figure 2. Synthesis of fully protected tetrasaccharide **17**.

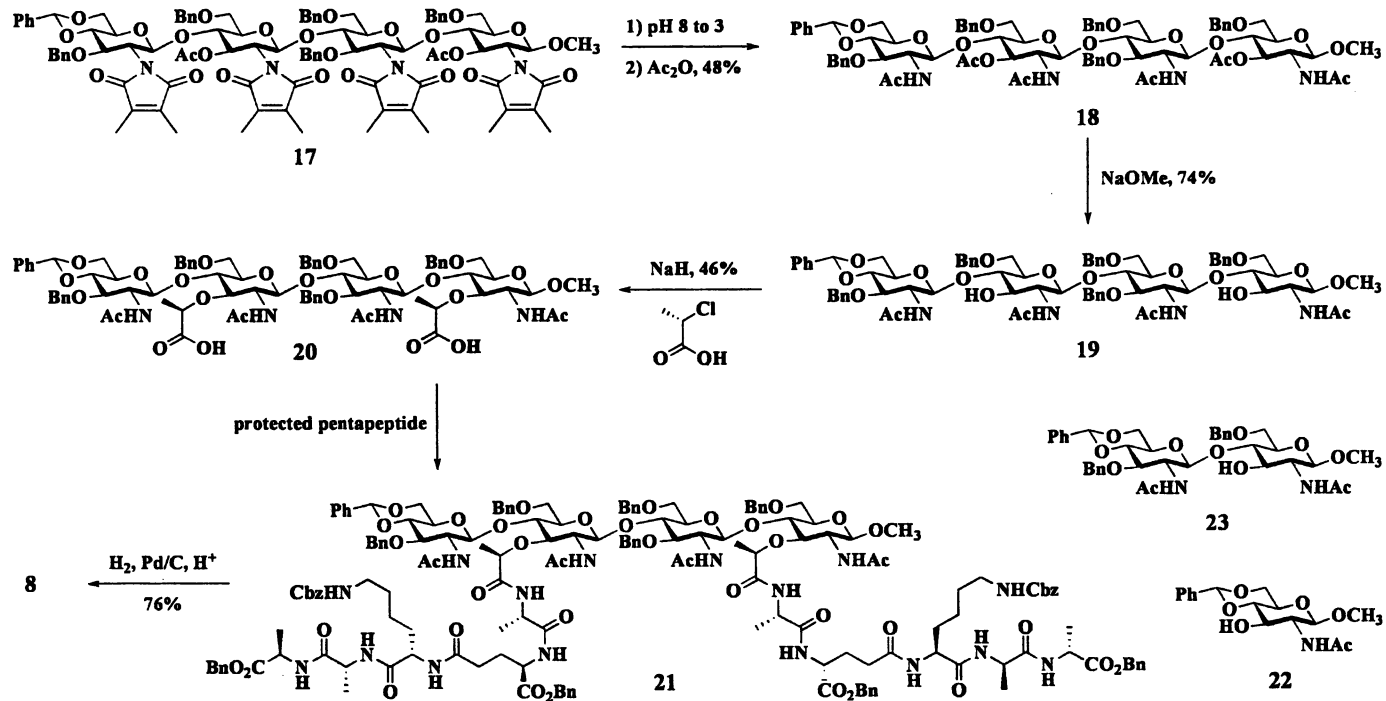
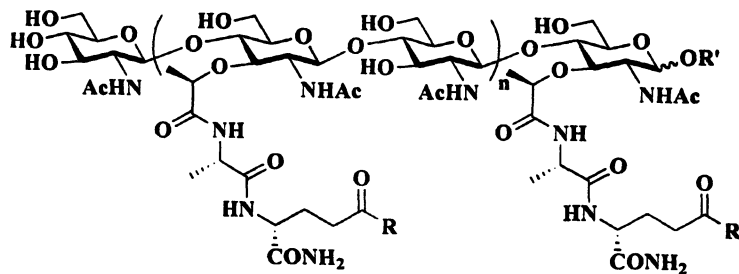


Figure 3. Synthesis of 8 as a fragment of the bacterial cell.



24 - 29

24, n = 2, R = OH, octasaccharide

25, n = 1, R = OH, tetrasaccharide

26, n = 1, R = -L-Lys

27, n = 1, R = -L-Lys-D-Ala

28, n = 1, R = -L-Lys-D-Ala-D-Ala

29, n = 0, R = -L-Lys, bissaccharide

R' = H or α -Pr

cetimide in the presence of TMSOTf (trimethylsilyl trifluoromethanesulfonate) as a promoter.

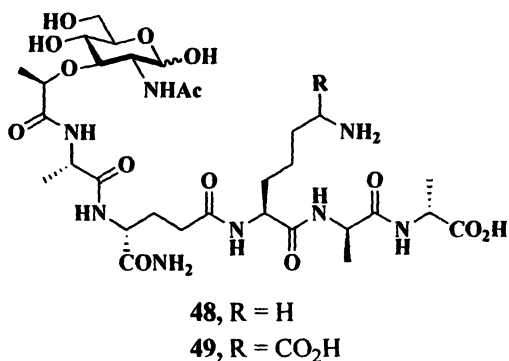
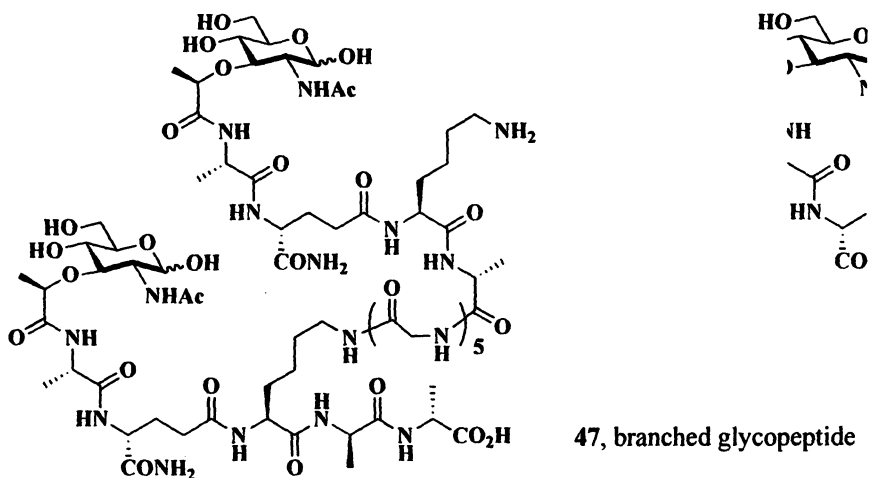
The key intermediate **32** was prepared by the glycosylation of the *N*-Troc-glycosyl donor **30** with *N*-Troc-muramyl acceptor **31**, which is already incorporated with the lactate arm, as shown in Figure 4. The tetrasaccharide **35** and octasaccharide **36** were then synthesized by multiplication of the disaccharide **32**. Deprotection of the *N*-Troc group in **32**, **35** and **36**, *N*-acetylation, saponification of the ethyl esters, peptide coupling, and then the final deprotection affording the bis-, tetra-, and octasaccharide with various peptide arms (**24-29**).

A further attempt was made by Fukase toward construction of hexadecasaccharide by glycosylation between the glycosyl acceptor and donor that were both derived from the octasaccharide **36**. Unfortunately, deprotection of the sixteen Troc groups in hexadecasaccharide by Zn-Cu in AcOH was not successful, giving a complex mixture of incompletely deprotected products having a dichloroethoxycarbonyl group. Other synthetic strategies and deprotection methods for the synthesis of the hexadecasaccharide having eight dipeptide units are presently under investigation.

Schmidt and co-worker have reported the synthesis of NAG-NAM-NAG-NAM as methyl lactate (**46**) (8). The β -1,4-linkage was constructed between the appropriately protected glucosamine and muramyltrichloroacetimidate derivatives based on the template of azidoglucose in the presence of $\text{BF}_3 \cdot \text{Et}_2\text{O}$ as a promoter (Figure 5). Yields varies (31-80%), depending on the reactivity of the glycosyl donors and acceptors. Transformation of four azido groups in the tetrasaccharide **44** into the acetamido group caused problems such as ring closure between the lactate and the newly generated amino group in the muramic acid moiety (**45**) under a typical azide reduction condition (H_2S in pyridine). However, selective reduction of **44** with palladium black was carefully investigated, thus affording the desired free amino group, which was subsequently acetylated. The final hydrogenolysis gave the desired compound **46**. The final compound was characterized as a per-*O*-acetate derivative after treatment with acetic anhydride in pyridine. Schmidt and co-workers have made contributions in syntheses of meso-diaminopimelic acid (DAP), DAP-containing muramyl peptide, and 1,6-anhydromuramyl peptide more recently (16).

Boons and co-workers synthesized an analog of the structure of the cross-linked peptidoglycan (17) as well as MPP (both Lys and meso-DAP type and their truncated peptide analogs, bis-, tri-, tetrapeptides) (3) using hyperacid sensitive Sieber amide resin (18) as solid-phase support.

In the synthesis of branched glycopeptide, the resin-bound compound **50** was obtained through a series of steps using standard coupling chemistry (PyBOP, HOBt, DIEPA), Fmoc-protected amino acids, starting from Fmoc-D-Ala, and a suitably protected NAM derivative (2-*N*-acetyl-1- β -O-allyl-4,6-benzylidene-3-muramic acid) at the end (Figure 6). Dotted arrows indicate the



direction of amide bond formation. Upon removal of the ivDde (1-(4,4-dimethyl-2,6-dioxycyclo-hex-1-ylidene)-3-methylbutyl) protecting group on the ϵ -side chain of lysine of **50** by reaction with 2% hydrazine, the pentaglycine bridge of the branched glycopeptides **51** was constructed in a stepwise fashion. After incorporation of the second peptide stem, the protected NMA derivative was added to complete the protected branched glycopeptides **51**. The removal of the branched glycopeptides from the solid support was done by treatment with 2% TFA and the anomeric allyl group was deprotected using Pd/C in EtOH:water:HCl. Finally, purification by size exclusion chromatography resulted in the target compound as a mixture of α/β anomers.

MPPs **48** and **49** were prepared by a similar approach employing either Fmoc-L-Lys(Mtt) (Mtt = 4-methyltrytyl) or Fmoc-DAP(Boc, ^tBu) (**19**).

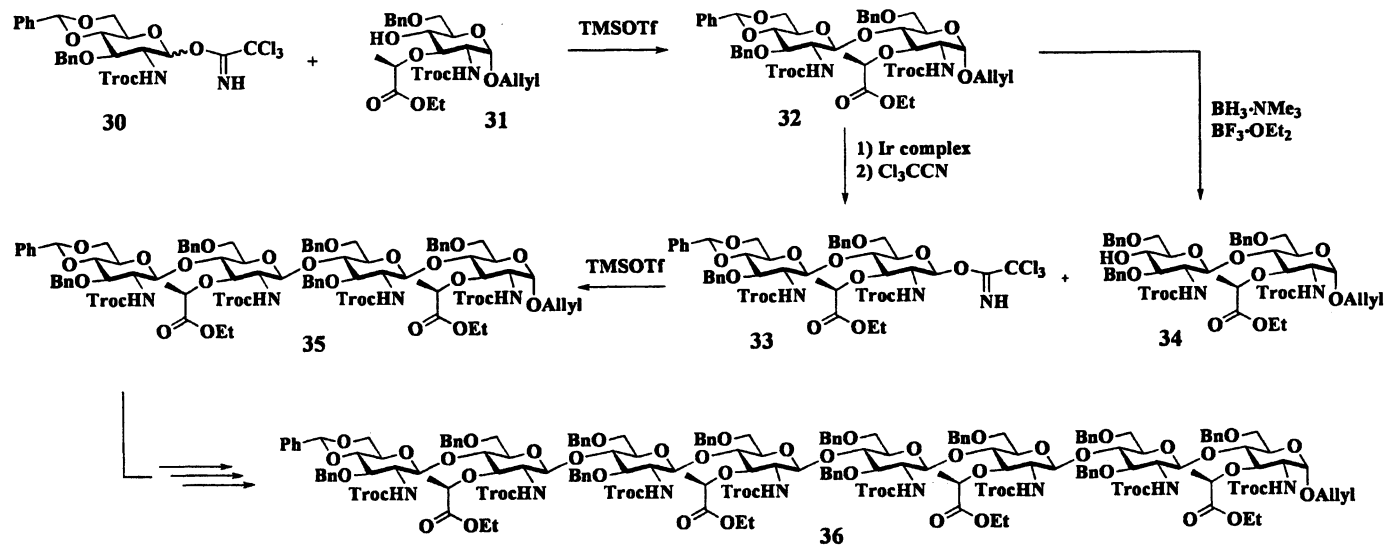


Figure 4. Synthesis of octasaccharide 36.

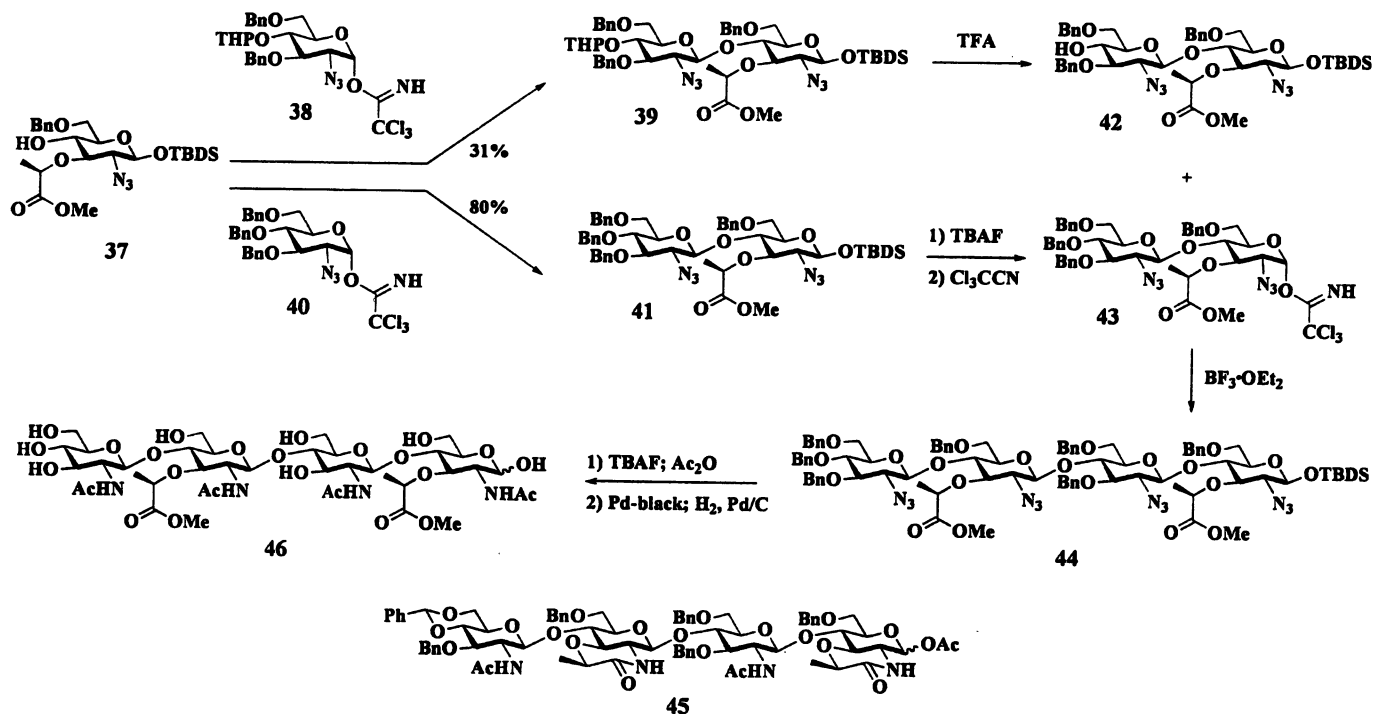


Figure 5. Schmidt's synthesis of NAG-NAM-NAG-NAM 46.

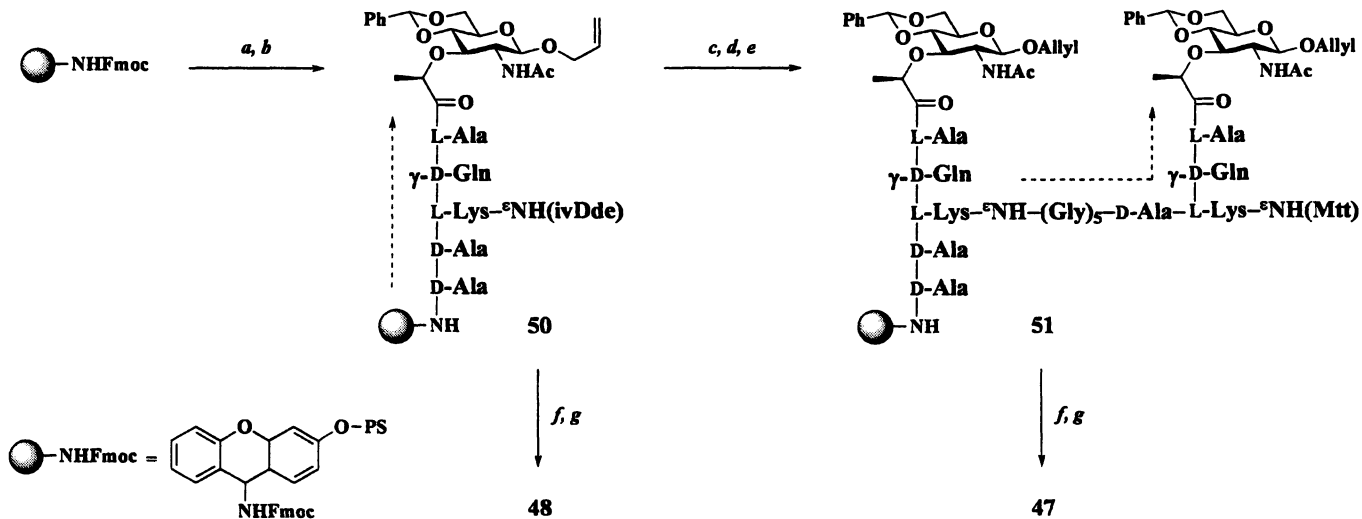
Syntheses of Lipid I, II and IV

Total syntheses of lipid I and II were accomplished by both chemical (20-22) and chemoenzymatic strategies (23) by several research groups. This subject has been reviewed previously (24-26).

The presence of the unsaturated undecaprenylside chain makes lipids I and II challenging synthetic targets. As a key strategy for the synthesis of lipid I (21), the undecaprenyl-linked diphosphate moiety was introduced at a late stage in the synthesis to avoid potential solubility complications caused by the enhanced lipophilic character of the undecaprenyl-linked substrate. The anticipated acid sensitivity of the anomeric diphosphate dictated the use of base-cleavable protective groups for all peripheral functionality, including the side chains of peptapeptide.

The muramyl derivative 53, with triple orthogonal protective groups, was converted to the desired phosphate (54) by successive anomeric deprotection and phosphorylation (Figure 7). A phosphitylation/oxidation protocol was utilized to introduce the α -phosphate ester. After hydrogenolysis of 53, the anomeric hydroxyl group was allowed to react with dibenzyl *N,N*-diethylphosphoramidate in the presence of tetrazol to afford the anomeric phosphite, which was oxidized *in situ* to the corresponding phosphate 54 by hydrogen peroxide. The carboxyl protecting group of lactate was removed by the DBU-mediated elimination, and the resulting acid was coupled to the protected pentapeptide to yield the desired glycopeptide (55). Among several methods to construct the lipid-tethered glycosyl diphosphates, the phosphoroimidazolidate protocol was selected, because it can proceed through initial activation of the carbohydrate component under mild condition. The phosphoryl group was freed, activated with CDI and subsequently reacted with the commercially available undecaprenyl monophosphate. Finally, removal of the peptidyl and hydroxyl protecting groups by treatment with aqueous NaOH furnished lipid I.

The total syntheses of lipid II were published independently by two groups, which followed essentially the same approach (20, 22). The main strategy is not different from the synthesis of lipid I, except for the use of compound 60 instead of 52, which has D-Lac-L-Ala at C₃ (Figure 8). When 57 was treated with trifluoroacetic acid and triethylsilane to convert to the glycosyl acceptor 59, lactone 58 was isolated as the major product along with a small amount of the desired 59 (7). The simple expedience of installing L-Ala at the lactyl position completely suppressed the acid-catalyzed cyclization of the reduction product. Thus, subjecting 60 to the same reductive ring opening conditions afforded the muramylmono-peptide ester 61 in 61% yield, which was combined with the glycosyl donor 62 under rigorously anhydrous Königs-Knorr conditions (silver triflate) resulting in the desired β -linked disaccharide 63.



(a) 20% piperidine in DMF. (b) PyBOP, HOBt, DIPEA, and AA (each 5) followed by 20% piperidine in DMF. AA is one of the following compounds: Fmoc-D-Ala, Fmoc-L-Lys(ivDde), Fmoc-D-isoGln, Fmoc-L-Ala, and MurNAc. (c) 2% hydrazine in DMF. (d) Five cycles of HBTU, NMP, and Fmoc-Gly followed by 20% piperidine in DMF. (e) PyBOP, HOBt, DIPEA and AA (each 5) followed by 20% piperidine in DMF. AA is one of the following compounds: Fmoc-D-Ala, Fmoc-L-Lys(Mtt), Fmoc-D-isoGln, Fmoc-L-Ala, and MurNAc. (f) 2% TFA, 1% TIS, DCM. (g) Pd/C, EtOH:H₂O:1(N)HCl (2:1:0.01).

Figure 6. Solid-phase synthesis of branched peptidoglycan derivative 47.

Compound **63**, after the removal of the Troc group, was subjected to same sequence of reactions shown in Figure 7 (i.e., introduction of benzylonophosphate, tetrapeptide coupling, lipid introduction, and final basic deprotection) to give the desired lipid II (**20**).

The total synthesis published by Schwartz and co-workers followed essentially the same approach, except they used fluoride-sensitive protective groups (such as TMSE (trimethylsilylethyl) and TEOC (trimethylsilylthoxycarbonyl) for pentapeptide side chains) to avoid the use of a strong base such as sodium hydroxide during the final deprotection step (**22**).

A chemoenzymatic preparation of lipid II was carried out by Walker, Kahne and co-workers (**23**). Compound **64**, which was prepared from muramyl derivative **52** (**27**) in a similar manner to lipid I described above (Figure 7), was reacted with undecaprenyl phosphoroimidazolidate activated by the CDI method (Figure 9). After deprotection of the peptide side chains using TBAF, the resultant lipid I was converted to lipid II by the reaction with UDP-(¹⁴C)-GlcNAc catalyzed by the purified MurG from *E. coli*. A series of lipid II analogues with different length of lipid side chains was prepared in this manner.

More recently, Walker, Kahne and co-workers reported the synthesis of heptaprenyl-lipid IV (**66**) in their studies of penicillin-binding proteins (**28**).

Though the enzymatic construction of lipid II from lipid I using a purified enzyme was successful, enzymic handle over longer glycan polymers with control over the product length does not exist. Therefore, all the glycosidic bonds in lipid IV were constructed synthetically. β -1,4-Glycosidic bond in tetrasaccharide **73** (Figure 10) as a key intermediate was constructed using the sulfoxide glycosylation method. The use of the bulky tetrachlorophthalimido (TCP) group to the mask amino group at C₂ and careful tuning of the reactivity of the hydroxyl in partially protected glycosyl donors (**68**, **69**, and **72**) and acceptors (**67** and **71**) at several different steps enabled a convergent synthesis of the tetrasaccharide **73** with a minimum number of protecting group manipulations. The tetrasaccharide **73** was then subjected to a similar sequence of reactions (Figures 3 and 7) and a final TBAF deprotection to result in the heptaprenyl-lipid IV (**66**).

The chemistry of polysaccharides remains far from having become routine. Each of these targets have to be designed carefully with detailed synthetic strategies delineated for each case. What has been described in this chapter is a summary of the successful efforts on the subject of the preparation of the pieces and components of the bacterial cell wall for the purposed of the elucidation of the biological processes that involve its construction or recognition. These literature entries no doubt make such efforts possible and represent pioneering efforts in elucidation of the biological events influenced by bacteria.

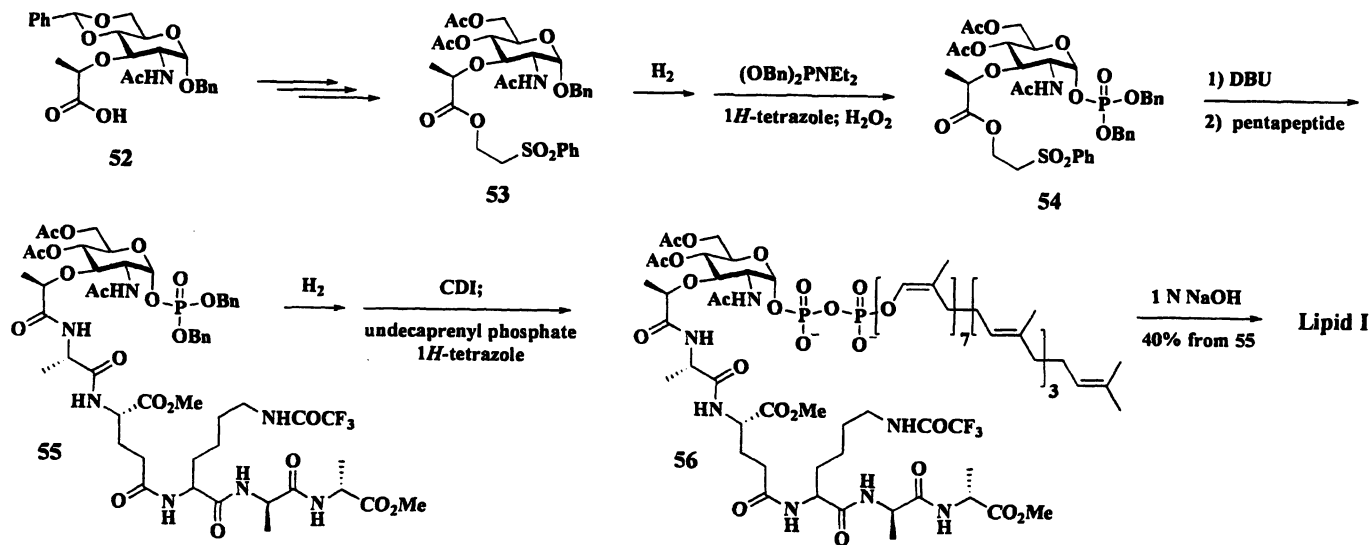


Figure 7. Synthesis of Lipid I.

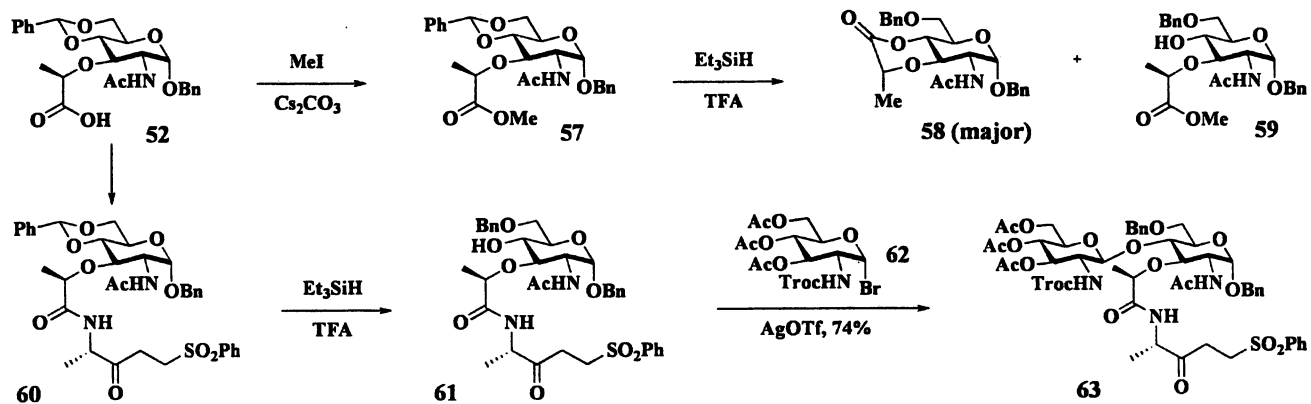


Figure 8. Reductive opening of 4,6-O-benzylidene acetal.

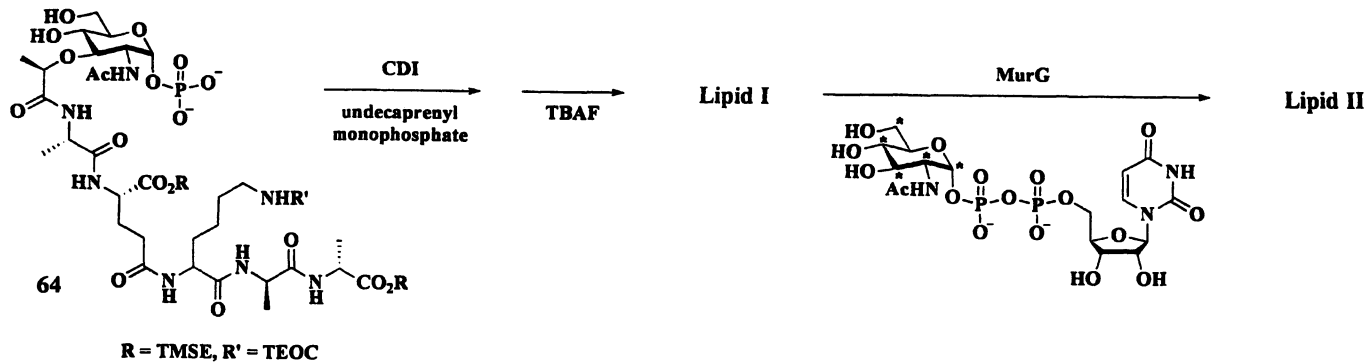
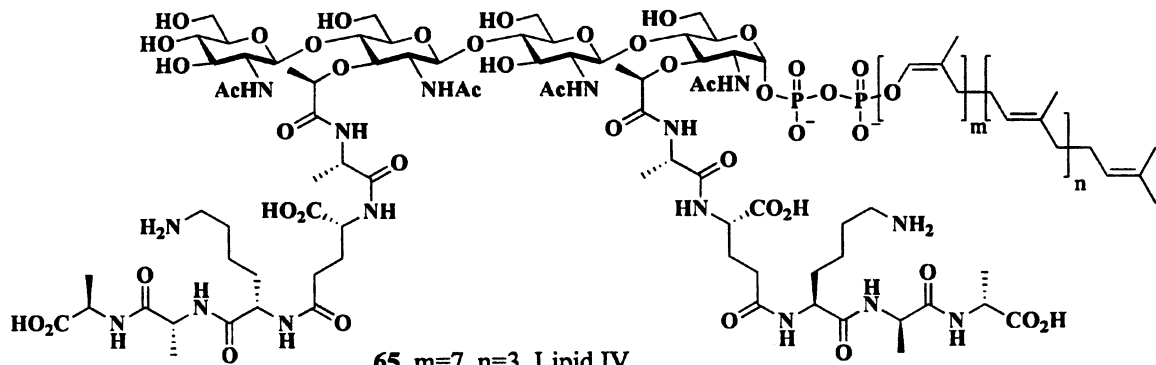


Figure 9. Chemoenzymatic synthesis of Lipid II using MurG from *E. coli*.



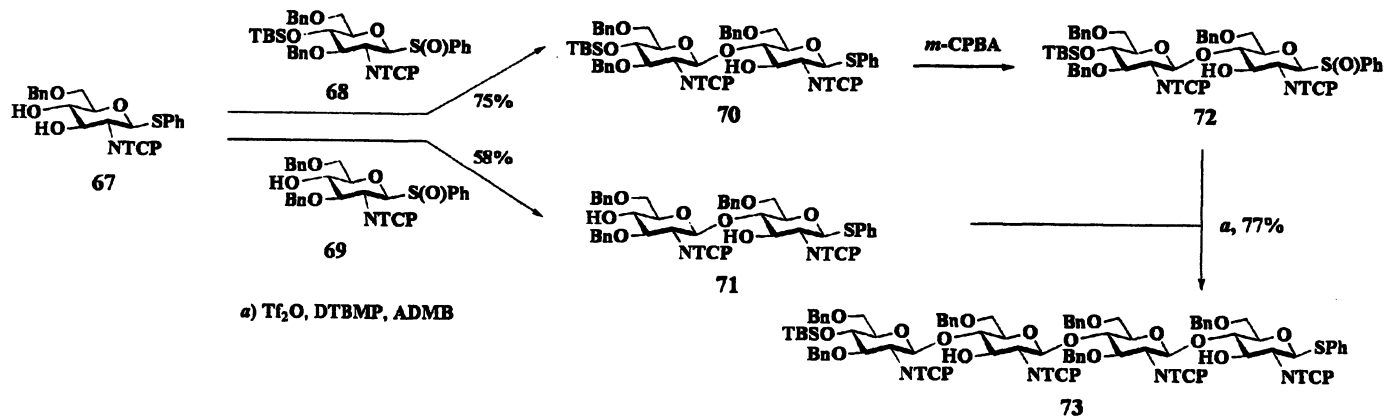


Figure 10. Construction of the backbone of Lipid IV tetrasaccharide 73.

References

1. van Heijenoort, J. *Nat. Prod. Rep.* **2001**, *18*, 503-519.
2. van Heijenoort, J. *Glycobiology* **2001**, *11*, 25R-36R.
3. Kumar, S.; Roychowdhury, A.; Ember, B.; Wang, Q.; Guan, R. J.; Mariuzza, R. A.; Boons, G. J. *J. Biol. Chem.* **2005**, *280*, 37005-37012.
4. Heseck, D.; Lee, M.; Morio, K. I.; Mobashery, S. *J. Org. Chem.* **2004**, *69*, 2137-2146.
5. Heseck, D.; Suvorov, M.; Morio, K.; Lee, M.; Brown, S.; Vakulenko, S. B.; Mobashery, S. *J. Org. Chem.* **2004**, *69*, 778-784.
6. Fuda, C.; Heseck, D.; Lee, M.; Morio, K.; Nowak, T.; Mobashery, S. *J. Am. Chem. Soc.* **2005**, *127*, 2056-2057.
7. Saha, S. L.; Van Nieuwenhze, M. S.; Hornback, W. J.; Aikins, J. A.; Blaszczyk, L. C. *Org. Lett.* **2001**, *3*, 3575-3577.
8. Termin, A.; Schmidt, R. R. *Liebigs Ann. Chem.* **1992**, 527-533.
9. Oikawa, M.; Liu, W. C.; Nakai, Y.; Koshida, S.; Fukase, K.; Kusumoto, S. *Synlett* **1996**, 1179-1180.
10. Fuda, C.; Heseck, D.; Lee, M.; Heilmayer, W.; Novak, R.; Vakulenko, S. B.; Mobashery, S. *J. Biol. Chem.* **2006**, *281*, 10035-10041.
11. Rekharsky, M.; Heseck, D.; Lee, M.; Meroueh, S. O.; Inoue, Y.; Mobashery, S. *J. Am. Chem. Soc.* **2006**, *128*, 7736-7737.
12. Cho, S.; Wang, Q.; Swaminathan, C. P.; Heseck, D.; Lee, M.; Boons, G. J.; Mobashery, S.; Mariuzza, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, in press.
13. Meroueh, S. O.; Bencze, K. Z.; Heseck, D.; Lee, M.; Fisher, J. F.; Stemmler, T. L.; Mobashery, S. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 4404-4409.
14. Inamura, S.; Fukase, K.; Kusumoto, S. *Tetrahedron Lett.* **2001**, *42*, 7613-7616.
15. Inamura, S.; Fujimoto, Y.; Kawasaki, A.; Shiokawa, Z.; Woelk, E.; Heine, H.; Lindner, B.; Inohara, N.; Kusumoto, S.; Fukase, K. *Org. Biomol. Chem.* **2006**, *4*, 232-242.
16. Kubasch, N.; Schmidt, R. R. *Eur. J. Org. Chem.* **2002**, 2710-2726.
17. Swaminathan, C. P.; Brown, P. H.; Roychowdhury, A.; Wang, Q.; Guan, R. J.; Silverman, N.; Goldman, W. E.; Boons, G. J.; Mariuzza, R. A. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 684-689.
18. Sieber, P. *Tetrahedron Lett.* **1987**, *28*, 2107-2110.
19. Chowdhury, A. R.; Boons, G. J. *Tetrahedron Lett.* **2005**, *46*, 1675-1678.
20. VanNieuwenhze, M. S.; Mauldin, S. C.; Zia-Ebrahimi, M.; Winger, B. E.; Hornback, W. J.; Saha, S. L.; Aikins, J. A.; Blaszczyk, L. C. *J. Am. Chem. Soc.* **2002**, *124*, 3656-3660.
21. VanNieuwenhze, M. S.; Mauldin, S. C.; Zia-Ebrahimi, M.; Aikins, J. A.; Blaszczyk, L. C. *J. Am. Chem. Soc.* **2001**, *123*, 6983-6988.

22. Schwartz, B.; Markwalder, J. A.; Wang, Y. *J. Am. Chem. Soc.* **2001**, *123*, 11638-11643.
23. Ye, X. Y.; Lo, M. C.; Brunner, L.; Walker, D.; Kahne, D.; Walker, S. *J. Am. Chem. Soc.* **2001**, *123*, 3155-3156.
24. Lazar, K.; Walker, S. *Curr. Opin. Chem. Biol.* **2002**, *6*, 786-793.
25. Welzel, P. *Chem. Rev.* **2005**, *105*, 4610-4660.
26. Narayan, R. S.; VanNieuwenhze, M. S. *Eur. J. Org. Chem.* **2007**, 1399-1414.
27. Ha, S.; Chang, E.; Lo, M. C.; Men, H.; Park, P.; Ge, M.; Walker, S. *J. Am. Chem. Soc.* **1999**, *121*, 8415-8426.
28. Zhang, Y.; Fechter, E. J.; Wang, T. S. A.; Barrett, D.; Walker, S.; Kahne, D. E. *J. Am. Chem. Soc.* **2007**, *129*, 3080-3081.

Chapter 4

New Glycosyl Thio-carboamino Peptides as New Tools for Glycobiology

Zbigniew J. Witzak

Department of Pharmaceutical Sciences, School of Pharmacy,
Wilkes University, Wilkes-Barre, PA 18766

New developments in glycobiology require determination of the precise functional domain of the carbohydrate ligand and follow up process for the specific binding affinity to glycoproteins through specific glycopeptide moieties. Small glycopeptide derivatives linked to the carbohydrate moiety *via* nonhydrolyzable link developed in our laboratory might serve as potential new tools to study this phenomenon. Our synthetic studies utilizing stereoselective Michael addition of reactive thiols to new conjugated enones and converting them into carbopeptides are presented in this short review.

Introduction

As recognition molecules, carbohydrates play critical roles in many biological functions. Discovery and identification of these critically important bioactive carbohydrates constitutes a unique challenge to determine the precise functional domain of the carbohydrate ligand to design and develop new classes of glycochemicals. This identification, including potent enzyme inhibitors and receptor ligands for peptide, oligonucleotide and small molecules through introduction of carbohydrate molecular diversity, *via* functionalization or

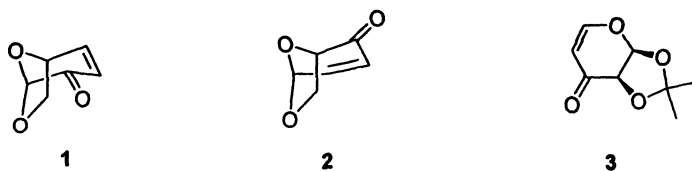
combinatorial chemistry into the drug discovery process, is a timely and widely accepted tool. However, the importance of glycoprotein and glycolipids found at the cell surface is critical in communication between bacteria and many viruses. Carbohydrate functionalized glycopeptides are excellent molecular tools to study these phenomena. Moreover, new synthetic procedures utilizing reactive carbohydrate enones as valuable synthons will open new alternative and efficient routes to accomplish this preparative task.

We have developed several methods for the preparation of *S*-linked thiodisaccharides and mixed C-S-thiodisaccharides linked via (1-4)-; (1-3)-; (1-2)-; and recently (1-5)- positions (1-6).

Our first method uses a stereoselective Michael addition of active thiols to conjugated enones, such as levoglucosenone (1) and its isomeric isolevoglucosenone (2). Both simple and small bicyclic enones are important and efficient chiral starting materials for the synthesis of many analogs of complex natural products (7-24). Our laboratory (7) was the first to synthesize the (+)-enantiomer of levoglucosenone and its 5-hydroxymethyl analog, starting from the known precursor, 5-hydroxymethyl-1,6-anhydro- α -*altro*-hexopyranose.

The high chemical reactivity of the conjugated system in levoglucosenone is an excellent reason to further develop new synthetic approaches for the synthesis of a variety of natural products targets that require stereoselective coupling with the sugar unit.

As our need for larger quantities and variety of conventionally functionalized enones increased, we have been constantly exploring methods that would make them more readily available for exploratory studies and multi-step synthesis. One of such reactive enones originally synthesized by Klemer and Jung (25) and currently explored by us is 4,6-dideoxy-1,2-*O*-isopropylidene-D-glycero-pent-4-enopyranose-3-ulose (3).

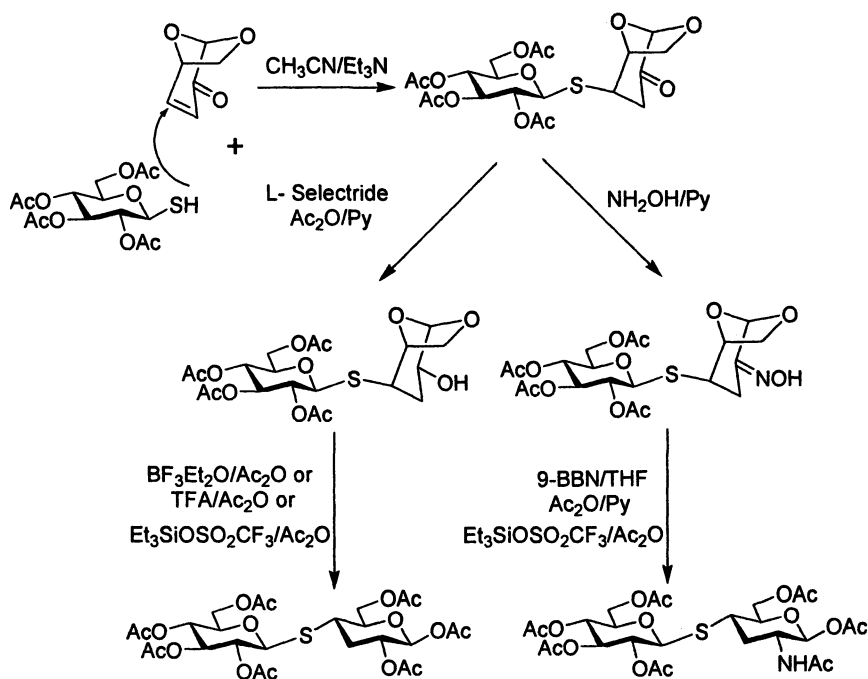


Scheme 1. Levoglucosenone 1, isolevoglucosenone 2, and 4,6-dideoxy-1,2-*O*-isopropylidene-D-glycero-pent-4-enopyranose-3-ulose 3

The enone (3) serves as new convenient building block for a stereoselective functionalization reaction, particularly as a Michael addition reaction acceptor (Scheme 1).

Synthetic Studies

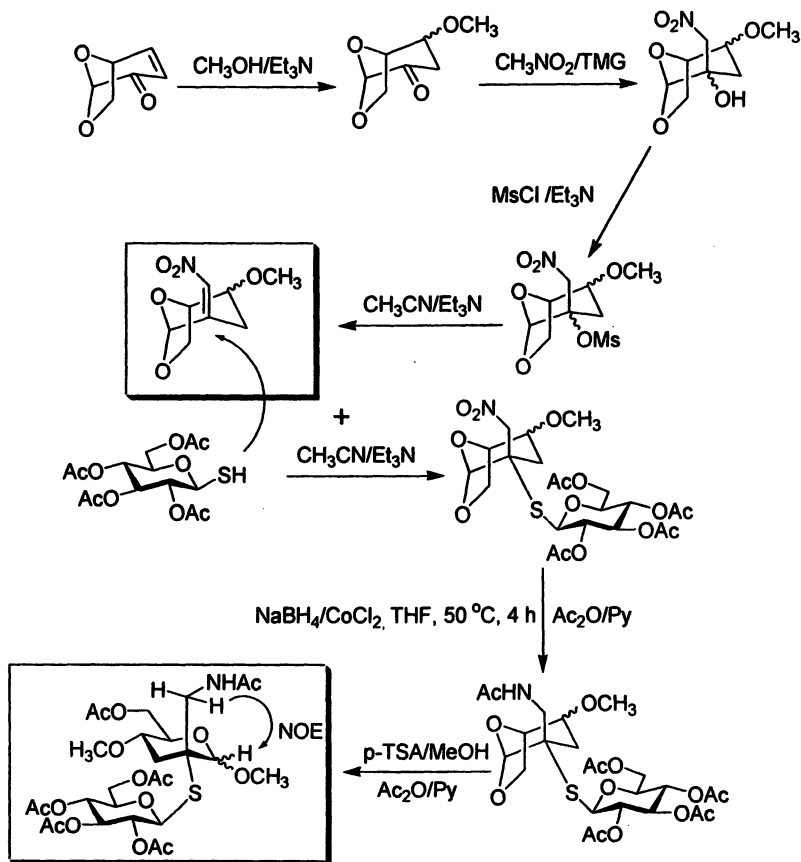
The stereoselective, one-step synthesis of (1,4)-*S*-thiodisaccharides (1-2) is a classical example (Scheme 2) of exploiting the excellent functionality of levoglucosenone as Michael reaction acceptor. The methodology of functionalization of C-2 either *via* conventional and stereoselective reduction, or oximation and then stereoselective reduction of the oximino functionality is unique in providing both classes of *S*-thio-disaccharides. Moreover, the C-2 amino functionalized nonhydrolyzable *S*-thio-disaccharide, after deprotection of C-2 acetamido function represents a convenient molecular scaffold for the generation of diverse libraries of peptidomimetics as a selected class of glycoconjugates.



Scheme 2. Stereoselective Michael addition of 1 thio- β -D-glucose to levoglucosenone with the formation of (1-4)-*S*-thiodisaccharides Ref. 1-2

Our laboratory also developed a new approach to β -(1-2)-2,3-dideoxy-2-*C*-acetamidomethyl-2-*S*-thiodisaccharides (Scheme 3) from new functionalized levoglucosenone chiral synthons at C-2 as α -nitroenone (3). This highly reactive

enone is conveniently produced *via* a three step approach involving Michael addition of methanol followed by nitromethane addition to the keto function at C-2 and mesylation/elimination under strongly basic conditions. The high reactivity of the synthesized enone was exploited in the stereoselective Michael addition of reactive thiols including 1-thio- β -D-glucose as depicted in Scheme 3.



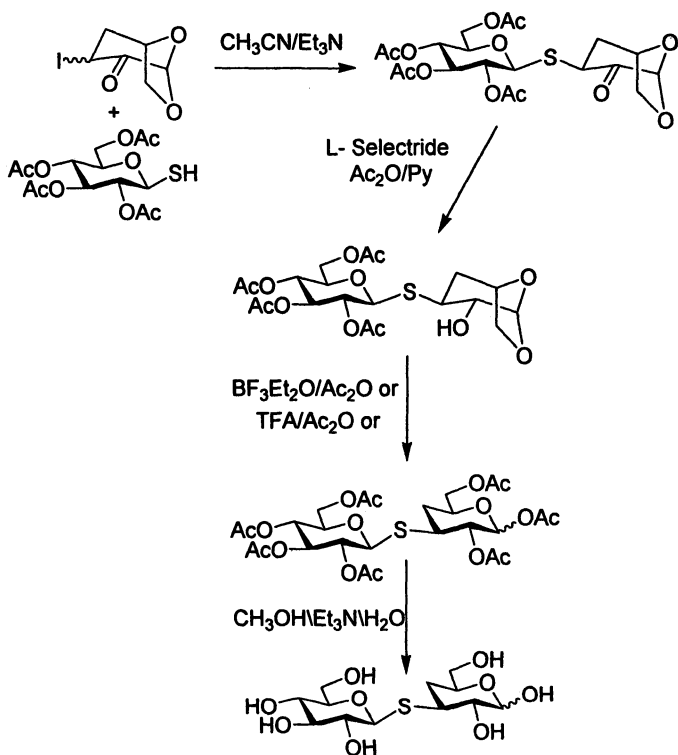
Scheme 3. Stereoselective Michael addition of 1-thio- β -D-glucose to 2-nitromethylene levoglucosene with formation of (1-2)-S-thio-3,4-dideoxy-2-acetamido-disaccharides Ref. 3

The Michael addition of 1-thioglucose to the α -nitroalkene proceeded smoothly with the formation of a new branched-chain thiodisaccharide with geminal nitro group at C-2. Reduction of the nitro group at C-2 of the branched-chain thiodisaccharide was especially challenging, but was achieved efficiently with

sodium borohydride/cobalt chloride complex. For the purpose of purification, acetylation of the reduction product was used to produce an acetamido derivative.

The stereochemistry of the new branched thiodisaccharide was assigned on the basis of NOE results displaying a 5% enhancement between the H-1 and H of acetamido group at C-2'.

Similar functionalization, but at the C-3 position in levoglucosenone, *via* iodine addition to the conjugate position and then nucleophilic displacement of iodine with a reactive thiol represents another valuable stereoselective strategy to α/β (1-3)-linked *S*-thiodisaccharides.



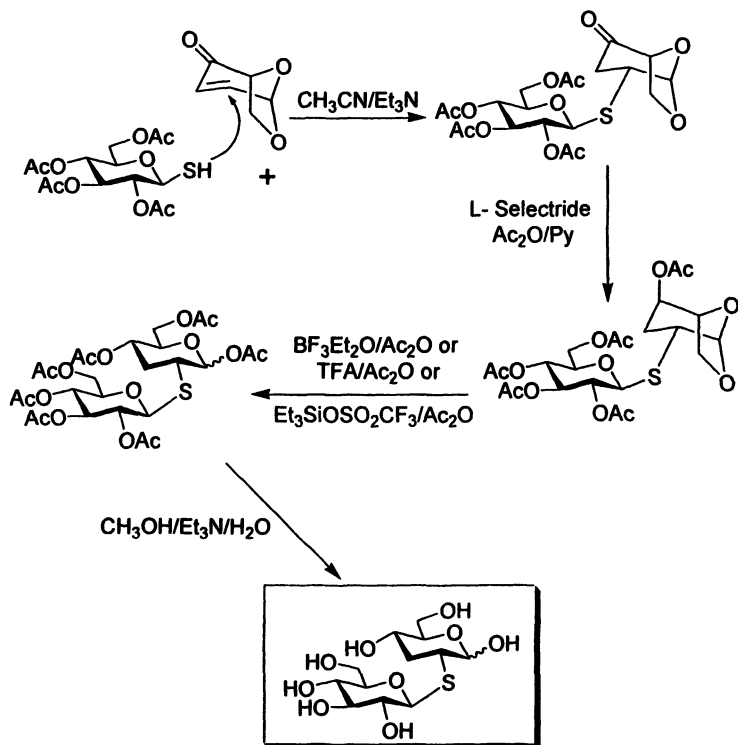
Scheme 4. Stereoselective $\text{S}_{\text{N}}2$ displacement of iodine in 3-iodolevoglucosenone by 1-thio- β -D-glucose with stereoselective formation of (1-3)-*S*-thiodisaccharides. Ref. 4

This approach starts with addition of iodine to levoglucosenone, which has been performed conveniently by the treatment of this enone with a solution of iodine in anhydrous pyridine resulting in the formation of 3-iodolevoglucosenone in moderate (55%) yield as reported by Isobe (15). The reaction of 3-iodoketone with thiols (4) (Scheme 4), proceeded smoothly with the formation of β -(1-3)-

S-thiodisaccharides in 60-72% yield. Proton-proton couplings in the $^1\text{H-NMR}$ spectra of thiodisaccharide confirmed that only the 3-equatorial product was obtained as single product. The *D-allo* stereochemistry with coupling constants of $J_{3a4} = 7.6$ Hz indicated an equatorial orientation of the substituent at C-3.

This stereospecificity has been observed previously in levoglucosenone conjugate addition (1-2) and proceeds by the attack on an incoming nucleophile at the iodine face opposite the 1,6-anhydro ring. The advantage of the stereoselective iodide displacement is the exclusive formation of an *S*-linkage from the less hindered face of the molecule with inversion of configuration at C-3. The shielding effect of the 1,6-anhydro bridge in the iodo precursor effectively prevents formation of the 3-axial product, thus forming only the 3-equatorial product.

Another stereoselective approach to (1-2)-*S*-thiodisaccharides developed in our laboratory exploited an excellent reactivity of isomeric isolevoglucosenone and is depicted in Scheme 5.



Scheme 5. Stereoselective Michael addition of 1 thio-β-D-glucose to isolevoglucosenone with the formation of (1-2)-*S*-thiodisaccharides Ref. 5

This stereoselectivity as observed previously in levoglucosenone conjugate addition proceeds by the attack of an incoming nucleophile (thiol) at the alkene face opposite the 1,6-anhydro ring. The sterically hindered 1,6-anhydro bridge in isolevoglucosenone is, therefore assumed to effectively prevent formation of the opposite stereoisomer.

Conventional reduction of C-4 keto function with L-Selectride, followed with acetylation proceeds with high stereoselectivity forming only single epimer as detected by ^1H NMR. Final deprotection by removal of 1,6-anhydro ring with tandem reagents trifluoroacetic acid/acetic anhydride *via* conventional acetolysis produced an anomeric mixture of crystalline heptaacetates in an α/β ratio of 1:3. This mixture was further deprotected with aqueous methanol with catalytic amount of triethylamine solution to form crystalline inseparable anomeric mixtures of (1-2)-*S*-thio-3'-deoxydisaccharides.

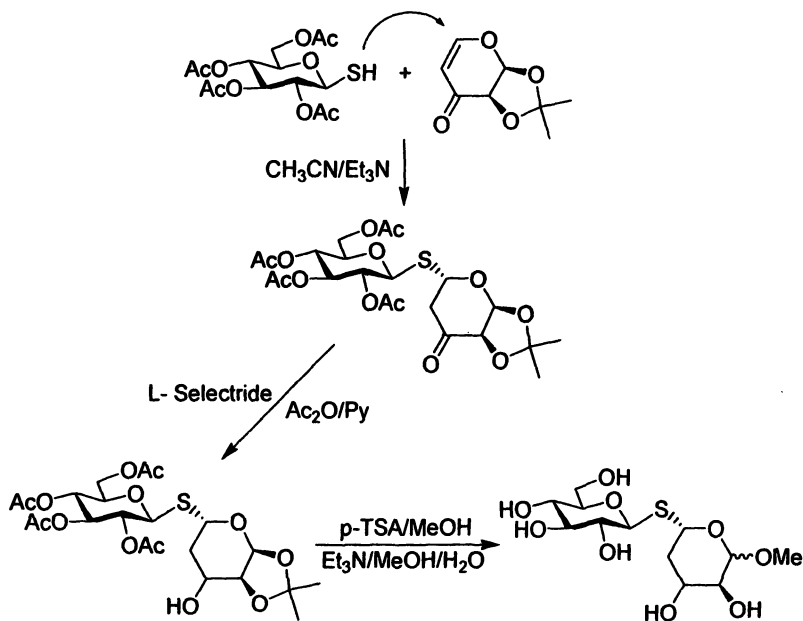
As previously shown in the case of levoglucosenone, the intermediate addition product with a keto functionality at C-4 is an excellent precursor for 4-amino functionalized derivatives, that could be extended *via* a conventional oximation/ reduction sequence. We are currently exploring this attractive route to the new generation of peptidomimetics bearing multiple peptide links at this terminal "tail" C-4 position. Additionally, this highly specific approach to (1-2)-thiolinked *S*-thio-disaccharides offers a unique possibility of further functionalization of the core structure from both ends, at the anomeric carbon at C-1, and at the C-4 of the non-reducing sugar in addition to both primary hydroxyl groups at C-6 and C-6'.

Another category of reactive enones is 4-deoxy-1,2-*O*-isopropylidene-L-glycero-pent-4-enopyrano-3-ulose, originally synthesized by Klemer and Jung (25) and currently explored by us as an extremely useful new chiral building block for stereoselective functionalization reactions, especially as Michael addition acceptors (Scheme 6).

This enone is another important starting material since it gives immediate access to the modified C-5 and C-3 positions. Also, because it can be selectively hydrolyzed at C-1 and C-2 to hydroxyl derivatives, further manipulation leading to the synthesis of various functionalized analogs can be readily achieved.

The Michael addition of 1-thio- β -D-glucose to this enone proceeded smoothly (6) with the formation of β -(1-5)-4-deoxy-5-*C*-thiodisaccharide in 94% yield. The proton-proton coupling constants in the ^1H NMR spectrum of the addition product confirmed that only the 5-axial adduct was obtained as a single product. This stereoselectivity, as observed previously in levoglucosenone conjugate additions proceeds by the attack of incoming nucleophile at the top of the alkene face of the enone ring as depicted in Scheme 6.

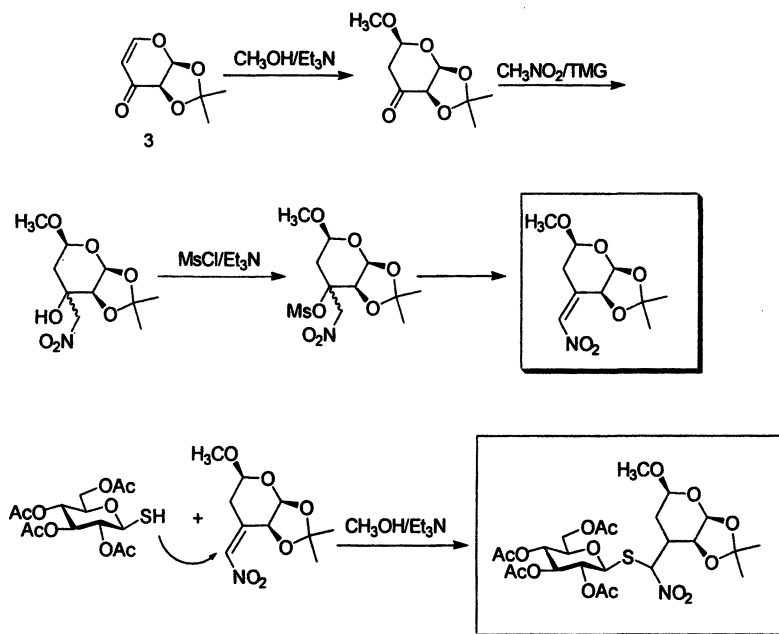
The intermediate ketone addition product offers potential for the synthesis of precursors of certain amino sugars as reported by us earlier (1-5), through conventional oximation and highly stereoselective reduction of the oxime function at C-3 to the corresponding amino group.



Scheme 6. Stereoselective Michael addition of 1-thio- β -D-glucose to arabinose enone **3** with formation of (1-5)-*S*-thiodissaccharides. Ref. 6

The enone **3** was further functionalized *via* tandem reaction by Michael addition of methanol to the conjugate system at C-5, followed by base catalyzed nitromethane addition to the keto function at C-2 and base catalyzed (TMG – tetramethylguanidine) mesylation/elimination with the formation of new α -nitroenone as depicted in Scheme 7. The intermediate α -nitroenone functionalized at C-2 as a conjugate system is indeed an excellent Michael acceptor of reactive nucleophiles, including 1-thio- β -D-glucose. The conjugate addition of this reactive thiol was performed in the same fashion as before (**2**, **3**) with the stereoselective formation of the first representative example of highly functionalized and previously unknown class of *C*-nitro-*S*-thiodisaccharide derivatives.

This new family of *C*-nitro-*S*-thio-addition products after subsequent reduction of the nitro group at C-2 to the corresponding amino group represents an excellent class of precursors for the preparation of a number of other convenient glycoconjugates. Moreover, this methodology is anticipated to be an extremely useful addition to the synthetic arsenal for preparation of bioactive conjugates containing nonhydrolyzable beta-linked thio-glucosyl residues.

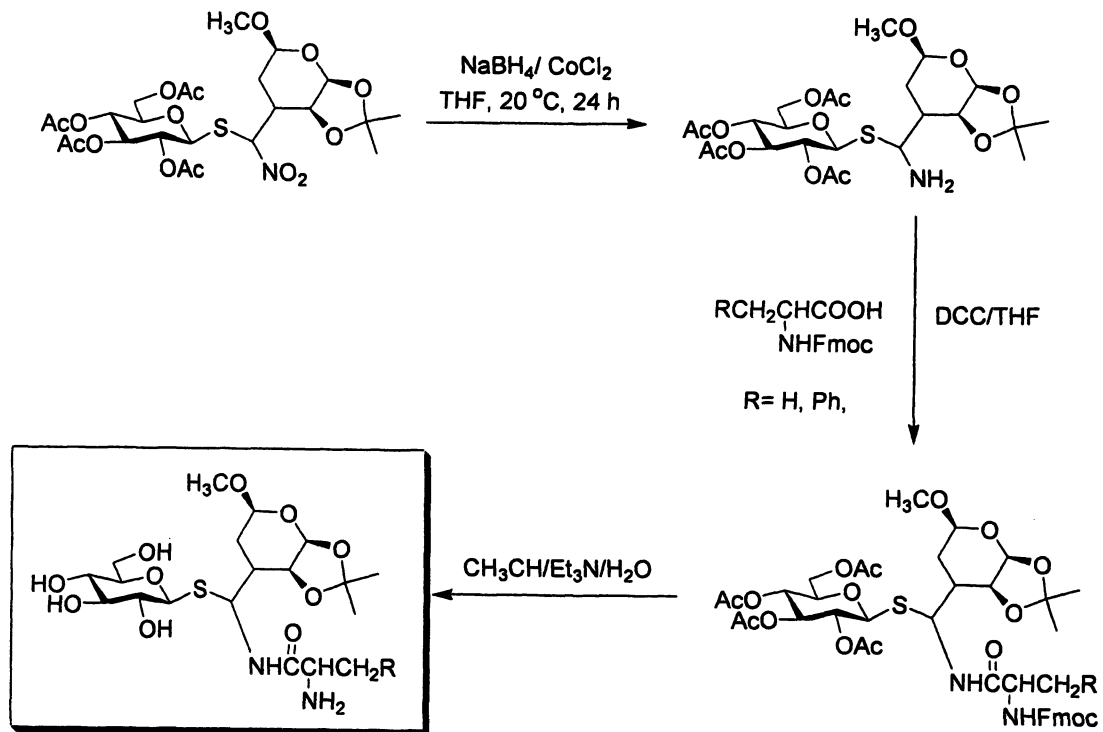


Scheme 7. Functionalization of 4,6-dideoxy-1,2-*O*-isopropylidene-*D*-glycero-pent-4-enopyranose-3-ulose **3** via nitroenone and Michael addition of 1-thio- β -*D*-glucose. Formation of (1-3)-*S*-thio-3'-nitro-3,4-dideoxysaccharides

The amino functionality of the reduction product is explored in the synthesis of the first members of thio-peptides as reported in the Scheme 8.

The reduction of the C- α -nitro function of new nitro-*S*-thiodisaccharide was performed in the same fashion as described by us earlier (3) with sodium borohydride/cobalt chloride complex. For the purpose of purification and isolation, the corresponding amino derivatives were acetylated *in situ* and selectively deprotected to free amino derivatives at C-2 by employing Hanessian methodology (26) using triethyloxonium fluoroborate.

Coupling of the free amino derivative with Fmoc protected commercially available amino acids (alanine and phenylalanine) was accomplished *via* standard reaction conditions activated by DCC in anhydrous tetrahydrofuran (THF) solution. Deprotection of synthesized peptides by treatment with aqueous methanol solution and catalytic amount of trimethylamine produced a new class of *S*-thio-carbo peptides in 86% yield. This particular family of new stable peptidomimetics is conveniently protected and could be used for further additional functionalization at the primary -OH at C-6 of the thioglucose moiety. Further deprotection of the 1,2-*O*-isopropylidene block created another



Scheme 8. Coupling of amino functionalized thiodisaccharide with Fmoc protected amino acids

opportunity for chain extension by conventional glycosylation with biologically important sugars molecules. These strategically important C-1 and C-2 positions can be functionalized even further to generate diverse library of peptidomimetics.

The same strategy of nitro functionalization/reduction and selective deprotection of acetamido functionality by Hanessian protocol (26) using triethyloxonium fluoroborate was applied to our earlier synthesized branched-chain amino sugars as suitable precursors for synthesizing new libraries of nonhydrolyzable *S*-thio-carbopeptido mimetics as new classes of geminal branch-chain functionalized glycoconjugates as depicted in Scheme 9.

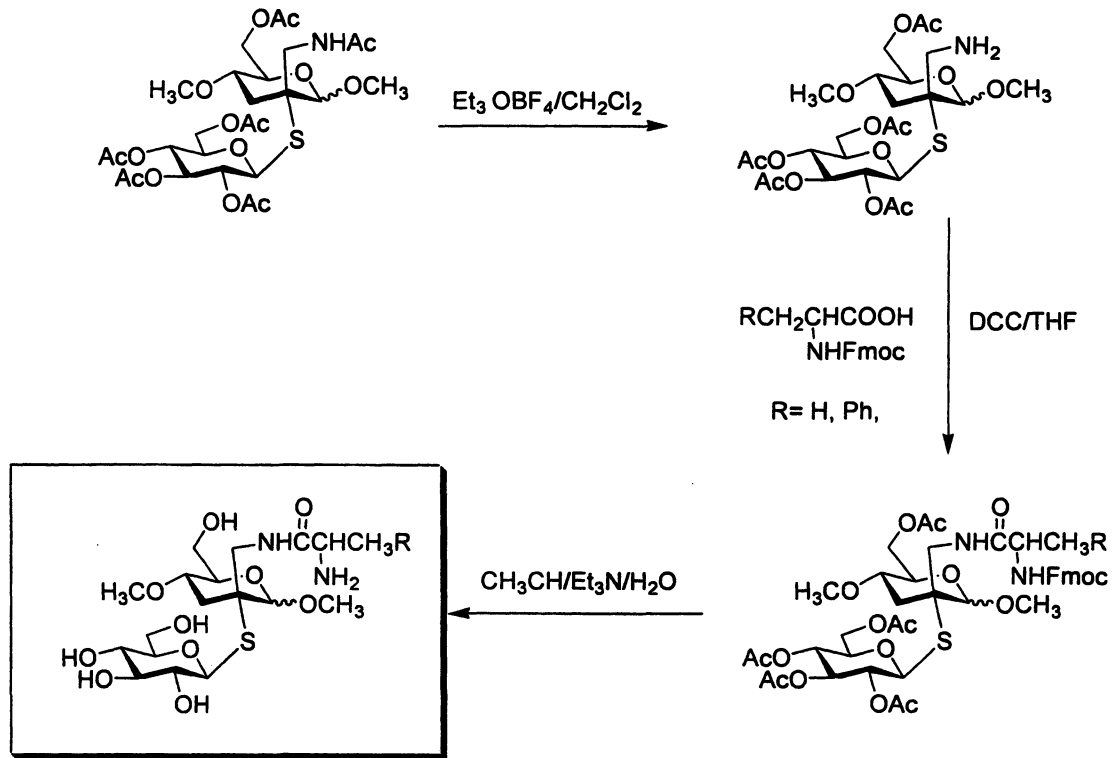
The coupling with Fmoc protected commercially available amino acids (alanine and phenylalanine) was performed in the same fashion as before for arabinose functionalized *S*-thio-carbopeptides. Final deprotection of synthesized peptides by treatment of aqueous methanol solution with catalytic amount of triethylamine produced the first representatives of a new class of stable and nonhydrolyzable *S*-thio-carbo peptides.

Finally, using 4-deoxy-(1-5)-5-*C*-thiodisaccharide (6) previously synthesized in our laboratory as a starting template (*via* conversion into amino derivative) after selective deprotection of 2-acetamido function at -C-2, we synthesized a third family of strategically important peptidomimetics. Again, coupling with Fmoc protected amino acids (alanine and phenylalanine) under the same reaction condition as before produces *S*-thio-peptides in 86% yield. Deprotection of the above target derivatives through standard aqueous methanol/triethylamine solution produces pure and stable derivatives.

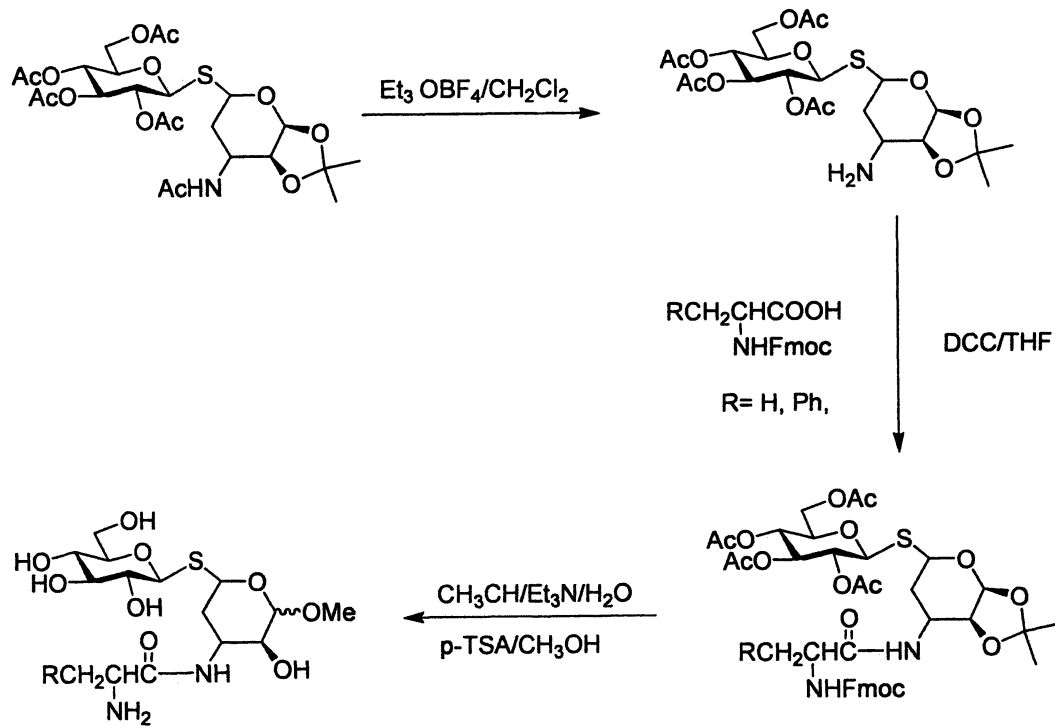
This particular family is composed of three specifically designated carbon/sulfur distance motifs as specific pharmacophores with increased level of lipophilicity, as compared to two previous analogs depicted in Schemes 8 and 9.

The measurement of lipophilicity value for all three families of *S*-thio-carbopeptides is essential for their potential affinity effect to cell membrane and the level of penetration/diffusion through membrane. The detailed values of measured lipophilicity calculated as $\log P$ *via* molecular modeling will be published separately as collaborative work effort. These particular data are essential for preliminary biological screening for all thioglycomimetics for selective *in vitro* cell viability assays as reported by us earlier (27)

The overall strategies of constructing *S*-thio-carbopeptide libraries summarized in Schemes 8-10 are representative of divergent syntheses, and enones or nitroenones used as starting materials are easily available and can be synthesized in a similar fashion. Moreover, these chemical strategies presented here for the purpose of introducing carbohydrate molecular diversity into glycoconjugates will have a significant impact on the preparation of many dynamic libraries of lead candidates having high potential rationales in the drug discovery process. Additionally, these new *S*-thio-carbo peptides are being tested as potential agents to inhibit HIV-induced cell killing and virus production in CEM or MT-2 cells (28).



Scheme 9. Coupling of amino functionalized thiodisaccharide with Fmoc protected amino acids

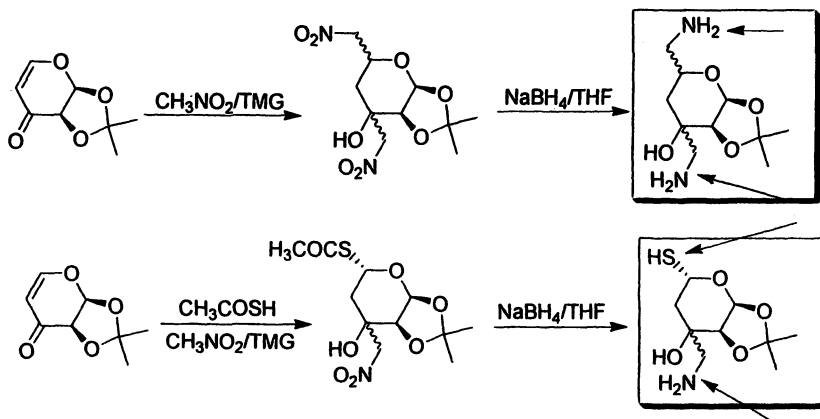


Scheme 10. Coupling of amino functionalized thiodisaccharide with Fmoc protected amino acids

New Perspectives

Although developments in the chemistry of carbohydrate enones levo- and isolevogluconone (during the last ten years) that uses modern reagents as tools in organic synthesis as presented in this short review, will definitely change our perception of their growing potential. The aggressive promotion and utilization of this chemistry must continue, to encourage more extensive study in many different directions.

Moreover, the excellent chiral functionality of arabinose enone and its functionalized new synthons will create additional possibilities of interdisciplinary approaches not only in pure synthetic organic chemistry but also in peptide and combinatorial chemistry. The latter is especially appealing for functionalizing this molecule by creating a number of useful scaffolds.



Scheme 11. Stereoselective functionalization of enone **3** via Michael addition of nitromethane/reduction and thiacetic acid/nitromethane reduction approach.

The most useful scaffolds would have modified functional groups such as -NH_2 , -SH , at C-5 and C-3'. Our laboratory is developing a new family of arabinose-based scaffolds with these functional groups at the above positions (Scheme 11).

Conclusion

Through a number of new developments and synthetic methods devoted to the subject during the last ten years, one can easily conclude that this fascinating

topic is growing and will continue to grow. The variety of methods for the functionalization of these classical building blocks as reactive enones provides a number of attractive stereoselective approaches to various classes of optically active derivatives of particular interest including sulfur and nitrogen peptides as well as rare carbohydrates.

Additionally, the combinatorial utilization of carbohydrate scaffolds based on enones functionalization will also constitute attractive and relatively cheap starting materials.

This rich selection of potential approaches, combined with further developments of new procedures and modern reagents, creates an enormous opportunity for the field of glycobiology to be at the frontier for many years to come.

References

1. Witczak, Z. J.; Sun, J.; Mielguj, R. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 2169.
2. Witczak, Z. J.; Chhabra, R.; Chen, H.; Xie, Q. *Carbohydr. Res.* **1997**, *301*, 167.
3. Witczak, Z. J.; Chhabra, R.; Boryczewski, D. *J. Carbohydr. Chem.* **2000**, *19*, 543.
4. Witczak, Z. J.; Kaplon, P. Kolodziej, M. *Monatshefte fur Chemie*, **2002**, *133*, 521.
5. Witczak, Z. J.; Chen, H.; Kaplon, P. *Tetrahedron: Asymmetry*, **2000**, *11*, 519.
6. Witczak, Z.J. Lorchak, D.; Nguyen, N. *Carbohydr. Res.* **2007**, *342*, 1929.
7. Witczak, Z. J.; Mielguj, R. *Synlett* **1996**, 108.
8. For reviews see; *Levoglucosenone and Levoglucosans Chemistry and Applications* Witczak, Z. J. Ed. ATL Press Science Publishers; Mt. Prospect, IL **1994**; Witczak, Z. J. in *Studies in Natural Products Chemistry*, Atta-Ur-Rahman, Ed. Vol. 14, Elsevier Science Publishers, Amsterdam, **1993**, pp. 267-282; Miftakhov, M. S.; Valeev, F. A.; Gaisina, I. N. *Uspekhi Khimi*, **1994**, *63*, 922; B. Becker, *J. Carb. Chem.* **2000**, *19*, 253. .
9. Witczak, Z. J. *Pure Appl. Chem.* **1995**, *66*, 2189.
10. Blake, A. J.; Cook, T. A.; Forsyth, A. C.; Gould, R. O.; Paton, R. M. *Tetrahedron* **1992**, *48*, 8053.
11. Blake, A. J.; Gould, R. O.; Paton, R. M.; Taylor, P. G. *J. Chem. Res. Synopses* **1993**, 289.
12. Isobe, M.; Fukami, N.; Nishikawa, T.; Goto, T. *Heterocycles* **1987**, *25*, 521.
13. Ward, D. D.; Shafizadeh, F. *Carbohydr. Res.* **1981**, *93*, 287.
14. Bamba, M.; Nishikawa, T.; Isobe, M. *Tetrahedron Lett.* **1996**, *37*, 8199.
15. Matsumoto, K.; Ebata, T.; Koseki, K.; Okano, K.; Kawakami, H.; Matsushita, H. *Carbohydr. Res.* **1993**, *246*, 345.
16. Ebata, T.; Matsumoto, K.; Yoshikoshi, H.; Koseki, K.; Kawakami, H.; Mashushita, H. *Heterocycles*, **1990**, *31*, 423.

17. Mori, M.; Chuman, T.; Kato, K. *Carbohydr. Res.* **1984**, *129*, 73.
18. Witczak, Z. J.; Li, Y. *Tetrahedron Lett.* **1995**, *36*, 2595.
19. Taniguchi, T.; Nakamura, K.; Ogasawara, K. *Synlett* **1996**, 971.
20. Blattner, R.; Page, D. M. *J. Carbohydr. Chem.* **1994**, *13*, 27.
21. Witczak, Z. J.; Chabra, R.; Chojnacki, J. *Tetrahedron Lett.* **1997**, *38*, 2215.
22. Nishikawa, T.; Araki, H.; Isobe, M. *Biosci. Biotechnol. Biochem.* **1998**, *62*, 190.
23. Takeuchi, M.; Taniguchi, T.; Ogasawara, K. *Synthesis* **1999**, 341.
24. Gomez, M.; Quincoces, J.; Peske, K.; Michalik, M. *J. Carbohydr. Chem.* **1999**, *18*, 851.
25. Klemer, A. Jung G. *Chem Ber.* **1981**, *114*, 1192
26. Hanesian S. *Tetrahedron Lett.* **1967**, *8*, 1549.
27. Witczak Z.J.; Kaplon, P.; Dey, P.M. *Carbohydr. Res.* **2003**, *338*, 11.
28. R. N. Comber, J.D. Friedrich, J. A. Secrist III, *J. Med. Chem.*, **1992**, *35*, 3567.

Chapter 5

Chemoenzymatic Synthesis of Sialosides and Their Applications

Hai Yu, Harshal A. Chokhawala, Shengshu Huang, and Xi Chen*

Department of Chemistry, University of California at Davis,
One Shields Avenue, Davis, CA 95616

Recent progress in the development of efficient one-pot three-enzyme chemoenzymatic synthesis and application of functionalized sialosides is described. By taking the advantage of relaxed substrate specificity of several bacterial sialoside biosynthetic enzymes, the method has been used for the preparative scale synthesis of α 2,3- and α 2,6-linked sialoside libraries containing naturally and non-naturally occurring sialic acid modifications. Starting from the hexose precursors (ManNAc or mannose) of sialic acids, a library of *p*NP-tagged sialyl disaccharides with various naturally occurring sialic acid forms, different sialyl linkages, and different penultimate monosaccharides have also been prepared and used in the substrate specificity studies of bacterial sialidases in a 96-well plate-based colorimetric high-throughput screening platform. The combination of efficient chemoenzymatic synthesis and high-throughput screening is a powerful approach to studying proteins that recognizing sialic acid-containing carbohydrates.

Introduction

Sialic acids are a family of nonulosonic acids that have been predominantly found as the outermost carbohydrate units on glycoproteins and glycolipids of vertebrates, or as components of polysaccharides in certain types of bacteria (1, 2). Sialic acids play vital roles in a variety of physiological and pathological processes in vertebrates, such as cellular recognition and communication (1, 2). They are also believed to be important virulence factors in bacteria, used by bacteria to mimic sialylated host cell surface carbohydrate structures to evade detection and attacking by the immune defense mechanisms of the host (3-6).

Sialic acids exhibit tremendous structural diversity in nature, and more than 50 structurally distinct sialic acid forms have been observed. Three basic sialic acid forms are: *N*-acetylneuraminic acid (Neu5Ac), *N*-glycolylneuraminic acid (Neu5Gc), and deaminoneuraminic acid (or keto-deoxynonulosonic acid, KDN). Single and multiple modifications, including *O*-acetylation and less frequent *O*-lactylation, *O*-methylation, *O*-sulfation, and *O*-phosphorylation, can take place at the hydroxyl groups on C-4, C-5, C-7, C-8, and/or C-9 positions of these three basic forms to generate diverse natural occurring sialic acid forms (1, 2).

Cell surface presentation of modified sialic acids is species- and tissue-specific, developmentally regulated, and is believed to be closely related to their biological functions, such as immunogenicity, inflammation, bacterial or viral infection, tumor growth, and metastasis. For example, studies showed that 9-*O*-acetylation of sialic acid can enhance the activation of the alternate pathway of complement (7-9), it is necessary for influenza C virus binding and subsequent invasion on host cell surface (10, 11), but prevents the attachment of malaria parasites (12) and influenza A and B viruses (13, 14). In another example, mouse hepatitis virus strain S is specific to 4-*O*-acetylated Neu5Ac (2, 15). 4-*O*-Acetylation of sialic acid has also been detected in human colon carcinomas (16). The loss of *O*-acetylation of sialyl Lewis X in human colon cancer facilitates metastasis (2). Modifications on sialic acids also affect the cleavage of sialic acid residues by sialidases or trans-sialidases, and often lead to the reduction or even resistance of cleavage by these enzymes (17, 18). The subtle structural modifications on the sialic acid residue may be a way of fine-tuning many biological processes mediated by these sialoglycoconjugates. Nevertheless, a clear understanding on how these structural modifications affect their biological significance is missing.

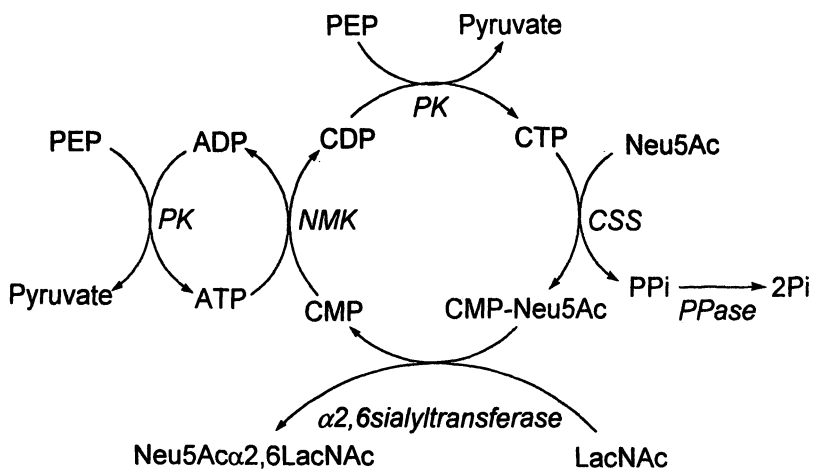
Studies on the mechanism and the significance of nature's sialic acid structural diversity have been limited due to the inaccessibility of homogenous sialosides and sialoglycoconjugates, especially those containing naturally occurring sialic acid modifications. These structures are extremely difficult to purify from natural source in homogenous forms. Despite recent promising developments in chemical synthesis of carbohydrates, selective formation of α -linked sialosides by chemical sialylation remains one of the most challenging

glycosylation reactions due to the hindered tertiary anomeric center and the lack of a neighboring participating group in sialic acids (19, 20). Protecting and deprotecting operations of the nine-carbon monosaccharides are also more complex than six-carbon monosaccharides. Therefore, current chemical synthesis of sialosides is still time-consuming and requires skillful expertise. On the other hand, sialyltransferase-catalyzed enzymatic sialylation offers great advantages. The intrinsic high regioselectivity and stereoselectivity, together with mild reaction condition (normally enzymatic reactions are performed at room temperature or 37 °C in aqueous solution with pH ranges from 6.0 to 8.5) of sialyltransferase-catalyzed reaction, make sialyltransferases very attractive biocatalysts for practical synthesis of sialosides. Earlier practice on sialyltransferase-catalyzed reactions, however, suffered from the low expression level and the narrow substrate specificity of many sialyltransferases, especially those from mammalian sources (21). Difficulties in obtaining expensive, unstable, and not readily accessible sugar nucleotides (CMP-sialic acid and its derivatives) which are donors for the sialyltransferase-catalyzed reaction also limit the scope of the sialosides that have been synthesized by earlier enzymatic approaches.

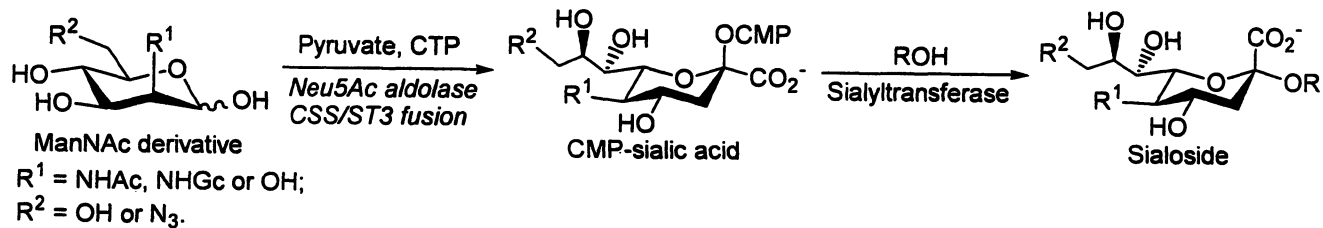
One-Pot Three-Enzyme Approach

Chi-Huey Wong's pioneer work on sialyltransferase-catalyzed enzymatic synthesis of sialyl-*N*-acetylglucosamine (Neu5Ac α 2,6LacNAc) with *in situ* regeneration of CMP-Neu5Ac is a landmark of efficient enzymatic synthesis of sialosides. This method avoided the stoichiometric use of expensive CMP-Neu5Ac donor and the product inhibition of sialyltransferase by CMP (22). In this system (Scheme 1), Neu5Ac was activated by an *E. coli* CMP-Neu5Ac synthetase and transferred by an α 2,6-sialyltransferase to produce sialylated product. The byproduct CMP of the α 2,6-sialyltransferase-catalyzed reaction was recycled to the sialyltransferase donor CMP-Neu5A by the function of two enzymes, including a nucleoside monophosphate kinase (NMK) and a pyruvate kinase (PK). A pyrophosphatase (PPase) was used to degrade the pyrophosphate produced. Five enzymes are involved in the process. The efficient of the synthesis relies on the activities of these enzymes. This method has mainly been applied for the synthesis of Neu5Ac-containing molecules.

A simplified one-pot two-step enzymatic approach was used by James Paulson's group to produce sialoside derivatives (23). Instead of recycling the CMP-sialic acid, the sialyltransferase donors with modifications at C-5 or C-9 position of Neu5Ac were enzymatically synthesized from ManNAc or its C-2 or C-6 modified derivatives, pyruvate, and CTP using a sialic acid aldolase and a CMP-Neu5Ac synthetase/sialyltransferase fusion protein. After removing the protein by membrane filtration, the filtrate containing produced CMP-sialic acid



Scheme 1. Synthesis of $\alpha 2,6$ -linked sialyl-N-acetyllactosamine using a one-pot multi-enzyme system with in situ regeneration of CMP-Neu5Ac. Abbreviations for enzymes: CSS, CMP-sialic acid synthetase; NMK, nucleoside monophosphate kinase; PK, pyruvate kinase; PPase, pyrophosphatase. Abbreviations for compounds: PEP, phosphoenolpyruvate; ADP, adenosine 5'-diphosphate; ATP, adenosine 5'-triphosphate; CMP, cytidine 5'-monophosphate; CDP, cytidine 5'-diphosphate; CTP, cytidine 5'-triphosphate; LacNAc, N-acetyllactosamine; Neu5Ac, N-acetylneuraminic acid; PPI, inorganic pyrophosphate.



Scheme 2. One-pot two-step enzymatic synthesis of sialosides.

or its derivative was mixed with a galactose-terminated oligosaccharide and a sialyltransferase selected from an *N. meningitidis* α 2,3-sialyltransferase, porcine ST3GalI, rat ST3Gal III, human ST6Gal I, and chicken ST6GalNAc I to produce a list of desired α 2,3- and α 2,6-linked sialosides (Scheme 2). Since the sialyltransferases used have restricted donor and acceptor specificity, different sialyltransferases had to be used to obtain different sialosides.

In the examples shown above, the efficiency of chemoenzymatic synthesis of sialoside derivatives was greatly hindered by the limited availability and the narrow substrate specificity of mammalian sialyltransferases. Recently, our group has established a convenient and efficient one-pot three-enzyme system (Scheme 3) for systematic chemoenzymatic synthesis of α 2,3- and α 2,6- linked sialoside libraries containing naturally occurring and non-natural sialic acid modifications (24, 25). In this approach, sialic acid modifications are chemically or enzymatically introduced at early stage, onto ManNAc or mannose (sialic acid precursors) (A in Scheme 3) to obtain B as precursors for the corresponding naturally existing or non-natural modified sialic acid forms. Enzymatic conversion of ManNAc/mannose derivatives B to sialic acid forms C is achieved by a sialic acid aldolase. The produced sialic acids are activated by a CMP-sialic acid synthetase to form CMP-sialic acids D by a CMP-sialic acid synthetase (CSS), and then transferred to a galactose- or GalNAc-terminated glycoside by a sialyltransferase (SiaT) to form structurally defined sialosides with naturally occurring and non-natural sialic acid forms (E in Scheme 3). Because the reaction conditions for these enzyme-catalyzed processes are similar (aqueous solution, neutral to weak basic condition, room temperature or 37 °C), the conversions catalyzed by three enzymatic can be performed in one pot without the isolation of intermediates. This approach thus simplifies the product purification process. Also, the sialic acid derivatives produced by reversible aldolase-catalyzed reaction can be immediately used the CSS that catalyzes the irreversible formation of CMP-sialic acids, which drives the reaction equilibrium of the aldolase reaction towards the formation of the desired sialic acid derivatives. This can avoid the addition of a large excess amount (5-10 equivalents) of pyruvate in a typical sialic acid aldolase catalyzed formation of sialic acids. In addition, this approach avoids the purification of relatively unstable CMP-sialic acid intermediates. Only the final sialoside product needs to be purified, usually by a BioGel P-2 gel filtration chromatography upon the completion of the enzymatic reactions. This approach, thus, simplifies the synthetic scheme and increases the efficiency of sialoside synthesis.

From Scheme 3, one can tell that the key to the success of this efficient one-pot three-enzyme chemoenzymatic approach is to identify and obtain active individual sialoside biosynthetic enzymes which have relaxed substrate specificity and can be expressed in simple expression system as active forms with a high expression level.

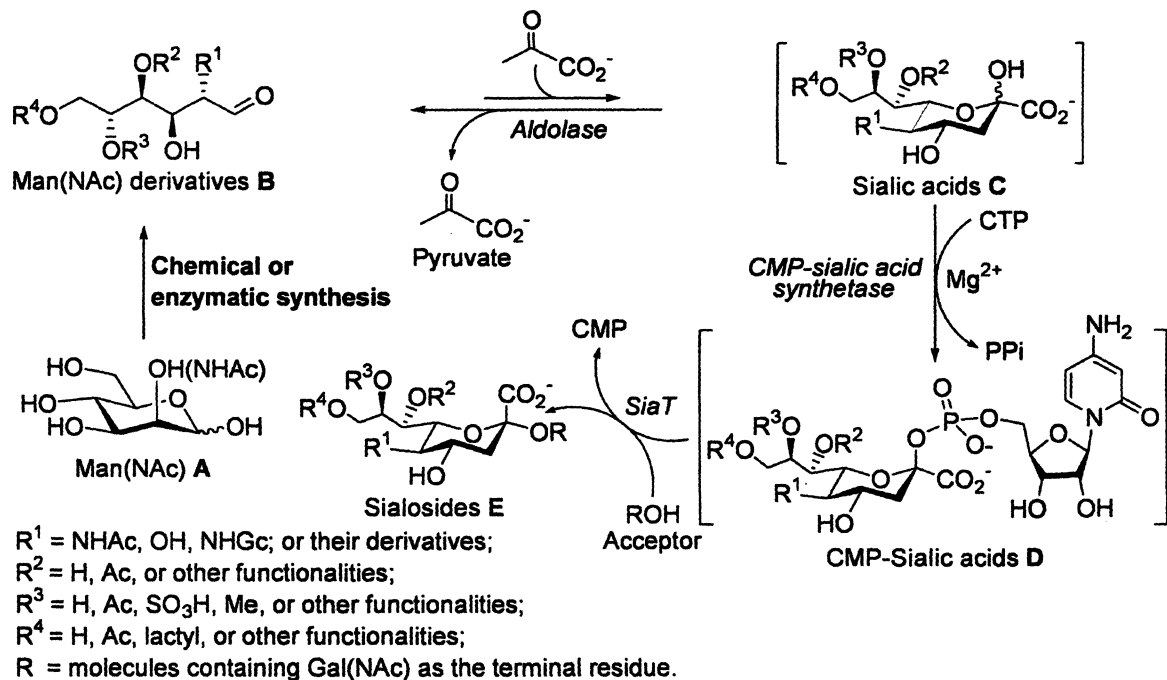
Cloning and characterization of a well reported *Escherichia coli* K-12 sialic acid aldolase indicated that this enzyme can tolerate a diverse modification of the substrates (26-28). Cloning, expression, and substrate specificity studies of three recombinant CMP-sialic acid synthetases cloned from *Neisseria meningitidis* (NmCSS), *Streptococcus agalactiae* serotype V (SaV CSS), and *Escherichia coli* K-1 (*E. coli* CSS) revealed that the NmCSS has the highest expression level, the highest solubility, the highest activity, and the most relaxed substrate specificity among these three enzymes (26). Various CMP-sialic acid derivatives with different modifications on the sialic acid residue have been successfully synthesized in preparative scales (50-200 mg) using a one-pot two-enzyme system containing the *E. coli* sialic acid aldolase and the NmCSS (26). By choosing an appropriate sialyltransferase, either α 2,6- or α 2,3- linked sialosides can be obtained in the one-pot three-enzyme system containing the *E. coli* sialic acid aldolase, the NmCSS, and the sialyltransferase.

Preparative Synthesis of α 2,6-Linked Sialosides Containing Naturally Occurring and Non-natural Sialic Acids

In order to prepare α 2,6-linked sialoside libraries with naturally occurring and non-natural sialic acid modifications, a flexible α 2,6-sialyltransferase enzyme is required. Although several mammalian sialyltransferases have been reported to have relatively broad donor substrate specificity (for example, rat liver α 2,6-sialyltransferases can tolerate a variety of modifications on the Neu5Ac moiety of CMP-Neu5Ac) (29-32), but they suffer from low expression level which limits their applications in preparative- and large- scale synthesis of sialosides.

Photobacterium damsela α 2,6-sialyltransferase (Pd2,6ST) was the first bacterial sialyltransferase which has been cloned and purified by the Yamamoto group (33, 34). This enzyme has a relaxed acceptor specificity (35, 36). For example, it has been applied for the enzymatic sialylation of Tn glycopeptides with GalNAc α -linked to either serine or threonine residue) (37). It was also shown to be able to transfer sialic acid to both *N*- and *O*-linked glycoproteins (38).

Our group has recently cloned a truncated Pd2,6ST containing 17-497 amino acid residues as N-hexohistidine tagged protein and explored its application in the one-pot three-enzyme system for preparative synthesis of functionalized α 2,6-sialosides (25). The tolerance of donor substrate modification by the purified Pd2,6ST was tested using the one-pot three-enzyme system, in which CMP-sialic acid derivatives were generated *in situ* from sialic acid precursors by the aldolase and NmCSS. An extremely relaxed donor substrate specificity was observed for Pd2,6ST. The preparative-scale reactions were then carried out at



Scheme 3. One-pot three-enzyme chemoenzymatic synthesis of sialosides containing natural and non-natural functionalities.

37 °C in Tris-HCl buffer (100 mM, pH 8.5 for substrates without base sensitive *O*-acetyl groups) containing 1.2 equiv. of ManNAc (or mannose and their derivatives) as sialic acid precursor, 1.0 equiv. of galactose (or GalNAc) terminal-containing oligosaccharide as a sialyltransferase acceptor, Mg²⁺ (20 mM), 5 equiv. of pyruvate, 1.2 equiv. of CTP, and appropriate amounts of aldolase, NmCSS, and Pd2,6ST. Tris-HCl buffer (100 mM) with a lower pH value (pH 7.5) was used for preparing sialosides containing base-labile *O*-acetylated or *O*-lactylated sialic acid residues to avoid de-*O*-acetylation. Reactions were monitored by thin-layer chromatography (TLC) analysis (EtOAc:MeOH:H₂O:HOAc = 5:2:1:0.1, by volume) and stained with *p*-anisaldehyde sugar stain. The final sialoside products were purified by a Bio-Gel P-2 gel filtration chromatography and the structures of all sialylated products were characterized by ¹H and ¹³C NMR as well as high resolution mass spectrometry (HRMS).

As shown in Table 1, Pd2,6ST showed very flexible donor substrate specificity and was able to accept a diverse array of modifications on CMP-activated sialic acid. Using 3-azidopropyl β-D-galactopyranose-(1→4)-β-D-glucopyranoside (LacβProN₃, 15) as an acceptor, various naturally occurring α2,6-linked sialosides containing Neu5Ac or its C-5, C-9, C-5/C-9 substituted analogs 16-27 were synthesized from their corresponding ManNAc or mannose analogs in one pot. The yields for sialosides with or without *O*-acylated substitutions were 75-99%. Two non-natural sialosides containing a 4,6-bis-*epi*-KDO 28 and an *N*-(benzyloxycarboxyamido) glycinylamido-neuraminic acid (NeuGlyCbz, 29) were also synthesized in high yields, 92% and 99%, respectively (Table 1). The successful synthesis of these two sialosides demonstrate that Pd2,6ST can transfer sialic acid residues with carbon backbones shorter than nine or with a bulky substitution at C-5 to galactoside with high efficiency. These examples further demonstrated the extremely flexible donor substrate specificity of all three enzymes in the system.

The azido group at the reducing end of functionalized sialosides with sialic acids α2,6-linked to LacβProN₃ can be easily converted, by catalytic hydrogenation, to a primary amino group which can be activated with coupling reagents such as succinimide esters (39, 40), squaric acid diesters (41, 42), maleimide (43, 44), and bis(*p*-nitrophenyl) esters (45-47) and linked to the amino group in proteins or other molecules. Alternatively, the azide group itself can be directly used for efficient conjugation with any biomolecule containing terminal alkyne functional group by "Click Chemistry" (48, 49) or molecules containing triphenylphosphine group by "Staudinger Ligation" (50, 51) (Scheme 4). Such biomolecules containing modified sialic acids serve as important probes to study protein-carbohydrate interactions in a multivalent setting and in producing sialic acid specific antibodies which in turn serve as histochemical tools for detecting organ and tissue specific sialosides.

Several non-natural α 2,6-linked sialosides 36-40 with azide or alkyne-modified sialic acid residues were also prepared in excellent yields (86%-93%) from their C2- or C6- modified ManNAc or mannose bearing corresponding azide or alkyne functional groups 30-34 using the one-pot three-enzyme approach and Gal β OMe (35) as an acceptor for Pd2,6ST (Scheme 5).

Due to its high efficiency, the one-pot three-enzyme approach described above should also be suitable for direct transferring modified sialic acid residues to glycoconjugates containing a terminal Gal or GalNAc residue.

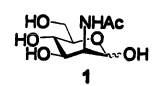
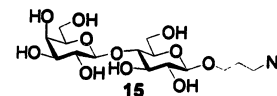
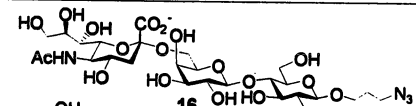
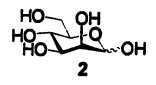
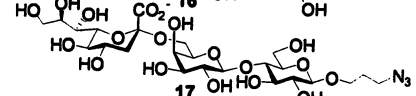
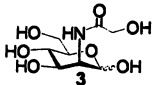
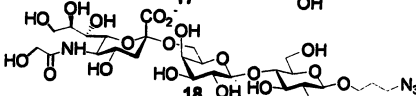
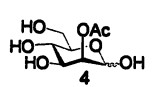
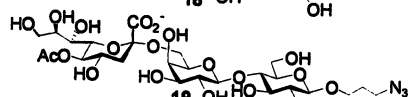
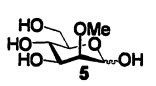
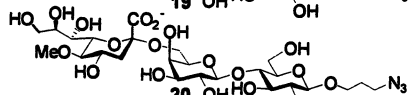
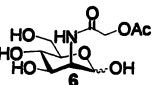
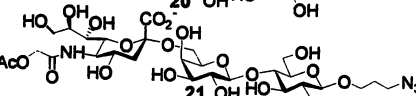
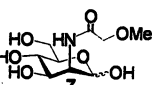
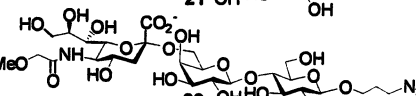
Preparative Synthesis of α 2,3-Linked Sialosides Containing Naturally Occurring and Non-natural Sialic Acids

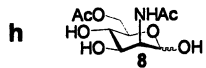
A number of α 2,3-sialyltransferases have been cloned from bacterial and mammalian sources and their substrate specificities have been studied. Among mammalian sialyltransferases, recombinant α 2,3-sialyltransferases from rat liver are the most widely studied ones. They are commercially available (Calbiochem Catalog #566227 and #566218) and have been used extensively in the synthesis of functionalized sialosides. Mammalian sialyltransferases normally have broad donor substrate specificity and are able to tolerate a variety of modifications at C-5 or C-9 of the sialic acid residue in CMP-sialic acid (30, 31). C5-Substituted Neu5Ac analogs (30) such as 5-*N*-formyl-Neu, 5-*N*-trifluoroacetyl-Neu, 5-*N*-benzyloxycarbonyl-Neu, 5-*N*-aminoacetyl-Neu, and 9-substituted Neu5Ac analogues (31) such as 9-acetamido-Neu5Ac, 9-benzamido-Neu5Ac, 9-hexanoylamido-Neu5Ac, and 9-azido-Neu5Ac, can all be transferred relatively easily by an α 2,3-sialyltransferase purified from rat liver to glycans on glycoproteins. CMP-activated 9-amino-Neu5Ac, however, is a weak substrate for this enzyme.

A recombinant α 2,3-sialyltransferase from *Neisseria gonorrhoeae* exhibits broad acceptor substrate specificity. It can use various oligosaccharides, glycoconjugates, and their sulfated derivatives as acceptors (52). The enzyme, however, has quite limited donor specificity and is quite specific for CMP-Neu5Ac. CMP-activated Neu5Gc, Neu5Boc, and KDN are very weak donor substrates for the enzyme. CMP-activated Neu5N₃ and Neu5Ac α 2,9Neu5Ac are not substrates.

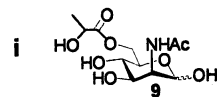
Characterization of a bifunctional bacterial sialyltransferase, Cst-II cloned from *C. jejuni* having both α 2,3- and α 2,8-sialyltransferase activities, has been reported. This sialyltransferase has been used for the synthesis of GD3 oligosaccharides Neu5Ac α 2,8Neu5Ac α 2,3LacOR and other gangliosides (53, 54).

Table 1. One-Pot Three-Enzyme Chemoenzymatic Synthesis of α 2,6-Linked Sialosides Containing Naturally Occurring and Non-natural Sialic Acid Residues Using Lac β ProN₃ as A Sialyltransferase Acceptor.

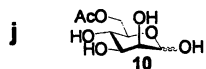
Entry	Donor Precursors	Acceptor	Product	Yield (%)
a				98
b		15		97
c		15		95
d		15		75
e		15		76
f		15		87
g		15		99



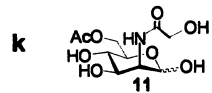
15



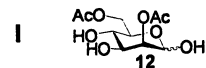
15



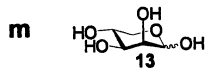
15



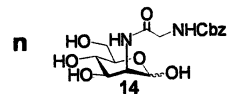
15



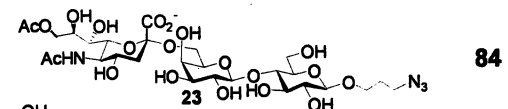
15



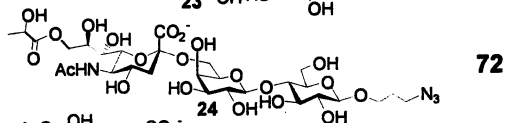
15



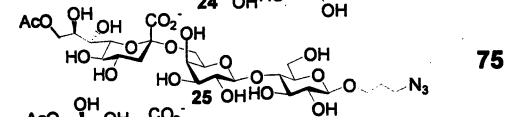
15



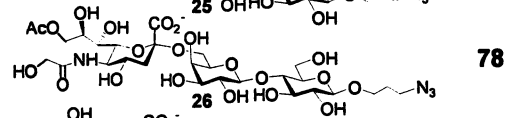
84



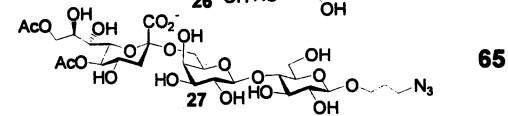
72



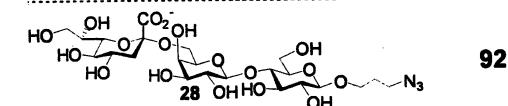
75



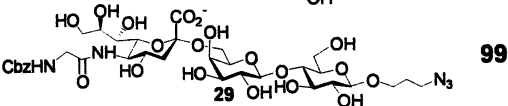
78



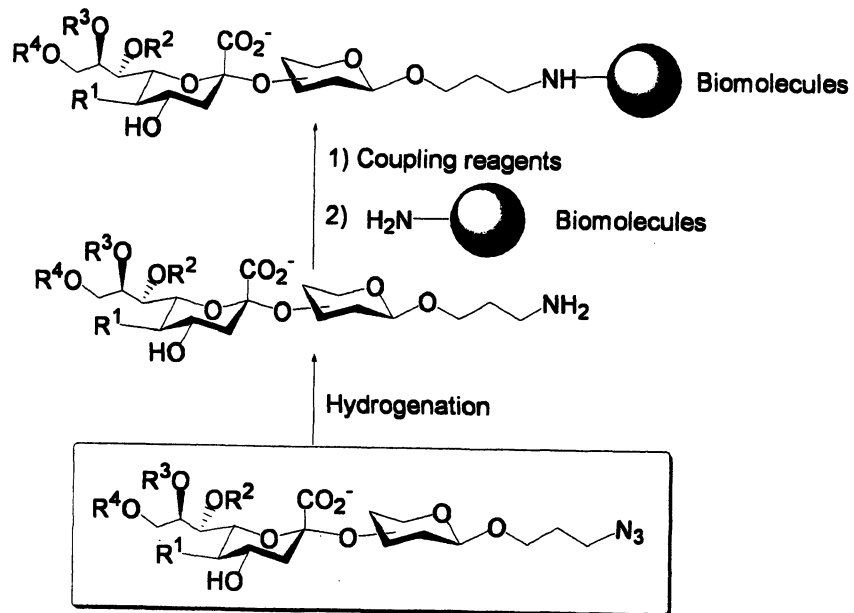
65

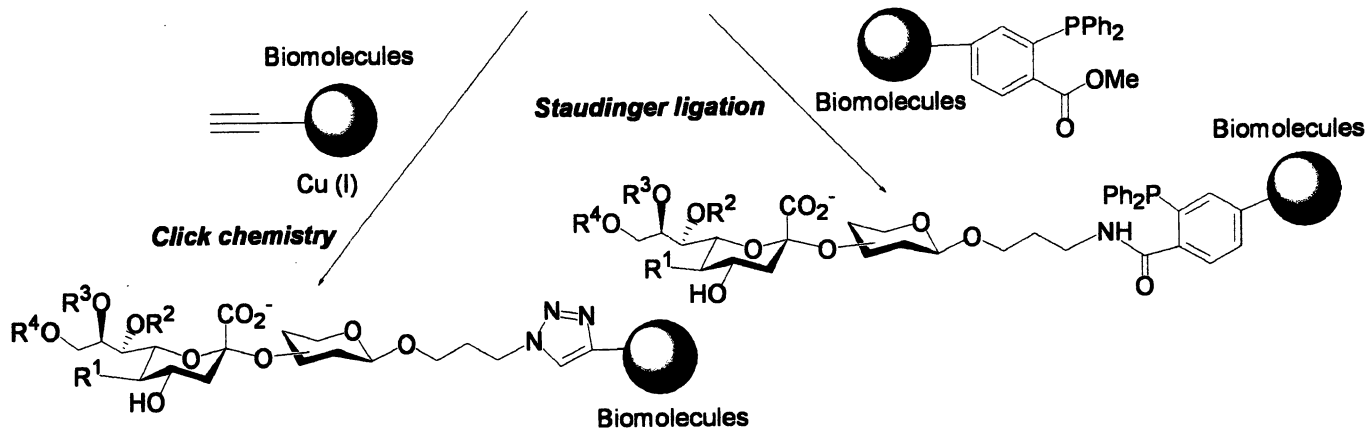


92

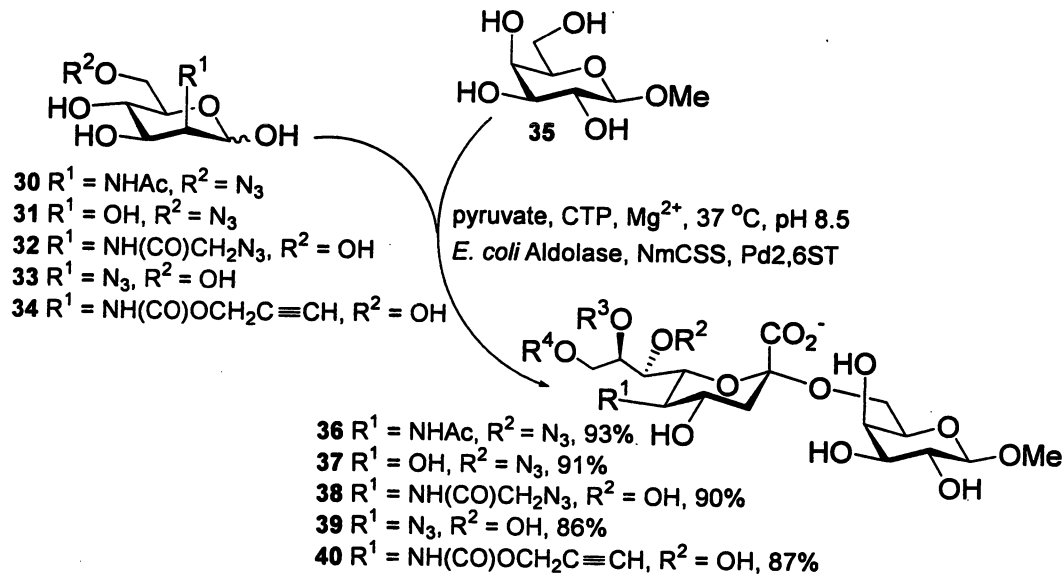


99





Scheme 4. Conjugation of functionalized sialosides with a terminal azido group to other biomolecules.



Scheme 5. One-pot three-enzyme chemoenzymatic synthesis of α 2,6-linked sialosides containing non-natural substituents using Gal β OMe as an acceptor.

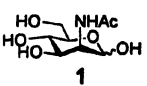
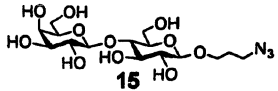
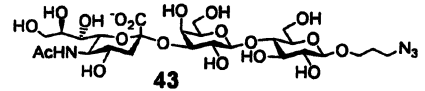
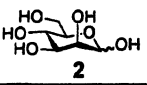
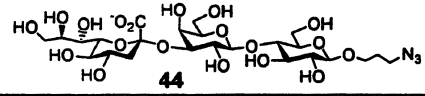
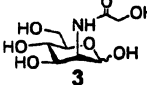
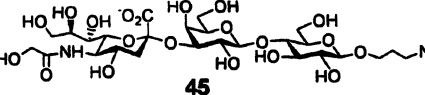
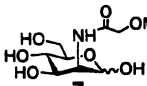
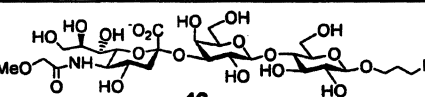
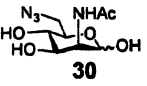
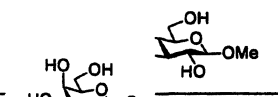
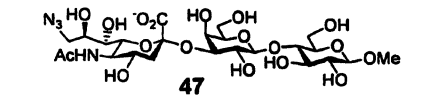
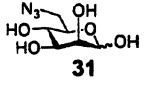
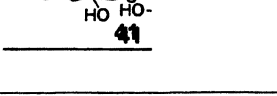
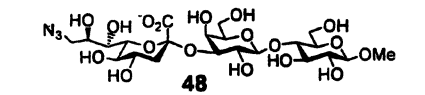
To avoid using expensive CMP-sialic acid as a starting material, Wakarchuk's group constructed and expressed in *E. coli* a fusion protein of CMP-Neu5Ac synthetase and α 2,3-sialyltransferase, both from *Neisseria meningitidis* (55). This fusion protein can sialylate various oligosaccharide acceptors in high yields using Neu5Ac, Neu5Gc, and *N*-propionyl-neuraminic acid as donor precursors. A large (100 gram) scale synthesis of α 2,3-sialyllactose was achieved in 68% yield using this fusion protein with regeneration of CMP-Neu5Ac (starting from lactose, Neu5Ac, phosphoenolpyruvate, catalytic amounts of ATP and CMP).

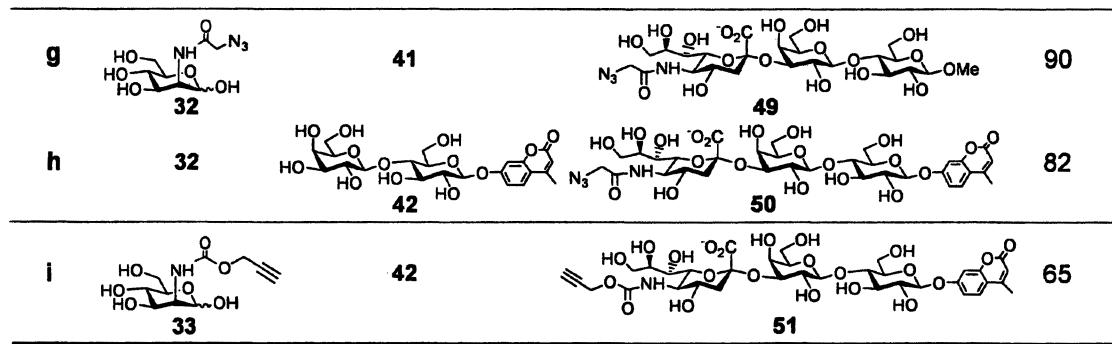
Our group has reported a novel multifunctional bacterial enzyme, *Pasteurella multocida* sialyltransferase (PmST1). PmST1 is an easy to express extremely soluble, and highly active sialyltransferase. Routinely, about 100 mg of PmST1 can be purified from one liter *E. coli* cell culture. PmST1 has four different activities including α 2,3-sialyltransferase (optimal activity at pH 7.5-9.0), α 2,6-sialyltransferase (optimal activity at pH 4.5-7.0), α 2,3-sialidase (optimal activity at pH 5.0-5.5), and trans-sialidase (optimal activity at pH 5.5-6.5) activities. Among them, α 2,3-sialyltransferase activity is the major function of the PmST1 with a specific activity of 60 U/mg protein (24).

By controlling the pH and reaction time, PmST1 was used in preparative scale (20-150 mg) synthesis of α 2,3-linked sialosides using the one-pot three-enzyme system similar to that described above for the Pd2,6ST. When the reaction was set up at the pH 7.0 to 9.0 for 1 to 2 h at 37 °C, only α 2,3-linked sialoside was formed and no significant sialidase or α 2,6SiaT activity was observed under these conditions (24).

As shown in Table 2, PmST1 has very flexible substrate specificity and can use CMP-activated sialic acids containing a number of modifications on the sialic acid residue, obtained *in situ* from their corresponding precursors using the recombinant *E. coli* sialic acid aldolase and NmCSS. Modified sialic acid derivatives with azide, alkyne, or methyl group were transferred by PmST1 to various lactose based acceptors containing azide, 4-methylumbelliferyl (MU), or methyl groups as aglycones to afford naturally/unnaturally occurring α 2,3-linked sialosides 43-51 from their corresponding sialic acid precursors 7 and 30-33 respectively in 65-91% yields using the one-pot three-enzyme system. The azide group in sialosides 43-49 and the alkyne group in sialoside 51 enable them to be conveniently attached to other molecules using the "Click Chemistry" or "Staudinger Ligation" strategy as described above. Sialosides 50, 51 containing the fluorophore (i.e. MU group) at the reducing end can facilitate enzyme activity assays and kinetics studies using High Performance Liquid Chromatography (HPLC) or Capillary Electrophoresis (CE) system with a fluorescence detector.

Table 2. One-Pot Three-Enzyme Chemoenzymatic Synthesis of α 2,3-Linked Sialosides Containing Naturally Occurring and Non-natural Sialic Acids

Entry	Donor Precursors	Acceptor	Product	Yield (%)
a				84
b		15		75
c		15		80
d		15		90
e				91
f				88



High-Throughput Substrate Specificity Studies of Sialidases Using Chemoenzymatically Synthesized *p*NP-Tagged Sialosides

Sialidases are enzymes that catalyze the removal of terminal sialic acids from sialosides (56, 57). Human sialidases play pivotal roles in sialic acid metabolism (58). They relate to a number of disease states such as sialidosis (59-62) and cancer (63-65). Bacterial and viral play significant roles in the pathogenesis and pathology of bacterial and viral infections (66, 67). Viral sialidases such as neuraminidases of influenza virus catalyze the removal of sialic acid from the surface of infected host cells to release the newly formed progeny virus (68).

A considerable interest has been drawn to the substrate specificity studies of sialidases due to their important roles in many physiological and pathological processes. Previous substrate specificity studies of sialidases have been focused on using a limited number of synthetic compounds, or glycoproteins and glycolipids isolated from natural sources before or after glycan remodeling (17, 69-73), which are heterogeneous compounds with varied carbohydrate structures at multiple glycosylation sites. It is necessary to investigate the substrate specificities of sialidases using structurally defined substrates containing diverse naturally occurring sialic acid forms linked with different sialyl linkages to varied penultimate carbohydrates. These compounds are extremely difficult to isolate in homogeneous forms from natural sources (74). Obtaining these compounds by chemical glycosylation also poses a big challenge (20).

Using the highly efficient one-pot three-enzyme chemoenzymatic approach developed in our laboratory, a *p*NP-tagged sialoside library containing different naturally occurring sialic acid forms α 2,3- or α 2,6- linked to *para*-nitrophenyl β -D-galactopyranoside (Gal β -*p*NP, 52) or GalNAc β -*p*NP was obtained (Scheme 6). When using Gal β -*p*NP 52 as an acceptor for PmST1, naturally occurring α 2,3-linked sialosides containing Neu5Ac or its C5- or C9- substituted analogs 54-58 were obtained in 75-91% yields from ManNAc, mannose, or their C6-modified analogs as sialic acid precursors. Similarly, using Gal β -*p*NP as an acceptor for Pd2,6ST, α 2,6-linked sialosides 59-63 were synthesized in 78-92% yields. Using GalNAc β -*p*NP 53 as an acceptor for Pd2,6ST, α 2,6-linked sialosides 64-68 were obtained in 33-92% yields. These *para*-nitrophenol-tagged sialosides containing different naturally occurring sialic acid forms, different sialyl linkages, and different penultimate monosaccharide are the representative terminal disaccharide units of naturally occurring sialoglycoconjugates.

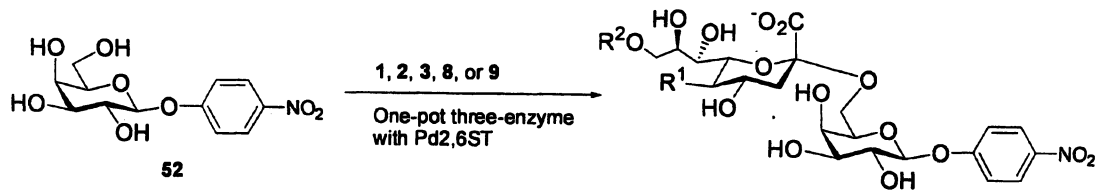
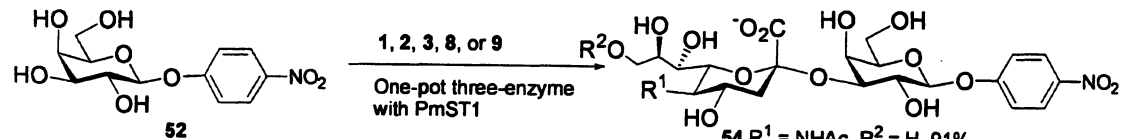
Using the sialosides obtained, a 96-well plate based coupled enzymatic assay for high-throughput colorimetric screening of structures that can be recognized by different sialidases was developed (Scheme 7) (75). Briefly, individual sialosides are incubated with an appropriate amount of a certain sialidase and an excess amount of exogalactosidase or hexosaminidase. If the

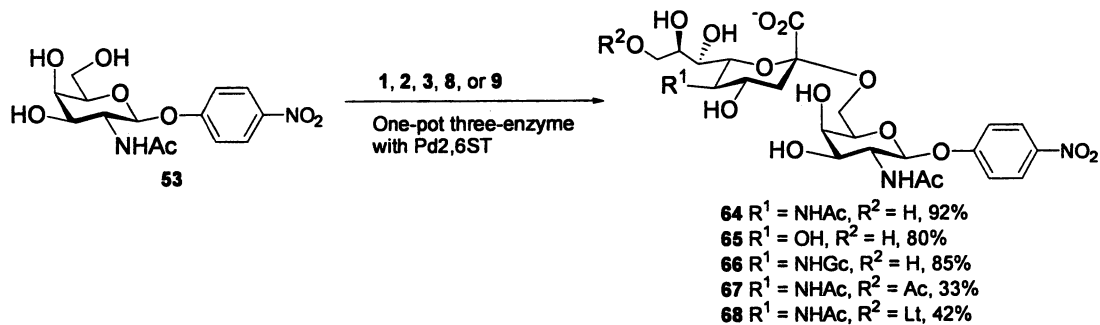
sialoside is a substrate of the sialidase, the terminal sialic acid residue is cleaved by the enzyme to give Gal β -*p*NP or GalNAc β -*p*NP, which is quickly hydrolyzed by the excess amount of galactosidase or hexosaminidase in the reaction mixture to give *p*NP and Gal or GalNAc. The assay is stopped by adjusting the pH of the reaction mixture to 9.6, and the amount of *p*NP formed is determined spectrophotometrically at $A_{405\text{ nm}}$. The amount of the sialic acid released from the sialoside by the sialidase is equivalent to the amount of *p*NP formed. A comparison of the color developed for different sialosides by different sialidases reveals the effect of the structural diversity of the sialoside substrate on the hydrolytic activity of different sialidases.

Seven different bacterial sialidases, including six commercially available sialidases from *Arthrobacter ureafaciens*, *Clostridium perfringens*, *Streptococcus sp.* IID, *Vibrio cholera*, *Salmonella typhimurium*, and *Streptococcus pneumoniae*, as well as PmST1 which also possesses sialidase activity (20), were used as model systems to test the application of the sialoside library and the 96-well plate based high-throughput colorimetric screening method.

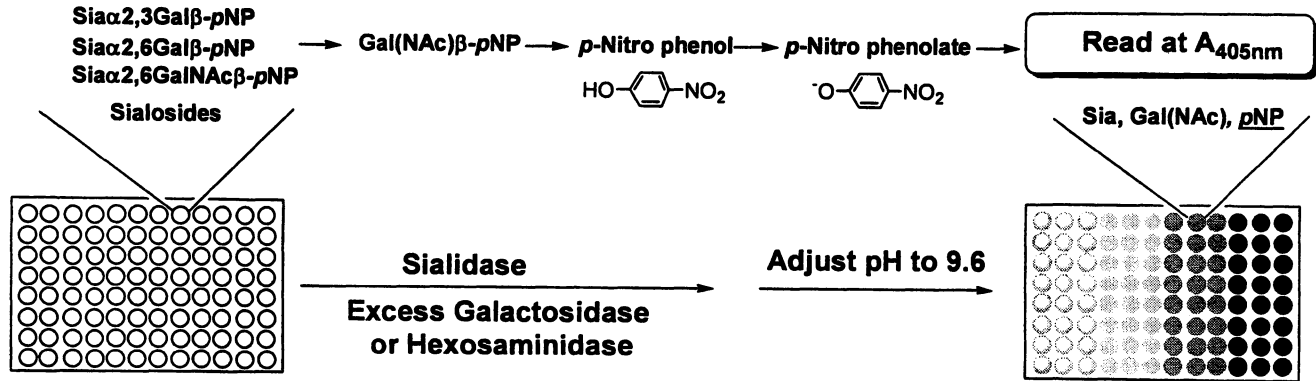
Results showed that four bacterial sialidases including those from *A. ureafaciens*, *C. perfringens*, *Streptococcus sp.* IID, and *V. cholera* were able to cleave both α 2,3- and α 2,6- linked sialosides with high efficiency. *S. typhimurium* and PmST1 sialidases were specific for cleaving α 2,3-linked sialosides. *S. typhimurium* sialidase was very efficient in cleaving α 2,3-linked sialosides, but it cleaved α 2,6-linked sialosides at a much slower rate (200 fold less). Among seven bacterial sialidases tested, PmST1 was the only enzyme that can cleave KDN from KDN α 2,3-linked galactoside efficiently (8.5% compared to Neu5Ac α 2,3-linked galactoside). *S. pneumoniae* sialidase cleaved KDN from KDN α 2,3Gal linkage at a much slower rate (0.7% compared to Neu5Ac α 2,3-linked galactoside). All other sialidases had no hydrolytic activity on KDN-containing sialosides, indicating the importance of the acylamido group at C-5 of sialic acid for the hydrolysis by sialidases.

Regardless of the acceptor sugar or the sialyl linkage, 9-*O*-acetylation on Neu5Ac had little effect on the hydrolysis of sialosides by sialidases from *A. ureafaciens*, *Streptococcus sp.* IID, and *S. pneumoniae* (the hydrolytic activity decreases to 60-80%), but it decreased the hydrolytic activity of sialidases from *C. perfringens*, *V. cholera*, *S. pneumoniae*, and PmST1 to 20-40%. In comparison, 9-*O*-lactylation modification dramatically reduced the hydrolysis rate for both *A. ureafaciens* and *C. perfringens* sialidases, but it was well tolerated by the *Streptococcus sp.* IID sialidase. This substrate specificity study also provides us information for selecting suitable sialidases for glycan analysis of biological samples.





Scheme 6. One-pot three-enzyme chemoenzymatic synthesis of pNP-tagged α 2,3- and α 2,6-linked sialoside libraries containing a list of naturally occurring sialic acid forms.



Scheme 7. High-throughput substrate specificity assays for sialidases

Conclusion

The one-pot three-enzyme chemoenzymatic system is a highly efficient and convenient approach for the synthesis of libraries of functionalized α 2,6- and α 2,3- linked sialosides. Starting from the six carbon sugar precursors of sialic acids (ManNAc, mannose, and their derivatives), sialosides containing various sialic acid modifications were synthesized by taking the advantage of the relaxed substrate specificity of bacterial sialoside biosynthetic enzymes. A library of pNP-tagged sialosides with different naturally occurring sialic acid forms and sialyl linkages were prepared using this approach and applied in the systematic study of the substrate specificity of sialidases in a 96-well plate-based colorimetric high-throughput screening platform. The chemoenzymatic synthetic approach will also be useful in glycan remodelling of glycoprotein or cell surface to study the biological function of certain glycan structures.

Acknowledgements

This work was supported by start-up funds from the Regents of the University of California, Mizutani Foundation for Glycoscience, and NIH R01GM076360.

References

1. Angata, T.; Varki, A. *Chem. Rev.* **2002**, *102*, 439-469.
2. Schauer, R. *Glycoconj. J.* **2000**, *17*, 485-499.
3. Vimr, E.; Lichtensteiger, C. *Trends Microbiol.* **2002**, *10*, 254-257.
4. Mandrell, R. E.; Griffiss, J. M.; Macher, B. A. *J Exp Med* **1988**, *168*, 107-126.
5. Mandrell, R. E.; McLaughlin, R.; Aba Kwaik, Y.; Lesse, A.; Yamasaki, R.; Gibson, B.; Spinola, S. M.; Apicella, M. A. *Infect. Immun.* **1992**, *60*, 1322-1328.
6. Mandrell, R. E.; Apicella, M. A. *Immunobiology* **1993**, *187*, 382-402.
7. Lewis, A. L.; Nizet, V.; Varki, A. *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 11123-11128.
8. Shi, W. X.; Chammas, R.; Varki, N. M.; Powell, L.; Varki, A. *J. Biol. Chem.* **1996**, *271*, 31526-31532.
9. Sharma, V.; Chatterjee, M.; Sen, G.; Kumar, C. A.; Mandal, C. *Glycoconj. J.* **2000**, *17*, 887-893.
10. Rogers, G. N.; Herrler, G.; Paulson, J. C.; Klenk, H. D. *J. Biol. Chem.* **1986**, *261*, 5947-5951.

11. Herrler, G.; Rott, R.; Klenk, H. D.; Muller, H. P.; Shukla, A. K.; Schauer, R. *Embo. J.* **1985**, *4*, 1503-1506.
12. Schauer, R.; Schmid, H.; Pommerencke, J.; Iwersen, M.; Kohla, G. *In Molecular immunology of complex carbohydrates 2*, edited by Wu, A. M., Plenum, New York, **2001**, 325-342.
13. Gottschalk, A. *Biochim. Biophys. Acta.* **1957**, *23*, 645-646.
14. Klenk, E.; Faillard, H.; Lempfrid, H. *Hoppe Seylers Z Physiol. Chem.* **1955**, *301*, 235-246.
15. Regl, G.; Kaser, A.; Iwersen, M.; Schmid, H.; Kohla, G.; Strobl, B.; Vilas, U.; Schauer, R.; Vlasak, R. *J. Virol.* **1999**, *73*, 4721-4727.
16. Higashi, H.; Hirabayashi, Y.; Fukui, Y.; Naiki, M.; Matsumoto, M.; Ueda, S.; Kato, S. *Cancer Res.* **1985**, *45*, 3796-3802.
17. Corfield, A. P.; Veh, R. W.; Wember, M.; Michalski, J. C.; Schauer, R. *Biochem. J.* **1981**, *197*, 293-299.
18. Varki, A.; Diaz, S. *Anal. Biochem.* **1984**, *137*, 236-247.
19. Boons, G. J.; Demchenko, A. V. *Chem. Rev.* **2000**, *100*, 4539-4566.
20. Kiefel, M. J.; von Itzstein, M. *Chem. Rev.* **2002**, *102*, 471-490.
21. Koeller, K. M.; Wong, C. H. *Nature* **2001**, *409*, 232-240.
22. Ichikawa, Y.; Shen, G. J.; Wong, C. H. *J. Am. Chem. Soc.* **1991**, *113*, 4698-4700.
23. Blixt, O.; Paulson, J. C. *Adv. Synth. Catal.* **2003**, *345*, 687-690.
24. Yu, H.; Chokhawala, H.; Karpel, R.; Wu, B.; Zhang, J.; Zhang, Y.; Jia, Q.; Chen, X. *J. Am. Chem. Soc.* **2005**, *127*, 17618-17619.
25. Yu, H.; Huang, S. S.; Chokhawala, H.; Sun, M. C.; Zheng, H. J.; Chen, X. *Angew. Chem. Int. Ed. Engl.* **2006**, *45*, 3938-3944.
26. Yu, H.; Yu, H.; Karpel, R.; Chen, X. *Bioorg. Med. Chem.* **2004**, *12*, 6427-6435.
27. Yu, H.; Chen, X. *Org. Lett.* **2006**, *8*, 2393-2396.
28. Huang, S.; Yu, H.; Chen, X. *Angew. Chem. Int. Ed. Engl.* **2007**, *46*, 2249-2253.
29. Gross, H. J.; Brossmer, R. *Glycoconj. J.* **1987**, *4*, 145-156.
30. Gross, H. J.; Brossmer, R. *Glycoconj. J.* **1995**, *12*, 739-746.
31. Gross, H. J.; Rose, U.; Krause, J. M.; Paulson, J. C.; Schmid, K.; Feeney, R. E.; Brossmer, R. *Biochemistry* **1989**, *28*, 7386-7392.
32. Dufner, G.; Schworer, R.; Muller, B.; Schmidt, R. R. *Eur. J. Org. Chem.* **2000**, 1467-1482.
33. Yamamoto, T.; Nakashizuka, M.; Kodama, H.; Kajihara, Y.; Terada, I. *J. Biochem. (Tokyo)* **1996**, *120*, 104-110.
34. Yamamoto, T.; Nakashizuka, M.; Terada, I. *J. Biochem. (Tokyo)* **1998**, *123*, 94-100.
35. Kajihara, Y.; Yamamoto, T.; Nagae, H.; Nakashizuka, M.; Sakakibara, T.; Terada, I. *J. Org. Chem.* **1996**, *61*, 8632-8635.

36. Endo, T.; Koizumi, S.; Tabata, K.; Kakita, S.; Ozaki, A. *Carbohydr. Res.* **2001**, *330*, 439-443.
37. Teo, C. F.; Hwang, T. S.; Chen, P. H.; Hung, C. H.; Gao, H. S.; Chang, L. S.; Lin, C. H. *Adv. Synth. Catal.* **2005**, *347*, 967-972.
38. Yamamoto, T.; Nagae, H.; Kajihara, Y.; Terada, I. *Biosci. Biotechnol. Biochem.* **1998**, *62*, 210-214.
39. Kondejewski, L. H.; Kralovec, J. A.; Blair, A. H.; Ghose, T. *Bioconjug. Chem.* **1994**, *5*, 602-611.
40. Meikle, P. J.; Bundle, D. R. *Glycoconj. J.* **1990**, *7*, 207-218.
41. Bergh, A.; Magnusson, B. G.; Ohlsson, J.; Wellmar, U.; Nilsson, U. J. *Glycoconj. J.* **2001**, *18*, 615-621.
42. Tietze, L. F.; Arlt, M.; Beller, M.; Glusenkamp, K. H.; Jahde, E.; Rajewsky, M. F. *Chem. Ber.* **1991**, *124*, 1215-1221.
43. Ragupathi, G.; Coltart, D. M.; Williams, L. J.; Koide, F.; Kagan, E.; Allen, J.; Harris, C.; Glunz, P. W.; Livingston, P. O.; Danishefsky, S. J. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 13699-13704.
44. Wan, Q.; Chen, J. H.; Chen, G.; Danishefsky, S. J. *J. Org. Chem.* **2006**, *71*, 8244-8249.
45. Graminski, G. F.; Carlson, C. L.; Ziemer, J. R.; Cai, F.; Vermeulen, N. M.; Vanderwerf, S. M.; Burns, M. R. *Bioorg. Med. Chem. Lett.* **2002**, *12*, 35-40.
46. Rich, J. R.; Wakarchuk, W. W.; Bundle, D. R. *Chem. Eur. J.* **2006**, *12*, 845-858.
47. Yu, H.; Chokhawala, H. A.; Varki, A.; Chen, X. *Org. Biomol. Chem.* **2007**, *5*, 2458-2463.
48. Fazio, F.; Bryan, M. C.; Blixt, O.; Paulson, J. C.; Wong, C. H. *J. Am. Chem. Soc.* **2002**, *124*, 14397-14402.
49. Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem. Int. Ed. Engl.* **2001**, *40*, 2004-2021.
50. Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2005**, *127*, 2686-2695.
51. Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010.
52. Izumi, M.; Shen, G. J.; Wacowich-Sgarbi, S.; Nakatani, T.; Plettenburg, O.; Wong, C. H. *J. Am. Chem. Soc.* **2001**, *123*, 10909-10918.
53. Gilbert, M.; Brisson, J. R.; Karwaski, M. F.; Michniewicz, J.; Cunningham, A. M.; Wu, Y. Y.; Young, N. M.; Wakarchuk, W. W. *J. Biol. Chem.* **2000**, *275*, 3896-3906.
54. Blixt, O.; Vasiliu, D.; Allin, K.; Jacobsen, N.; Warnock, D.; Razi, N.; Paulson, J. C.; Bernatchez, S.; Gilbert, M.; Wakarchuk, W. *Carbohydr. Res.* **2005**, *340*, 1963-1972.
55. Gilbert, M.; Bayer, R.; Cunningham, A. M.; Defrees, S.; Gao, Y. H.; Watson, D. C.; Young, N. M.; Wakarchuk, W. W. *Nat. Biotechnol.* **1998**, *16*, 769-772.

56. Takada, K.; Hamada, T.; Hirota, H.; Nakao, Y.; Matsunaga, S.; van Soest, R. W.; Fusetani, N. *Chem. Biol.* **2006**, *13*, 569-574.
57. Corfield, T. *Glycobiology* **1992**, *2*, 509-521.
58. Monti, E.; Bassi, M. T.; Bresciani, R.; Civini, S.; Croci, G. L.; Papini, N.; Riboni, M.; Zanchetti, G.; Ballabio, A.; Preti, A.; Tettamanti, G.; Venerando, B.; Borsani, G. *Genomics* **2004**, *83*, 445-453.
59. Milner, C. M.; Smith, S. V.; Carrillo, M. B.; Taylor, G. L.; Hollinshead, M.; Campbell, R. D. *J. Biol. Chem.* **1997**, *272*, 4549-4558.
60. Bonten, E.; van der Spoel, A.; Fornerod, M.; Grosveld, G.; d'Azzo, A. *Genes. Dev.* **1996**, *10*, 3156-3169.
61. Pshezhetsky, A. V.; Richard, C.; Michaud, L.; Igdoura, S.; Wang, S.; Elsliger, M. A.; Qu, J.; Leclerc, D.; Gravel, R.; Dallaire, L.; Potier, M. *Nat. Genet.* **1997**, *15*, 316-320.
62. Achyuthan, K. E.; Achyuthan, A. M. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* **2001**, *129*, 29-64.
63. Kakugawa, Y.; Wada, T.; Yamaguchi, K.; Yamanami, H.; Ouchi, K.; Sato, I.; Miyagi, T. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 10718-10723.
64. Meuillet, E. J.; Kroes, R.; Yamamoto, H.; Warner, T. G.; Ferrari, J.; Mania-Farnell, B.; George, D.; Rebbaa, A.; Moskal, J. R.; Bremer, E. G. *Cancer Res.* **1999**, *59*, 234-240.
65. Miyagi, T.; Wada, T.; Yamaguchi, K.; Hata, K. *Glycoconj. J.* **2004**, *20*, 189-198.
66. Vimr, E. R. *Trends Microbiol.* **1994**, *2*, 271-277.
67. Corfield, T. *Glycobiology* **1992**, *2*, 509-521.
68. Suzuki, Y. *Biol. Pharm. Bull.* **2005**, *28*, 399-408.
69. Paulson, J. C.; Weinstein, J.; Dorland, L.; van Halbeek, H.; Vliegthart, J. F. *J. Biol. Chem.* **1982**, *257*, 12734-12738.
70. Corfield, A. P.; Higa, H.; Paulson, J. C.; Schauer, R. *Biochim. Biophys. Acta* **1983**, *744*, 121-126.
71. Corfield, A. P.; Sanderwewer, M.; Veh, R. W.; Wember, M.; Schauer, R. *Biol. Chem. Hoppe-Seyler* **1986**, *367*, 433-439.
72. Kleineidam, R. G.; Furuhata, K.; Ogura, H.; Schauer, R. *Biol. Chem. Hoppe-Seyler* **1990**, *371*, 715-719.
73. Xu, G. Y.; Suzuki, T.; Maejima, Y.; Mizoguchi, T.; Tsuchiya, M.; Kiso, M.; Hasegawa, A.; Suzuki, Y. *Glycoconj. J.* **1995**, *12*, 156-161.
74. Chappell, M. D. H., R. L. *J. Am. Chem. Soc.* **1997**, *119*, 3393-3394.
75. Chokhawala, H. A.; Yu, H.; Chen, X. *ChemBioChem* **2007**, *8*, 194-201.

Chapter 6

Convergent Chemoenzymatic Synthesis of Complex *N*-Glycopeptides

Lai-Xi Wang

Institute of Human Virology and Department of Biochemistry and
Molecular Biology, University of Maryland School of Medicine,
725 West Lombard Street, Baltimore, MD 21201

This article describes recent advances in the endoglycosidase-catalyzed transglycosylation method for the synthesis of large, homogeneous *N*-linked glycopeptides. The exploration and exploitation of sugar oxazolines as activated, transition state analog substrates for an efficient enzymatic transglycosylation were discussed in details, which holds a great promise for glycosylation engineering of glycoproteins.

Introduction

Glycosylation is one of the most ubiquitous posttranslational modifications of proteins in eukaryotes. Carbohydrates can be attached to a protein in two major forms. In the *O*-linked form, the first monosaccharide GalNAc of the glycan is connected to the hydroxyl group of serine or threonine in an α -*O*-glycosidic linkage; in the *N*-linked form, the first GlcNAc in *N*-glycan is attached to the amide side chain of an asparagine residue in a conserved Asn-Xaa-Ser/Thr sequon, where Xaa can be any natural amino acid but proline. *N*-linked glycosylation is the predominant form (1). Numerous examples have demonstrated that the covalently linked oligosaccharides can affect protein's structure and *in vivo* biological activities, and are directly involved in many important biological events such as cell adhesion, viral/bacterial infection, and immune response (2-5). Since glycosylation is not under the direct genetic

control, the posttranslational glycan modification events (i.e., trimming and processing) can be affected by a number of factors, such as protein's sequence, folding, the availability of respective processing enzymes, and the *in situ* activity of these enzymes in different subcellular locations. As a result, glycoproteins are usually not produced in a unified form, but as a mixture of heterogeneous glycoforms, which possess the same polypeptide backbone but differ in the pendent sugar chains. To meet the urgent need of homogeneous glycopeptides and glycoproteins for structural and functional studies, various elegant biological (e.g., cellular glycosylation engineering), chemical, and chemoenzymatic methods have been developed in the past two decades, which have been well reviewed in a number of excellent review articles (6-17). Despite tremendous progress in the field, however, it should be pointed out that each synthetic method has its own limitations, and the synthesis of large, homogeneous glycopeptides and glycoproteins is still a challenging task. This article describes a highly convergent chemoenzymatic approach for the construction of homogeneous glycopeptides and glycoproteins, which employs an endoglycosidase-catalyzed transglycosylation as the key step. In contrast to typical glycosyltransferase-catalyzed synthesis that adds monosaccharide one at a time, the unique advantage of the endoglycosidase-catalyzed method is the transfer of a large intact oligosaccharide to a peptide or protein moiety in a single step, thus providing a highly convergent approach. Recent advances in the chemoenzymatic method, as well as its scope and limitations, is highlighted.

ENGase-Catalyzed Transglycosylation for Glycopeptide Synthesis

Endo- β -*N*-acetylglucosaminidases (ENGases) (EC 3.2.1.96) are a special class of endoglycosidases that hydrolyze the β -1,4-glycosidic bond of the core *N,N'*-diacetylchitobiose moiety in N-glycoproteins to release the N-glycans. The enzymes are found in microorganisms, plants, animals, and human cells (18). Besides hydrolytic activity, a few ENGases have been found to possess transglycosylation activity, i.e., the ability to transfer the releasing oligosaccharide moiety to a suitable acceptor other than water. These include: Endo-A from *Arthrobacter protophormiae* (19,20); Endo-M from *Mucor hiemalis* (21-23); and Endo-CE from *C. elegans* (Endo-CE) (24). These ENGases also showed distinct substrate specificity. For example, Endo-A and Endo-CE are specific for high-mannose or hybrid type N-glycans (20,24), whereas Endo-M can work on both high-mannose and complex type *N*-glycans, with preference for the complex type (23). Initial studies showed that Endo-A and Endo-M could transfer the releasing oligosaccharides from natural *N*-glycans to some monosaccharides such as GlcNAc and Glc (19,25). Later on, it was found that GlcNAc-containing peptides could serve as an acceptor for the trans-

glycosylation to form the corresponding *N*-glycopeptides carrying a native oligosaccharide moiety (25-27). In contrast to common glycosyltransferases that transfer only monosaccharides one at a time, the biggest advantage of the ENGase-catalyzed reaction is the transfer of a large intact oligosaccharide to a GlcNAc-containing peptide or protein in a single step, without the need of using any protecting groups (Figure 1).

A number of large homogeneous glycopeptides of biological interest have been synthesized by this convergent chemoenzymatic method, including the glycosylated calcitonin (28), a *C*-linked analog of *N*-glycopeptides (27,29), a *glc*-containing glycopeptide (30), a glycosylated fragment of the nicotinic acetylcholine receptor (nAChR) (31), and a glycosylated substance P (32). The chemoenzymatic synthesis and application of these bioactive glycopeptides have been summarized in our previous review article (33). Recent applications of the chemoenzymatic method include the synthesis of HIV-1 envelope glycoprotein fragments for elucidating the effects of glycosylation on peptide's global conformations and biological properties (34,35). A concise chemoenzymatic synthesis of the CD52 antigen carrying full-size high-mannose and complex type *N*-glycans was also accomplished (Figure 2) (36). These large, homogeneous glycopeptides are otherwise difficult to obtain by other means for structural and functional studies. However, as ENGases are inherently glycohydrolases, a broader application of the ENGases for synthetic purpose has been hampered by the relatively low transglycosylation yield (usually 5-20%) and by the product hydrolysis by the enzyme. Although the incorporation of organic solvents in the reaction medium can enhance the transglycosylation yield to some extent (37), the overall efficiency is generally low. In addition, the limitation to the use of only natural *N*-glycans or *N*-glycopeptides as the donor substrates further constrains the usefulness of the approach, as these natural donor substrates themselves are difficult to obtain. A major advance in enhancing the transglycosylation efficiency was made recently by exploring synthetic sugar oxazolines, the presumed transition-state analog substrates, for the enzymatic transglycosylation. This new development has led to a highly efficient chemoenzymatic approach for glycopeptide synthesis and glycoprotein glycosylation engineering (38-40).

Transition-State Analog Substrates for ENGase-Catalyzed Transglycosylation

To address the problems associated with the ENGase-catalyzed transglycosylation, e.g., the low transglycosylation yield, the issue of product hydrolysis, and the limitation to the use of only natural *N*-glycans as the donor substrates, we sought to explore highly activated species as donor substrates, which might be kinetically more favorable for the transglycosylation than the

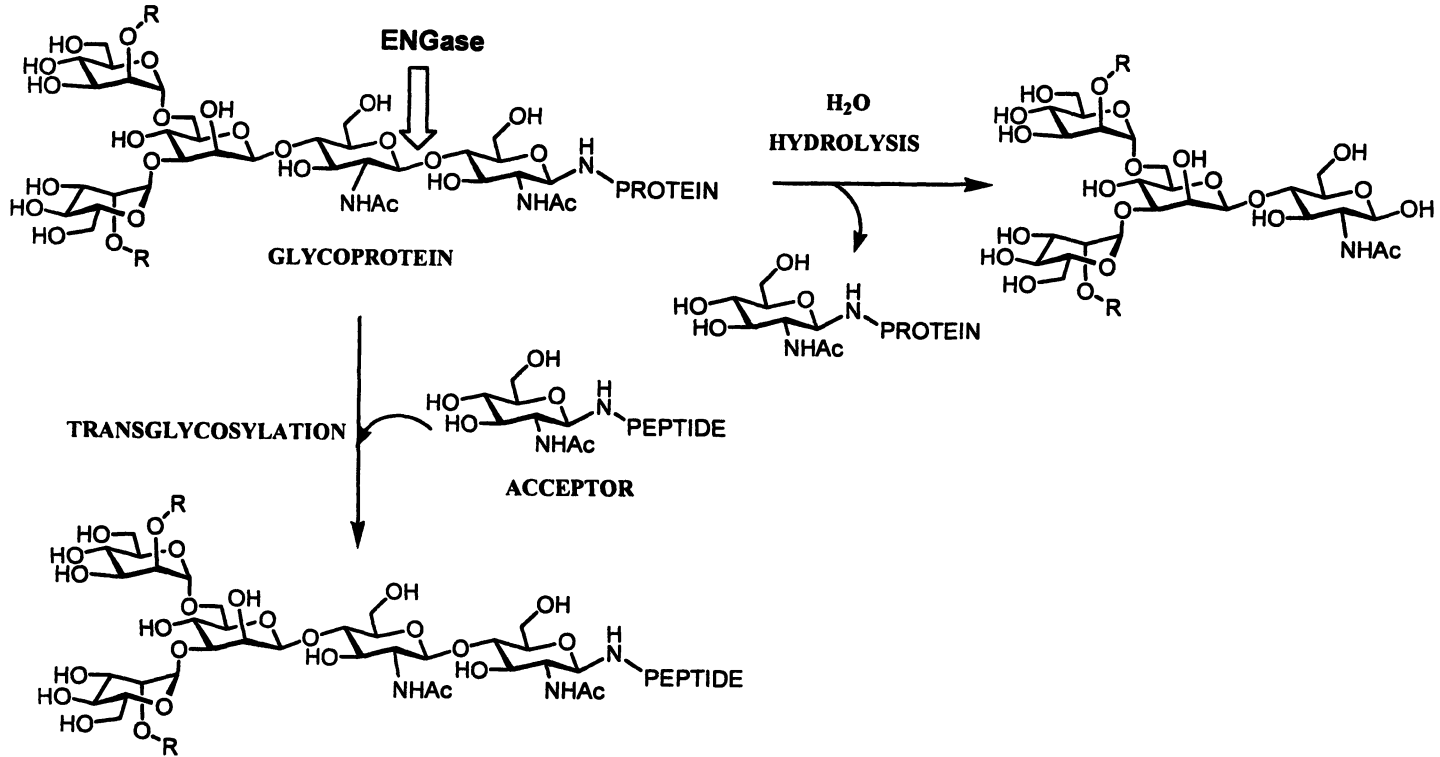


Figure 1. ENGase-catalyzed hydrolysis and transglycosylation

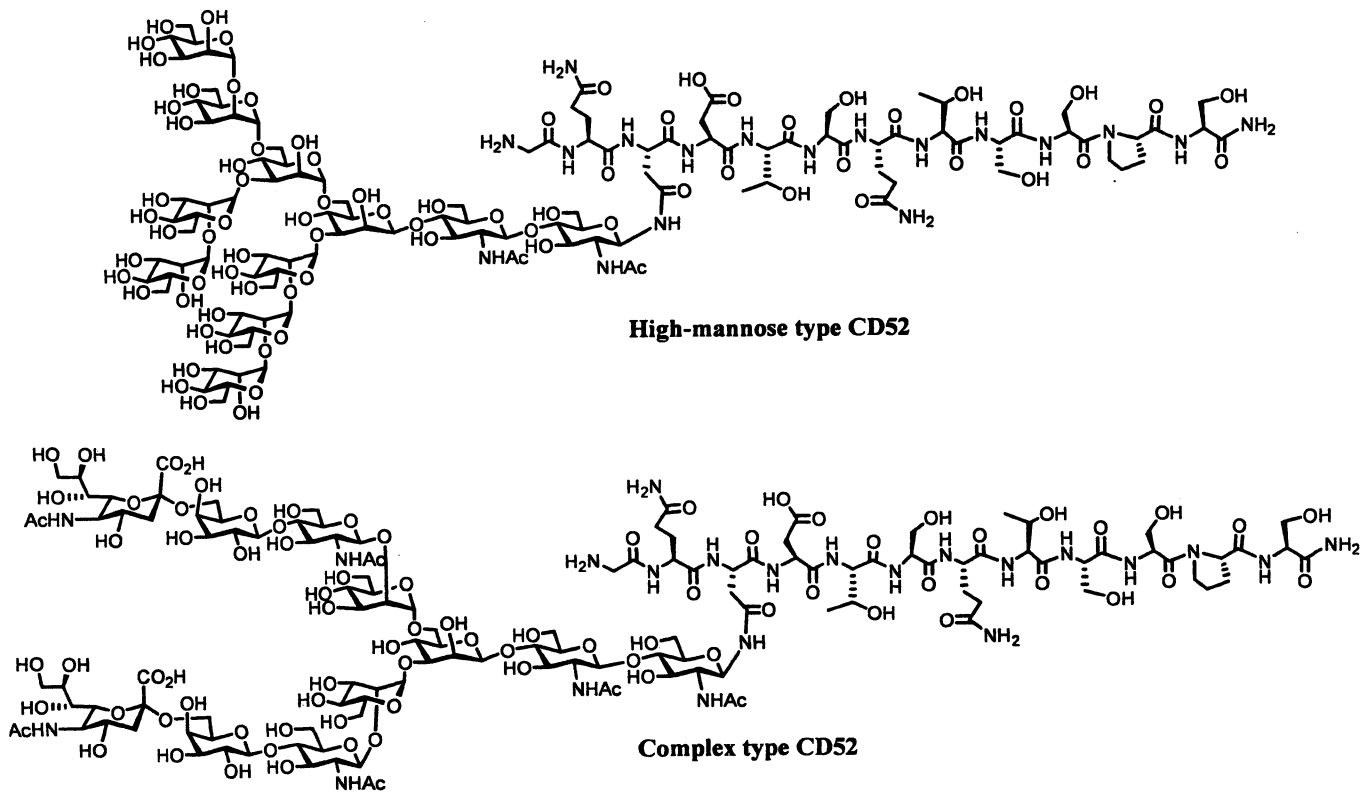


Figure 2. Structures of synthetic CD52 antigen carrying natural glycans

ground-state substrates. In analogy to some family 18 chitinases (41-43) and family 20 *N*-acetyl- β -hexosaminidases (44,45), which hydrolyze the carbohydrate substrates via a substrate-assisted mechanism involving the formation of an oxazolinium ion intermediate, the ENGase-catalyzed reaction (hydrolysis and/or transglycosylation) might also proceed *via* the substrate-assisted mechanism through the participation of the 2-*N*-acetamido group to form an 1,2-oxazoline ion intermediate (Figure 3). Structural studies on closely related enzymes such as Endo-H and Endo-F provided hints supporting this mechanism (46,47). We anticipated that synthetic sugar oxazolines corresponding to the *N*-glycan core structure should be highly active as donor substrates and the transglycosylation rate of the transition-state analogs should be much higher than the hydrolysis of the transglycosylation product (viewed as ground-state substrate). More direct evidence was provided by Fujita and co-workers, showing that a disaccharide Man β 1,4GlcNAc-oxazoline could serve as substrate of Endo-A and Endo-M for transglycosylation to GlcNAc-PNP to form a trisaccharide derivative (48). In addition, there are also precedents that disaccharide oxazoline derivatives as transition-state mimics could serve as donor substrates of some endoglycosidases such as chitinases and hyaluronidases for oligo- and poly-saccharide synthesis (49-54). These findings prompted us to explore the corresponding oligosaccharide oxazolines, the presumed transition-state mimics, as donor substrates for *N*-glycopeptide and *N*-glycoprotein synthesis.

A concise synthesis of the di- and tetra-saccharide oxazolines corresponding to the core of *N*-glycans was described (Scheme 1). A two-step approach was employed to construct the Man β 1,4GlcN disaccharide, which was regarded as a difficult linkage for chemical synthesis (12). This involves the stereocontrolled β -glycosidation of the 2-*O*-acetyl-glucose 1 with acceptor 2, and the subsequent inversion of the *Glc* C-2 configuration to provide the desired disaccharide 4. Compound 4 was converted into the peracetylated intermediate 5 by protecting group manipulations, and the oxazoline formation was efficiently achieved by treatment with TMS-Br/BF₃ Et₂O in the presence of collidine to give the disaccharide oxazoline 6 after de-*O*-acetylation. To synthesize the tetra-saccharide oxazoline, compound 4 was changed to the diol 7, which was then glycosidated with two equivalents of the mannosyl imidate 8 to provide the tetrasaccharide intermediate 9. Protecting group manipulations and subsequent oxazoline formation gave the tetrasaccharide oxazoline 11 in good yield (Scheme 1).

The ENGase-catalyzed transglycosylation was first examined with a GlcNAc-containing heptapeptide (12) derived from HIV-1 gp120 as the acceptor. It was found that the disaccharide oxazoline 6 was a good substrate for ENGase-catalyzed transglycosylation to form the corresponding glycopeptide 13 in 72% yield (Scheme 2). The enzymatic transglycosylation proceeded in a regio- and stereo-specific manner and the newly formed glycosidic linkage was

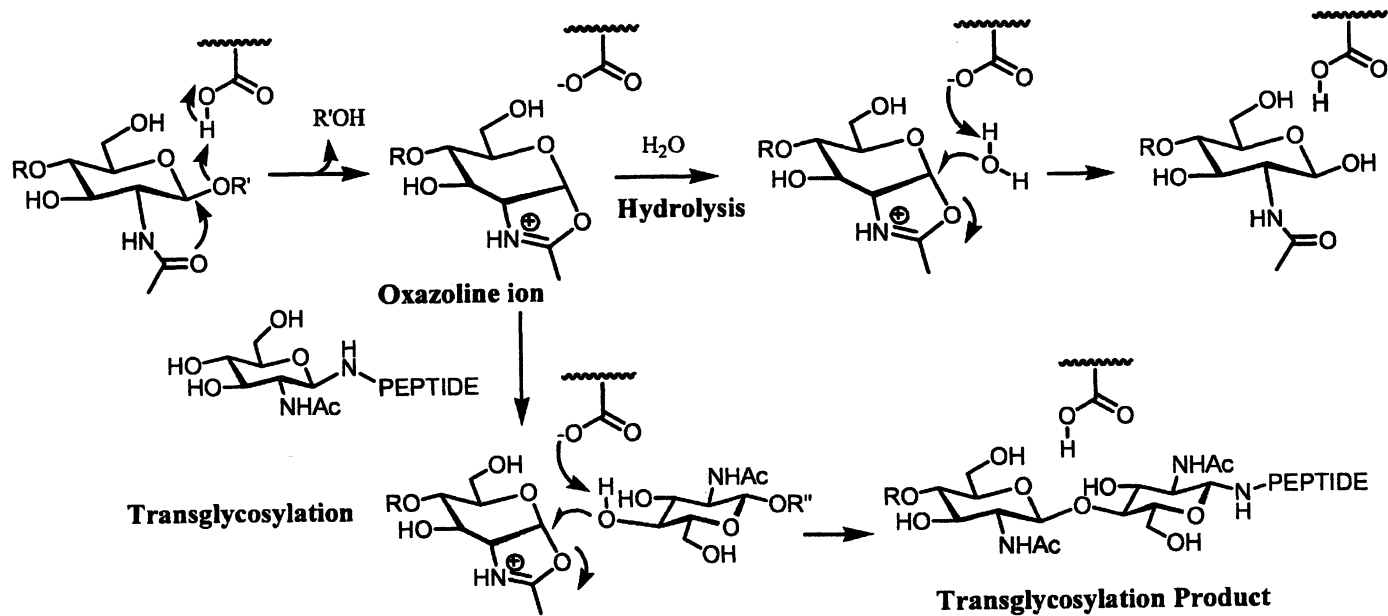
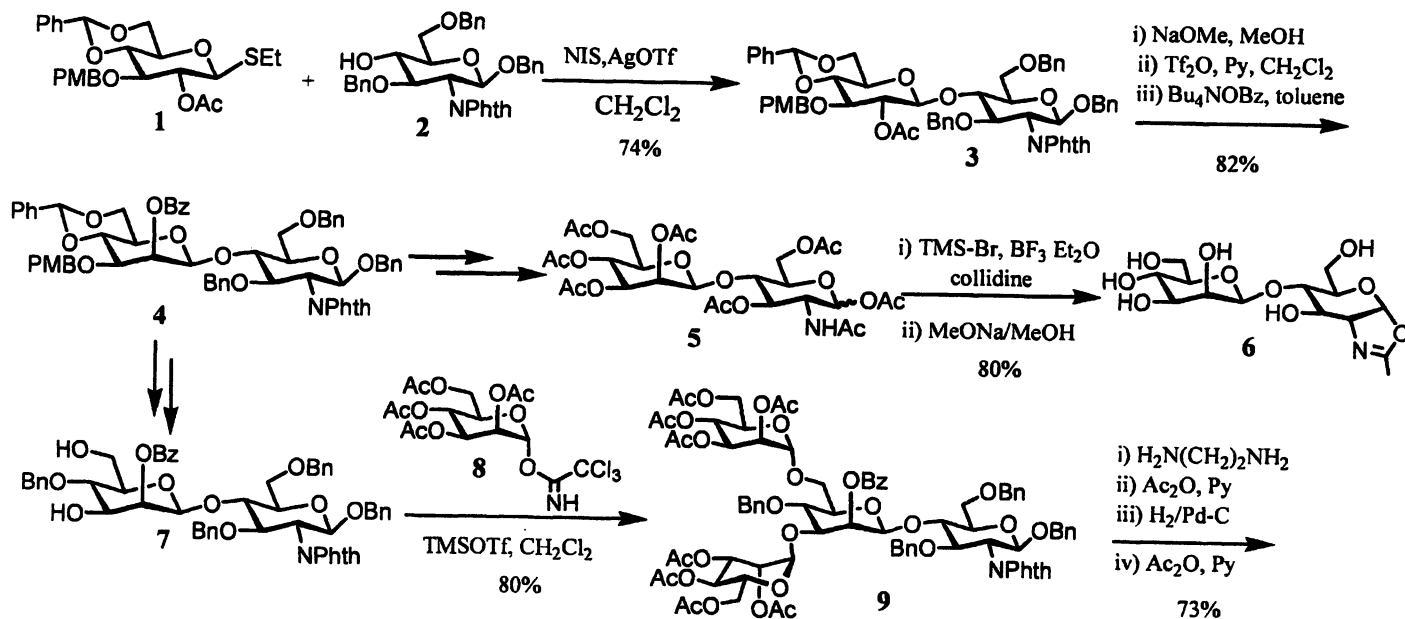
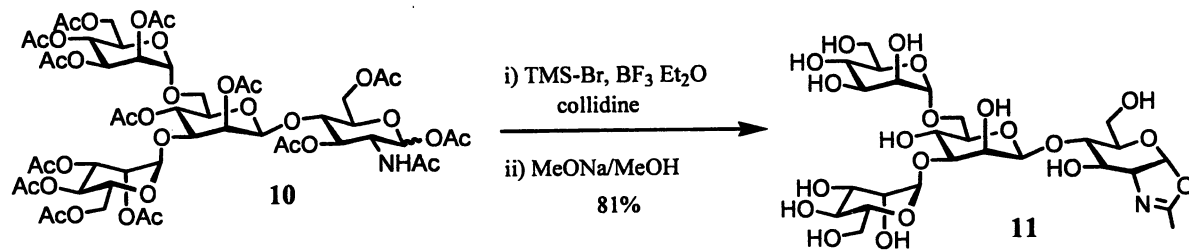


Figure 3. Substrate-assisted mechanism for *ENGase*-catalyzed reactions





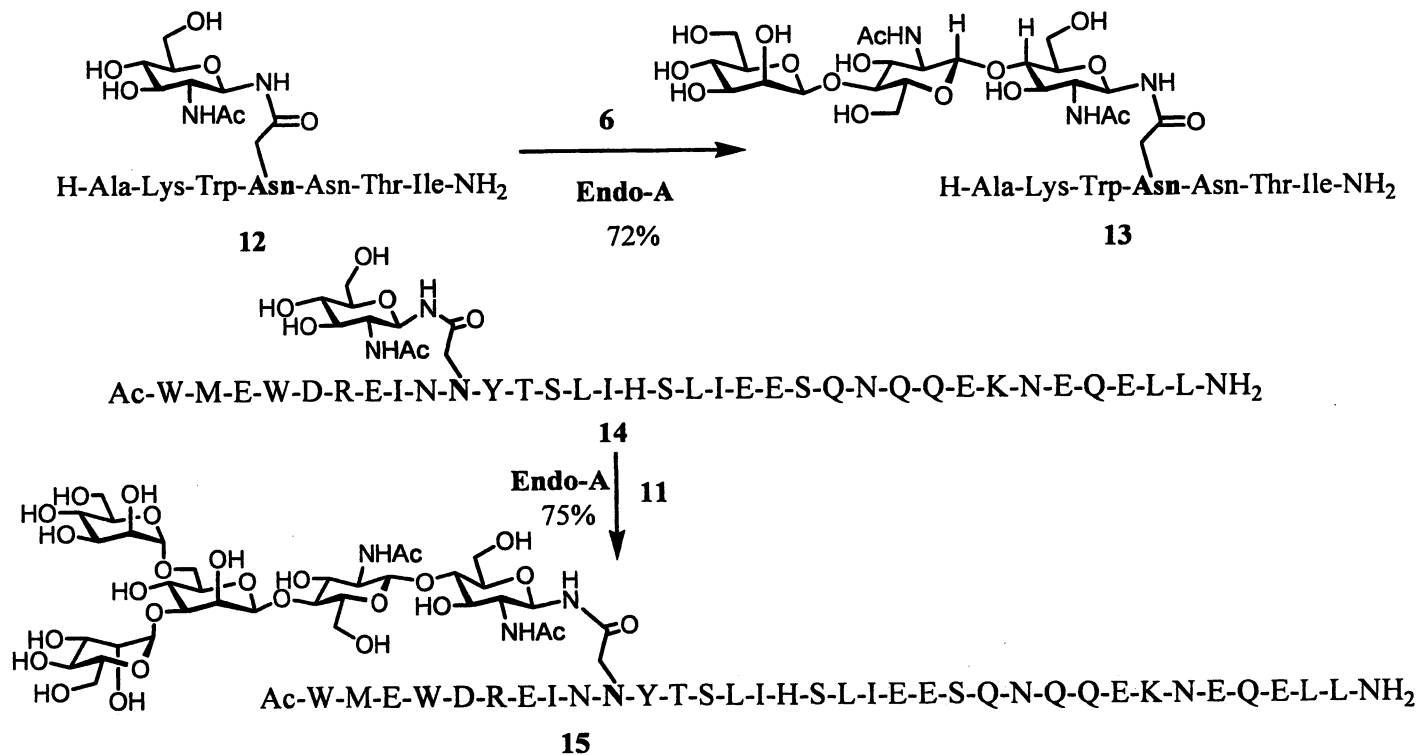
Scheme 1. Synthesis of di- and tetra-saccharide sugar oxazolines

confirmed by 2D-NMR to be the desired β -1,4-glycosidic linkage found in the natural form (38). Both Endo-A and Endo-M are efficient for the transglycosylation with the sugar oxazoline. Interestingly, while the disaccharide oxazoline **6** acted as an efficient substrate for transglycosylation, the resulting glycopeptide **13** carrying the core trisaccharide was actually resistant to the hydrolysis by ENGases, thus allowing the accumulation of the product. This approach was equally efficient for larger GlcNAc-peptide acceptor. As exemplified by a 34-mer peptide acceptor, GlcNAc-C34 (**14**), its Endo-A catalyzed reaction with the tetrasaccharide oxazoline **11** proceeded smoothly under a mild condition (phosphate buffer, pH 6.5, 37 °C) to give the glycopeptide **15** that carries a core N-linked pentasaccharide Man₃GlcNAc₂. Again, although the tetrasaccharide **11** was an excellent substrate for ENGases for transglycosylation, the resulting glycopeptide **15** was only slowly hydrolyzed by Endo-A (38). The huge difference in enzymatic reaction rates between the transition state analog substrate and the ground-state substrate (the product) suggests that the synthetic oligosaccharide oxazolines are kinetically more favorable for transglycosylation than the ground-state substrates such as natural N-glycans and N-glycopeptides. Thus, the employment of sugar oxazolines for ENGase-catalyzed transglycosylation not only expanded the substrate availability, but also led to substantially enhanced yield of the products, allowing a highly efficient, convergent synthesis of even large glycopeptides.

Modified Sugar Oxazoline as Substrates

In order to evaluate whether ENGases can recognize modified sugar oxazolines as substrates for transglycosylation, an array of sugar oxazolines, which differ in the modification or configurations of selected hydroxyl groups were prepared (55) (Figure 4). For the purpose, an alternative synthetic route was developed that started with a readily available disaccharide (cellobiose) derivative and took advantage of the existing β -1,4-glycosidic linkage in cellobiose for the construction of the core Man β 1,4-GlcN linkage through selective functional group transformations (55).

Using a small GlcNAc-tripeptide, Asn(GlcNAc)-Ile-Thr (**21**) as the acceptor, the Endo-A catalyzed reactions with various sugar oxazolines were examined in a phosphate buffer (pH 6.5) using an excess of the respective sugar oxazolines (donor: acceptor = 3:1). The di- and tetra-saccharide oxazolines **6** and **11** were tested before and were shown to be efficient substrates (38). The new tri-saccharide oxazoline **16**, with an additional mannose on the 6'-position of the core disaccharide, was also found to be an efficient substrate by Endo-A for the transglycosylation to form the glycopeptide **22** in a satisfactory yield. Interestingly, the modified oxazoline **17**, which has an aromatic substituent attached at the 6'-position, could be still recognized by Endo-A as a substrate to



Scheme 2. *ENGase*-catalyzed transglycosylation with synthetic sugar oxazolines

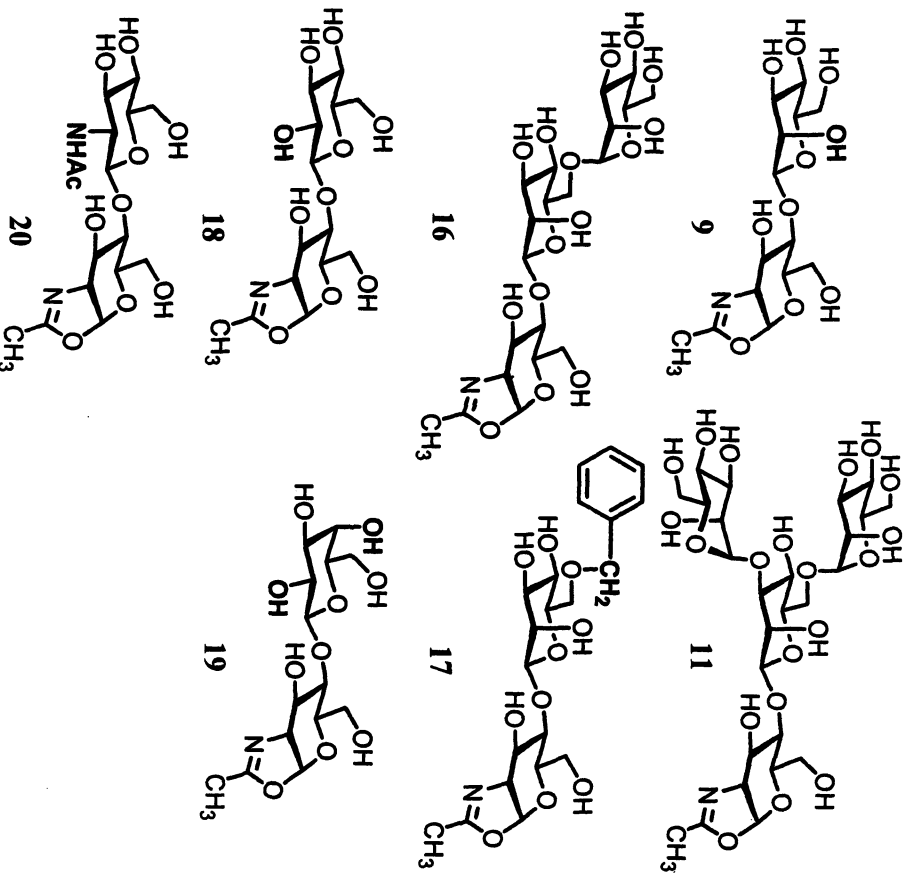
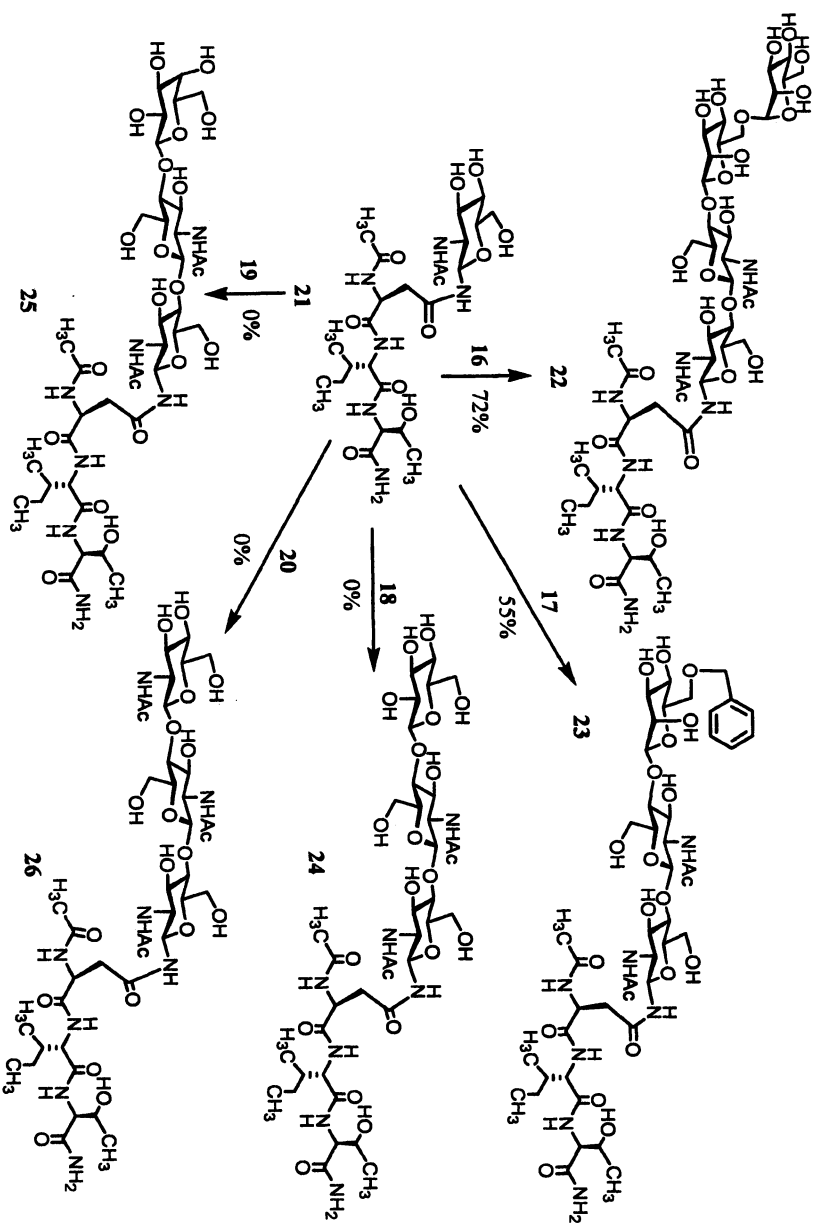
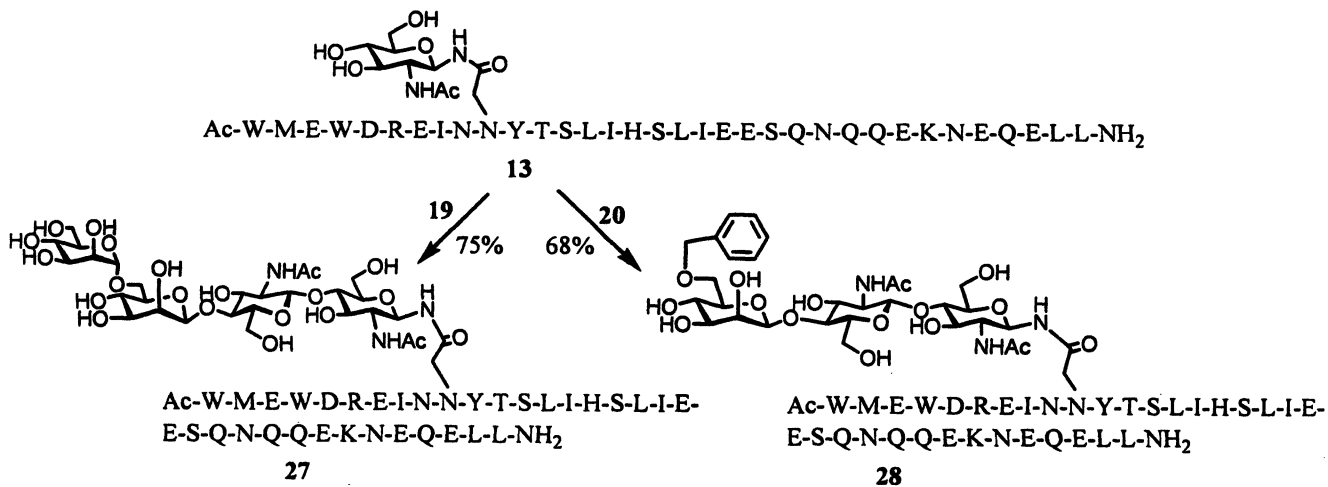


Figure 4. Structures of modified oligosaccharide oxazoline derivatives

form the modified glycopeptide **23**. In contrast, the change of the configurations at the 2' and/or 4'-hydroxyl groups (oxazolines **18** and **19**), or modification at the C-2' position (oxazoline **20**), all resulted in a total loss of substrate activity to Endo A (Scheme 3). The results suggest that the Man β 1,4GlcNAc oxazoline moiety is the minimum structure recognized by the enzyme Endo-A for the transglycosylation. But the enzyme could tolerate some modifications on the 3' and 6'-hydroxyl groups of the core disaccharide, implicating an expanded scope of substrates for transglycosylation. An immediate application of this finding is the synthesis of homogeneous glycopeptides with tags on the oligosaccharide moiety. An example was the synthesis of novel glycoforms of C34, a potent HIV-1 inhibitor (**55**) (Scheme 4).



Scheme 3. Endo-A-catalyzed transglycosylation with different oxazolines



Scheme 4. Endo-A catalyzed synthesis of modified glycopeptide C34

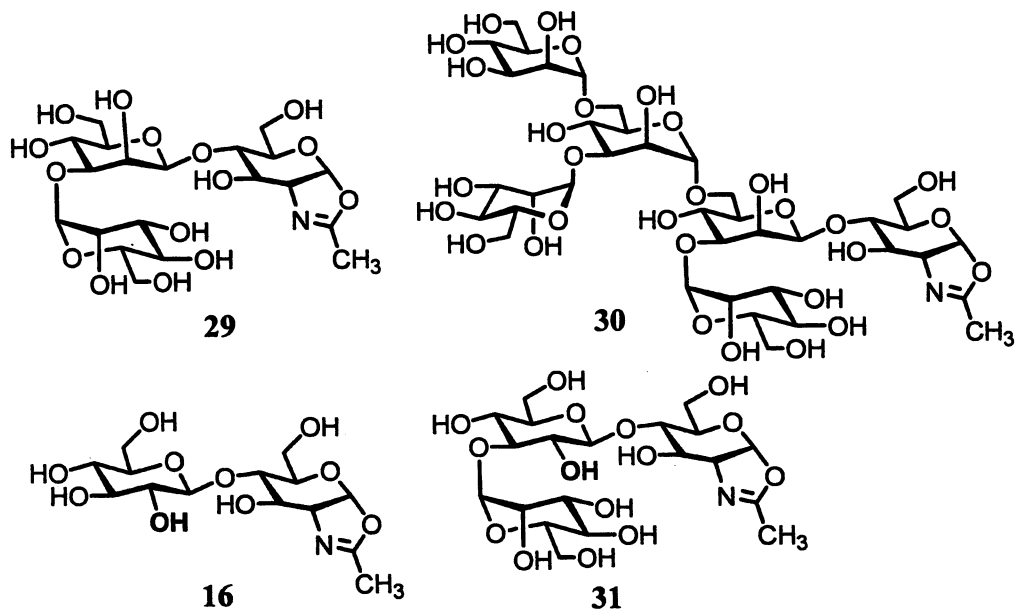


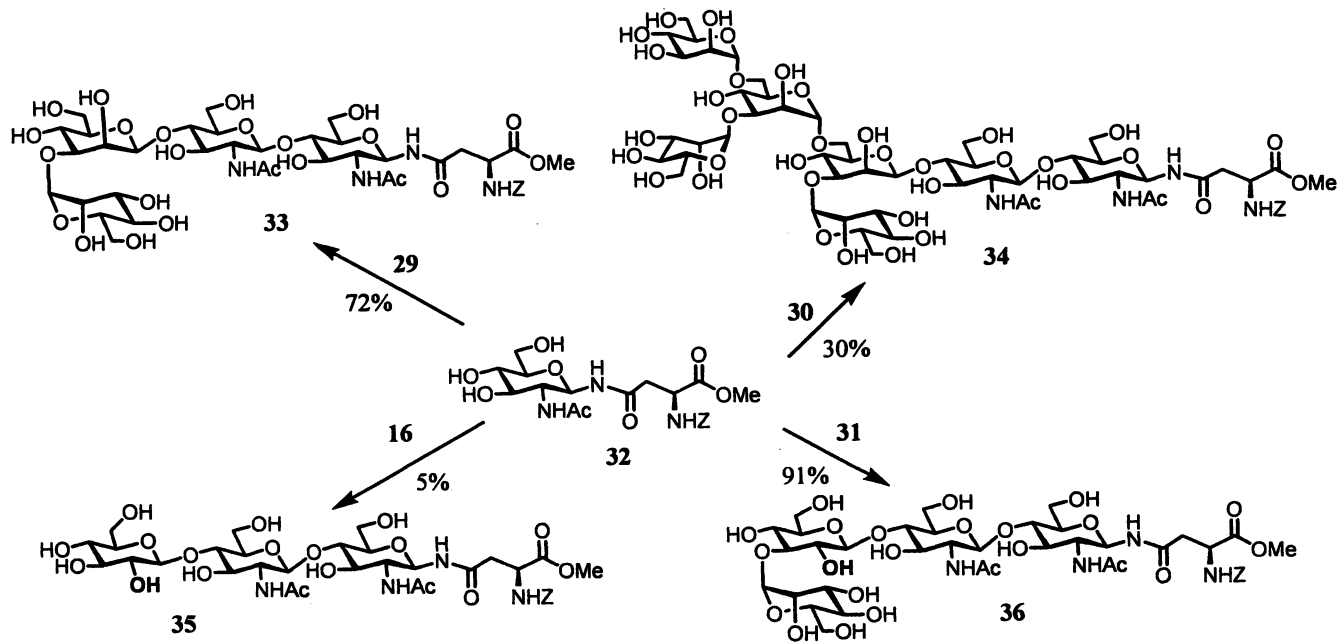
Figure 5. Structures of additional synthetic sugar oxazoline derivatives

Independently, Fairbanks and co-workers have prepared another set of sugar oxazolines (Figure 5) and examined their feasibility for Endo-M catalyzed transglycosylation (40,56). The tri- and hexasaccharide oxazolines (29 and 30) corresponding to the tri- and hexasaccharide core N-glycans (Man₂GlcNAc and Man₅GlcNAc), respectively, were synthesized (56). In addition, the *glc*-isomer (31) of the trisaccharide 29 was also synthesized (Figure 5).

The Endo-M catalyzed transglycosylation was tested using a GlcNAc-Asn derivative 32 as an acceptor (Scheme 5). As expected, the trisaccharide oxazoline 29 was an efficient substrate for Endo-M and its reaction with acceptor 32 gave the transglycosylation product 33 in a satisfactory yield. However, the larger sugar oxazoline 30 (Man₅GlcNAc-oxazoline) turned out to afford a poor yield, probably because of the rapid hydrolysis of the product by Endo-M. This study indicates the limitation of the method for synthesizing glycopeptides containing more extended natural oligosaccharides, when the product becomes an excellent substrate for enzymatic hydrolysis. On the other hand, it was observed that the Endo-M could tolerate the epimerization of the C-2 configuration of the β -mannose in the core disaccharide oxazoline (40). For example, the *glc*-containing disaccharide oxazoline 16 was able to serve as a substrate of Endo-M for transglycosylation to form the product 35, albeit at only 5% yield. This is in contrast to Endo-A, which could not take compound 16 as a substrate as previously reported (55). Very interestingly, when the disaccharide 16 was extended to a trisaccharide derivative with an additional α -mannosyl residue attached to the 3'-position, the resulting trisaccharide oxazoline 31 became an excellent substrate for Endo-M, and the transglycosylation with acceptor 32 gave a 91% yield of the product 36. It was also observed that the *glc*-containing glycopeptide products 35 and 36 could not be hydrolyzed by Endo-M. These results again demonstrated that the sugar oxazolines, as highly activated transition state mimic, could tolerate certain modifications without loss of their ability to serve as donor substrates for enzymatic transglycosylation, yet the resulting glycopeptide products could become resistant to hydrolysis due to the slightly structural modifications. The huge difference in reactivity of the donor substrates and the resulting glycopeptides thus allows the accumulation of the transglycosylation products.

ENGase-Catalyzed Double Glycosylation

The chemoenzymatic approach was also successfully applied to the synthesis of glycopeptides carrying two N-glycans (57). This was exemplified by the highly convergent synthesis of a large HIV-1 gp120 fragment, a 47-mer V3 domain glycopeptide carrying two core N-linked pentasaccharide (Scheme 6). First, the 47-mer peptide that contains two GlcNAc moieties was prepared by an Fmoc-based solid phase peptide synthesis. After retrieval from the resin, deprotection, and subsequent cyclization, the large GlcNAc-containing peptide



Scheme 5. Endo-M catalyzed transglycosylation with synthetic oxazolines

37 was purified by reverse-phase HPLC. The Endo-A catalyzed trans-glycosylation between the acceptor **37** and the Man₃GlcNAc-oxazoline (**11**) (**38**) was examined with an excess sugar oxazoline (donor/acceptor 3:1) and the reaction was monitored by RP-HPLC. It was found that the enzymatic reaction proceeded very efficiently to give a 86% yield of the desired glycopeptide **38**, which carries two *N*-glycans (Scheme 6). The product was characterized by ESI-MS and enzymatic transformations. The high-yield, simultaneous enzymatic double glycosylation with the sugar oxazolines on the large peptide was a surprise to us, as the addition of two *N*-glycans to the sterically-hindered GlcNAc residues in the cyclic peptide seemed to be difficult at a glance. The synthetic V3 glycopeptides were used for probing the effects of glycosylation on the properties of the V3 domain. It was observed that glycosylation could dramatically increase the amount of α -helical structure compared to the nonglycosylated V3 domain in a mixed water/TFE solvent. It was also found that the attachment of the core pentasaccharides could afford protection of the polypeptide domain against protease digestion (**57**). This study provides a fine example of how glycosylation could affect the structure and properties of a polypeptide domain. The ability to synthesize large HIV-1 glycopeptides carrying multiple *N*-glycans is key to the search for a glycopeptide-based HIV-1 vaccine (**58**).

The apparent advantage of the chemoenzymatic method is the single-step, highly efficient ligation of a pre-assembled oligosaccharide and a free polypeptide moiety by an enzyme under mild conditions, without the need of any protecting groups. Particularly appealing of the chemoenzymatic approach are the highly convergent nature of the ligation and the totally independent synthetic manipulations of the oligosaccharide and polypeptide portions. In contrast, the conventional glycopeptide synthesis, either the use of pre-formed glycosyl-amino acid building blocks in a stepwise solid phase peptide synthesis or the convergent coupling of a pre-assembled oligosaccharide with a fully protected polypeptide containing only free aspartic acid residues, all requires final deprotection of the protected polypeptide with a strong acid, e.g., TFA. This in many cases causes problems because the *O*-glycosidic linkages in the oligosaccharides are labile to strong acidic conditions. This is particularly true when some most acid-sensitive linkages such as the α -linked fucose and sialic acid moieties are involved. Thus, the convergent ENGase-catalyzed ligation, which uses pre-assembled oligosaccharides and free polypeptides as substrates, may provide an ultimate solution to the long-standing problem of "incompatibility" of protecting group manipulations in glycopeptide synthesis.

Extension of the Chemoenzymatic Approach to Glycosylation Remodeling of Glycoproteins

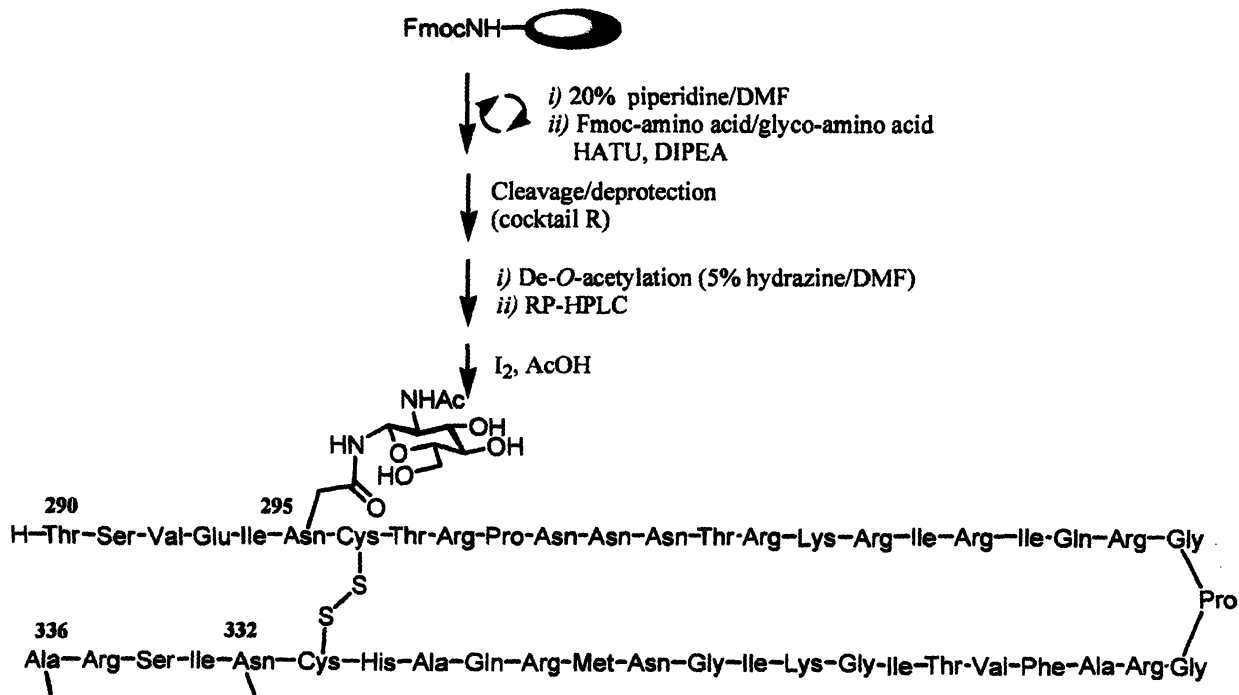
The potential of the chemoenzymatic approach for glycoprotein synthesis was exemplified by the successful glycosylation remodeling of bovine

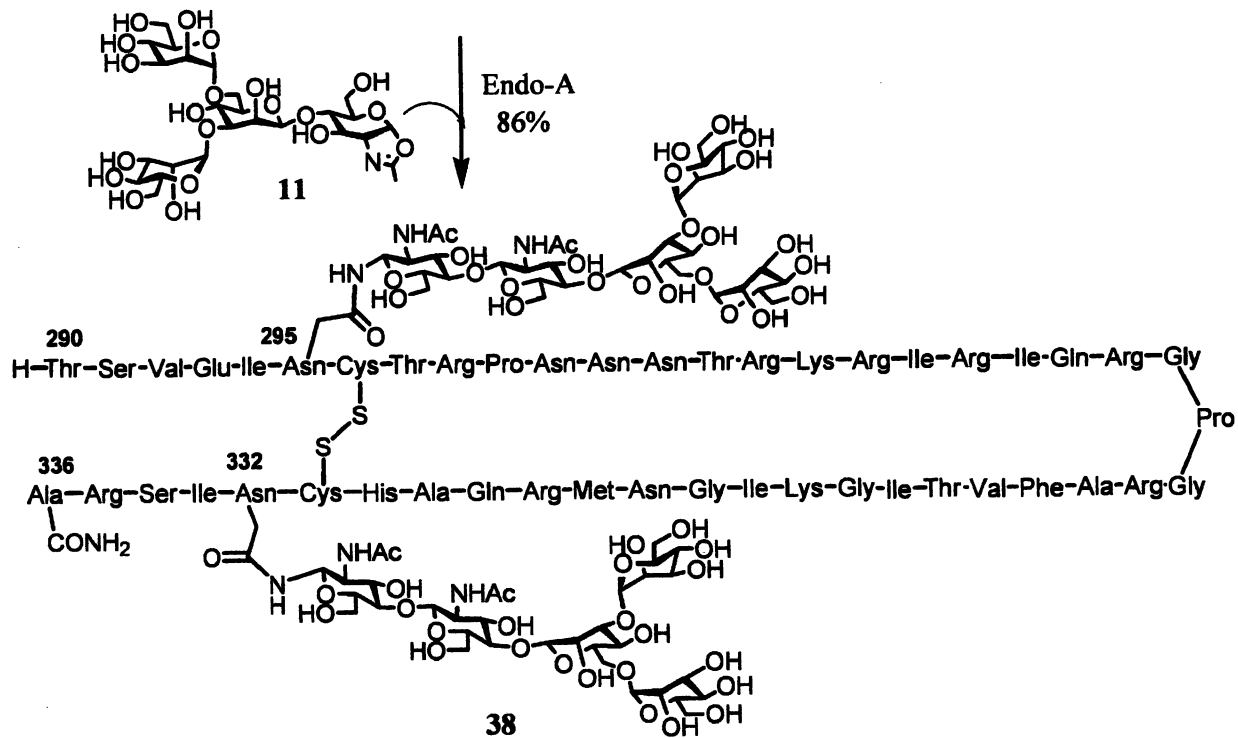
ribonuclease B (39). Ribonuclease B is a small glycoprotein that consists of 124 amino acids with a single N-glycosylation site at Asn-34, but native ribonuclease B is a mixture of several glycoforms of the high-mannose type N-glycan ranging from Man₅ to Man₉. Takegawa and co-workers have previously attempted to transfer a Man₆GlcNAc moiety to the GlcNAc-RNase B by Endo-A using the natural Man₆GlcNAc₂Asn as the donor substrate (59), but the yield was estimated to be only about 5% by SDS-PAGE analysis. To examine the feasibility of the ENGase-catalyzed transglycosylation with various oligosaccharide sugar oxazolines for glycoprotein remodeling, a novel hexasaccharide oxazoline **42** (Gal₂Man₃GlcNAc-oxazoline) was synthesized, which has two galactose residues β -1,4-linked to the terminal mannose residues in the core Man₃GlcNAc (Scheme 7). This hexasaccharide derivative might be regarded as a mimic of a bi-antennary complex type *N*-glycan lacking the internal GlcNAc moieties. This unnatural hexasaccharide oxazoline was expected to serve as a substrate for ENGases as it reserves the core Man₃GlcNAc structure and the modification on the outer mannosyl residues might not abolish its recognition by the enzyme.

The acceptor GlcNAc-RB (**43**) was easily prepared by treatment of commercially available bovine ribonuclease B with Endo-H (an endoglycosidase that cleaves high-mannose type *N*-glycans at the chitobiose core) to remove the heterogeneous *N*-glycans but leave only the inner most GlcNAc at the Asn-34 site. Transglycosylation with GlcNAc-RB was examined by using synthetic sugar oxazolines **11** and **42** (Scheme 8). Indeed, it was found that the transglycosylation with the two oxazolines proceeded very fast in the presence of Endo-A under mild conditions (phosphate buffer, pH 6.5, 23 °C), giving the desired homogeneous glycoproteins **44** and **45**, respectively, in excellent yields. The identity of the products was confirmed by their ESI-MS and the NMR analysis of the intact oligosaccharide from the pronase digestion of the products (39). It was also observed that the glycoprotein **45**, once formed, was completely resistant to the hydrolysis by Endo-A. This could be explained by the fact that Endo-A hydrolyzes only high-mannose type natural *N*-glycans, but glycoprotein **45** carries an unnatural *N*-glycan mimicking the complex type sugar chain. The results again attested to the great potential of the novel chemoenzymatic approach for constructing tailor-made homogeneous glycoproteins carrying well defined oligosaccharides.

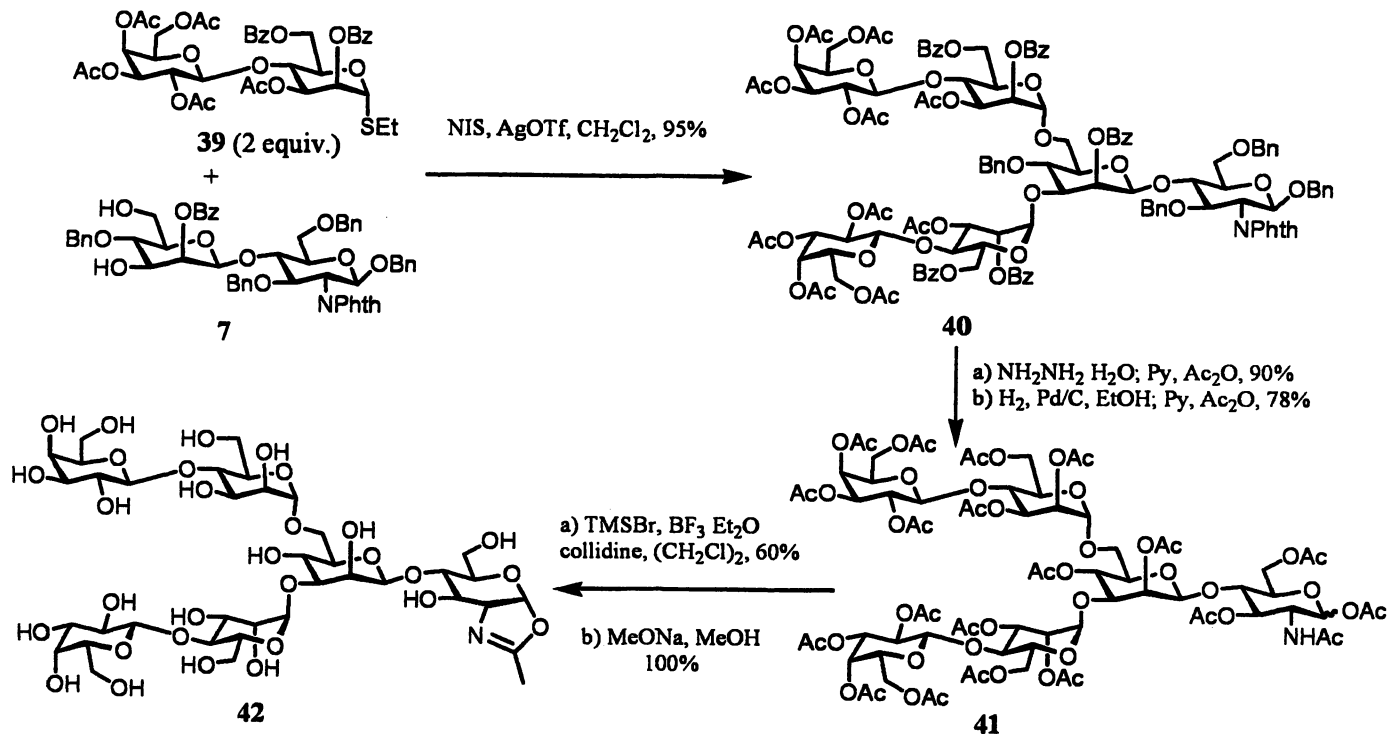
Conclusion

Homogeneous glycopeptides are indispensable tools for various structural and biological studies of glycoproteins. However, the construction of large, homogeneous glycopeptides is still a challenging task. The endoglycosidase-catalyzed transglycosylation provides a novel method for the ligation of a pre-formed oligosaccharide to a polypeptide in a regio- and stereospecific manner,

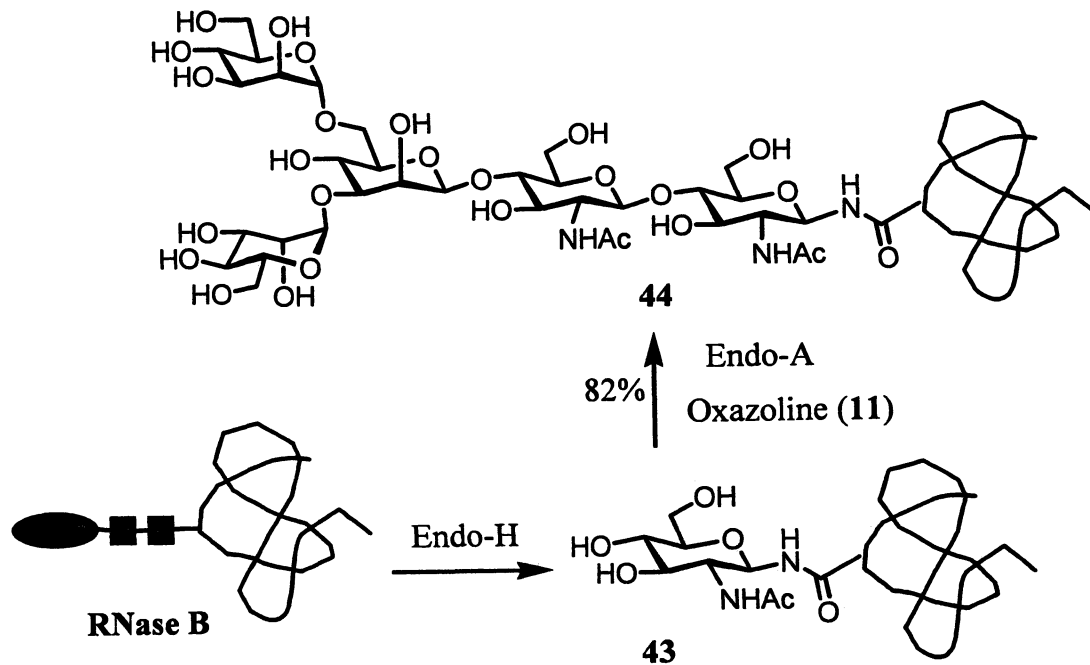




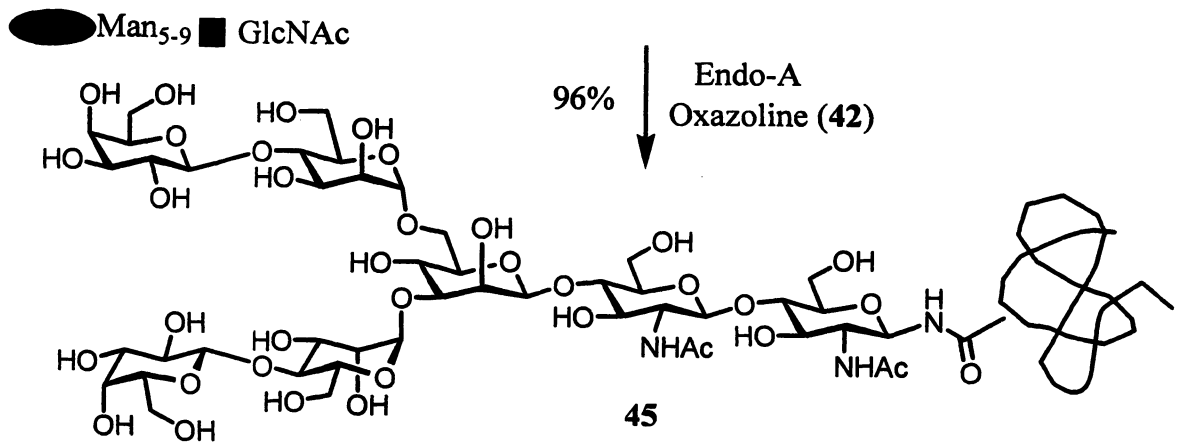
Scheme 6. Synthesis of cyclic HIV-1 V3 glycopeptides carrying two N-glycans



Scheme 7. Synthesis of an unnatural hexasaccharide oxazoline



Scheme 8. Glycosylation remodeling of bovine ribonuclease B. Continued on next page.



Scheme 8. Continued.

without the need of any protecting groups. The most notable feature of the chemoenzymatic approach is the highly convergent nature of the ligation and the totally independent synthetic manipulations of the oligosaccharide and polypeptide/protein portions. The use of synthetic sugar oxazolines, the presumed transition-state analogs, as donor substrates not only expands the availability of donors, but also leads to a substantial enhancement of the transglycosylation yield, allowing an efficient synthesis of both natural and unnatural glycopeptides and even glycoproteins. The chemical synthesis of a given oligosaccharide oxazoline should be within the reach of an established carbohydrate group, while the GlcNAc-polypeptide acceptor could be synthesized by the common solid-phase peptide synthesis using a GlcNAc-Asn building block without extra difficulty. In addition, a given GlcNAc-protein could be prepared by several approaches, including a) Endo-H or Endo-F treatment of recombinant glycoproteins to remove the N-glycans; b) total protein synthesis via native chemical ligation or expressed protein ligation by incorporating a GlcNAc-tag at a pre-determined site during the synthesis (60,61); and c) overproduction of a GlcNAc-containing protein in *E. coli* through the novel *in vivo* suppressor tRNA technology (62). As to the limitations of the chemoenzymatic method, an unsolved problem is still the possibility of product hydrolysis, particularly when the natural glycopeptide formed is a good substrate for the enzymatic hydrolysis. Future studies should be directed to structural and mechanistic studies on various endoglycosidases, so that novel endoglycosidase mutants can be discovered that possess enhanced transglycosylation activity but lack the hydrolytic activity, as exemplified for some glycosynthases (63,64). A broader application of the novel chemoenzymatic approach for complex N-glycopeptide synthesis and for glycoprotein engineering is anticipated in the future.

References

1. Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim. Biophys. Acta* **1999**, *1473*, 4-8.
2. Varki, A. *Glycobiology* **1993**, *3*, 97-130.
3. Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683-720.
4. Helenius, A.; Aebi, M. *Science* **2001**, *291*, 2364-2369.
5. Bertozzi, C. R.; Kiessling, L. L. *Science* **2001**, *291*, 2357-2364.
6. Stanley, P. *Glycobiology* **1992**, *2*, 99-107.
7. Gerngross, T. U. *Nat. Biotechnol.* **2004**, *22*, 1409-1414.
8. Koeller, K. M.; Wong, C. H. *Nat. Biotechnol.* **2000**, *18*, 835-841.
9. Seitz, O. *ChemBioChem* **2000**, *1*, 214-246.
10. Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H. *Chem. Rev.* **2000**, *100*, 4495-4538.
11. Hang, H. C.; Bertozzi, C. R. *Acc. Chem. Res.* **2001**, *34*, 727-736.
12. Davis, B. G. *Chem. Rev.* **2002**, *102*, 579-601.

13. Grogan, M. J.; Pratt, M. R.; Marcaurelle, L. A.; Bertozzi, C. R. *Annu. Rev. Biochem.* **2002**, *71*, 593-634.
14. Wong, C. H. *J. Org. Chem.* **2005**, *70*, 4219-4225.
15. Guo, Z.; Shao, N. *Med. Res. Rev.* **2005**, *25*, 655-678.
16. Liu, L.; Bennett, C. S.; Wong, C. H. *Chem. Commun.* **2006**, 21-33.
17. Buskas, T.; Ingale, S.; Boons, G. J. *Glycobiology* **2006**, *16*, 113R-136R.
18. Suzuki, T.; Yano, K.; Sugimoto, S.; Kitajima, K.; Lennarz, W. J.; Inoue, S.; Inoue, Y.; Emori, Y. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 9691-9696.
19. Takegawa, K.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. *Biochem. Int.* **1991**, *25*, 829-835.
20. Takegawa, K.; Yamaguchi, S.; Kondo, A.; Iwamoto, H.; Nakoshi, M.; Kato, I.; Iwahara, S. *Biochem. Int.* **1991**, *24*, 849-855.
21. Kadowaki, S.; Yamamoto, K.; Fujisaki, M.; Izumi, K.; Tochikura, T.; Yokoyama, T. *Agric. Biol. Chem.* **1990**, *54*, 97-106.
22. Kadowaki, S.; Yamamoto, K.; Fujisaki, M.; Tochikura, T. *J. Biochem. (Tokyo)* **1991**, *110*, 17-21.
23. Yamamoto, K.; Kadowaki, S.; Fujisaki, M.; Kumagai, H.; Tochikura, T. *Biosci. Biotechnol. Biochem.* **1994**, *58*, 72-77.
24. Kato, T.; Fujita, K.; Takeuchi, M.; Kobayashi, K.; Natsuka, S.; Ikura, K.; Kumagai, H.; Yamamoto, K. *Glycobiology* **2002**, *12*, 581-587.
25. Yamamoto, K.; Kadowaki, S.; Watanabe, J.; Kumagai, H. *Biochem. Biophys. Res. Commun.* **1994**, *203*, 244-252.
26. Haneda, K.; Inazu, T.; Yamamoto, K.; Kumagai, H.; Nakahara, Y.; Kobata, A. *Carbohydr. Res.* **1996**, *292*, 61-70.
27. Wang, L. X.; Tang, M.; Suzuki, T.; Kitajima, K.; Inoue, Y.; Inoue, S.; Fan, J. Q.; Lee, Y. C. *J. Am. Chem. Soc.* **1997**, *119*, 11137-11146.
28. Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, I.; Kawakami, T.; Aimoto, S.; Yamamoto, K.; Inazu, T. *J. Am. Chem. Soc.* **1999**, *121*, 284-290.
29. Wang, L. X.; Fan, J. Q.; Lee, Y. C. *Tetrahedron Lett.* **1996**, *37*, 1975-1978.
30. Deras, I. L.; Takegawa, K.; Kondo, A.; Kato, I.; Lee, Y. C. *Bioorg. Med. Chem. Lett.* **1998**, *8*, 1763-1766.
31. O'Connor, S. E.; Pohlmann, J.; Imperiali, B.; Saskiawan, I.; Yamamoto, K. *J. Am. Chem. Soc.* **2001**, *123*, 6187-6188.
32. Haneda, K.; Inazu, T.; Mizuno, M.; Iguchi, R.; Tanabe, H.; Fujimori, K.; Yamamoto, K.; Kumagai, H.; Tsumori, K.; Munekata, E. *Biochim. Biophys. Acta* **2001**, *1526*, 242-248.
33. Wang, L. X.; Singh, S.; Ni, J. In *Synthesis of Carbohydrates through Biotechnology*; Wang, P. G., Ichikawa, Y., Eds.; American Chemical Society: Washington, D. C., 2004, p 73-92.
34. Singh, S.; Ni, J.; Wang, L. X. *Bioorg. Med. Chem. Lett.* **2003**, *13*, 327-330.
35. Wang, L. X.; Song, H.; Liu, S.; Lu, H.; Jiang, S.; Ni, J.; Li, H. *ChemBioChem* **2005**, *6*, 1068-1074.

36. Li, H.; Singh, S.; Zeng, Y.; Song, H.; Wang, L. X. *Bioorg. Med. Chem. Lett.* **2005**, *15*, 895-898.
37. Akaike, E.; Tsutsumida, M.; Osumi, K.; Fujita, M.; Yamanoi, T.; Yamamoto, K.; Fujita, K. *Carbohydr. Res.* **2004**, *339*, 719-722.
38. Li, B.; Zeng, Y.; Hauser, S.; Song, H.; Wang, L. X. *J. Am. Chem. Soc.* **2005**, *127*, 9692-9693.
39. Li, B.; Song, H.; Hauser, S.; Wang, L. X. *Org. Lett.* **2006**, *8*, 3081-3084.
40. Rising, T. W.; Claridge, T. D.; Moir, J. W.; Fairbanks, A. J. *ChemBioChem* **2006**, *7*, 1177-1180.
41. Brameld, K. A.; Shrader, W. D.; Imperiali, B.; Goddard, W. A., 3rd *J. Mol. Biol.* **1998**, *280*, 913-923.
42. Terwisscha van Scheltinga, A. C.; Armand, S.; Kalk, K. H.; Isogai, A.; Henrissat, B.; Dijkstra, B. W. *Biochemistry* **1995**, *34*, 15619-15623.
43. Tews, I.; Terwisscha van Scheltinga, A. C.; Perrakis, A.; Wilson, K. S.; Dijkstra, B. W. *J. Am. Chem. Soc.* **1997**, *119*, 7954-7959.
44. Mark, B. L.; Vocadlo, D. J.; Knapp, S.; Triggs-Raine, B. L.; Withers, S. G.; James, M. N. *J. Biol. Chem.* **2001**, *276*, 10330-10337.
45. Williams, S. J.; Mark, B. L.; Vocadlo, D. J.; James, M. N.; Withers, S. G. *J. Biol. Chem.* **2002**, *277*, 40055-40065.
46. Rao, V.; Guan, C.; Van Roey, P. *Structure* **1995**, *3*, 449-457.
47. Waddling, C. A.; Plummer, T. H., Jr.; Tarentino, A. L.; Van Roey, P. *Biochemistry* **2000**, *39*, 7878-7885.
48. Fujita, M.; Shoda, S.; Haneda, K.; Inazu, T.; Takegawa, K.; Yamamoto, K. *Biochim. Biophys. Acta* **2001**, *1528*, 9-14.
49. Kobayashi, S.; Kiyosada, T.; Shoda, S. *J. Am. Chem. Soc.* **1996**, *118*, 13113-13114.
50. Kobayashi, S.; Morii, H.; Itoh, R.; Kimura, S.; Ohmae, M. *J. Am. Chem. Soc.* **2001**, *123*, 11825-11826.
51. Kobayashi, S.; Fujikawa, S.; Ohmae, M. *J. Am. Chem. Soc.* **2003**, *125*, 14357-14369.
52. Ochiai, H.; Ohmae, M.; Kobayashi, S. *Carbohydr. Res.* **2004**, *339*, 2769-2788.
53. Kobayashi, S.; Ohmae, M.; Ochiai, H.; Fujikawa, S. *Chem. Eur. J.* **2006**, *12*, 5962-5971.
54. Ochiai, H.; Fujikawa, S. I.; Ohmae, M.; Kobayashi, S. *Biomacromolecules* **2007**.
55. Zeng, Y.; Wang, J.; Li, B.; Hauser, S.; Li, H.; Wang, L. X. *Chem. Eur. J.* **2006**, *12*, 3355-3364.
56. Rising, T. W.; Claridge, T. D.; Davies, N.; Gamblin, D. P.; Moir, J. W.; Fairbanks, A. J. *Carbohydr. Res.* **2006**, *341*, 1574-1596.
57. Li, H.; Li, B.; Song, H.; Breydo, L.; Baskakov, I. V.; Wang, L. X. *J. Org. Chem.* **2005**, *70*, 9990-9996.

58. Wang, L. X. *Curr. Opin. Drug Discov. Devel.* **2006**, *9*, 194-206.
59. Takegawa, K.; Tabuchi, M.; Yamaguchi, S.; Kondo, A.; Kato, I.; Iwahara, S. *J. Biol. Chem.* **1995**, *270*, 3094-3099.
60. Dawson, P. E.; Muir, T. W.; Clark-Lewis, I.; Kent, S. B. *Science* **1994**, *266*, 776-779.
61. Schwarzer, D.; Cole, P. A. *Curr. Opin. Chem. Biol.* **2005**, *9*, 561-569.
62. Zhang, Z.; Gildersleeve, J.; Yang, Y. Y.; Xu, R.; Loo, J. A.; Uryu, S.; Wong, C. H.; Schultz, P. G. *Science* **2004**, *303*, 371-373.
63. Perugino, G.; Trincone, A.; Rossi, M.; Moracci, M. *Trends Biotechnol.* **2004**, *22*, 31-37.
64. Hancock, S. M.; Vaughan, M. D.; Withers, S. G. *Curr. Opin. Chem. Biol.* **2006**, *10*, 509-519.

Chapter 7

Synthesis of Glycolipid Antigens

Suvarn S. Kulkarni* and Jacquelyn Gervay-Hague*

Department of Chemistry, University of California at Davis,
One Shields Avenue, Davis, CA 95616

Carbohydrates and glycoconjugates play important roles in complex life processes. Development of carbohydrate-based therapeutics is a major focus of our laboratory requiring the synthesis of glycolipid antigens for biological evaluation. Stereoselective formation of *O*- and *C*-glycosides is one of the central challenges in carbohydrate synthesis. In this pursuit, we employ glycosyl iodides as unique glycosyl donors to construct functionally enriched appendages on the anomeric center of the sugar moiety in a highly efficient and stereoselective manner. These constructs can be used to rapidly synthesize a panoply of glycoside analogs for biological studies. In this minireview, glycosyl iodide mediated syntheses of α -linked *O*-glycolipids such as KRN 7000 and BbGL2 as well *C*-glycolipid analogs are described.

Introduction

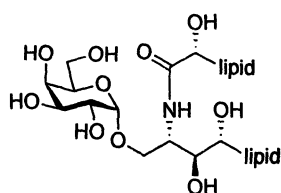
1.1 Discovery of an immunopotent α -linked galactolipid KRN 7000

Glycolipids play important roles in various disease states and immune response (1). For example, Agelasphins (AGL) (1), a group of novel galactosyl ceramides (GalCer) isolated from the marine sponge *Agelas mauritianus* (2,3) showed potent anti-tumor activity in mice. These glycolipids were the first reported galactosyl ceramides possessing unique α -glycosidic linkages. This discovery prompted syntheses of AGL analogs in order to identify compounds having better activity profiles and lower toxicity (4-6). Among the initial analogs synthesized, KRN7000 (2) (Figure 1), which lacks the hydroxyl group on the ceramide lipid chain, was found to be the most potent (5). Structure-activity studies also revealed that the α -linked galactose moiety was essential for anti-tumor activity (6). Further biomedical and clinical studies suggested that KRN7000 and its analogs held promise in treating other diseases such as hepatitis B (7), malaria (8), and diabetes (9). Extensive mechanistic studies indicated that the anti-tumor activity resulted from stimulation of specialized immune effector cells called Natural Killer T (NKT) cells in a CD1d-dependent manner (10).

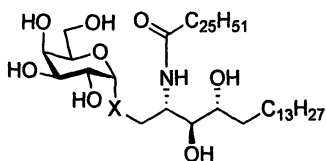
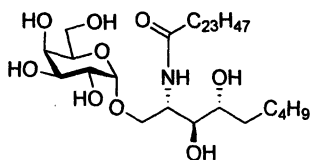
1.2 CD1d dependent mode of activation of NKT cells

CD1d is a member of the CD1 protein family that presents lipid antigens to NKT cells to activate the immune response. CD1 molecules function similarly to the evolutionarily related peptide-presenting proteins, major histocompatibility complex (MHC) molecules. Five members of the CD1 family including CD1a, CD1b, Cd1c, CD1d and CD1e are found in humans whereas only two homologs of CD1d (CD1d1 and CD1d2) are present in mice (11). The crystal structure of mouse CD1d1 (12), reveals that the lipid chains are buried in a hydrophobic pocket of the protein and the hydrophilic sugar head-groups are exposed for presentation to NKT cells. Antigen presenting cells expressing CD1d recognize KRN7000 and the binding complex interacts with the T-cell receptor (TCR) of NKT cells stimulating the release of cytokines that initiate two major types of response. One group of cytokines, consisting of interferon- γ (INF- γ) and interleukin-2 (IL-2), causes an inflammatory response, termed a T helper 1 (Th1) response whereas the other includes IL-4 and IL-10 that lead to an immunomodulatory or a T helper 2 (Th2) response. In general, proinflammatory Th1 responses are mostly responsible for cell-mediated immunity resulting in

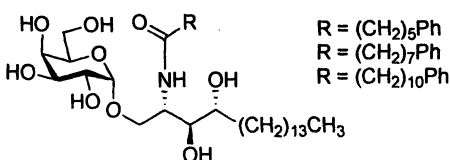
anti-virus, anti-bacterial, anti-tumor and adjuvant activity, while Th2 cytokines correlate with prevention of certain autoimmune diseases. Interestingly, different KRN7000 analogs selectively stimulate cytokine production. For example, OCH (3) (Figure 1), which has only 9 carbons in the base chain of ceramide, produces predominately IL-4 and exhibits greater efficacy than KRN7000 against the autoimmune disease experimental allergic encephalomyelitis (13-15). Whereas, the C-glycoside analog of KRN7000, cKRN7000 (4) upregulates IFN- γ and is 100 times more potent than KRN7000 in inhibiting tumor growth in mice (16-18). Similarly, an aryl containing fatty acid chain analog of α -GalCer (5), elicits enhanced IFN- γ secretion relative to IL-4, thus proving useful for adjuvant development (19).



1. Agelasphins

2. KRN7000, X = O
4. cKRN7000, X = CH₂

3. OCH

5. IFN- γ -selective analogsFigure 1. Various immunogenic α -GalCer analogs

1.3 Bacterial glycolipid ligands for NKT cell activation

α -Galactosyl ceramides are rarely found in mammalian cells; only a few have been identified in cancer cells or fetuses; most are derived from bacteria. It

is believed that CD1d recognition of α -GalCer is one pathway to activate immune response against bacterial infection. Recently, there has been interest in studying immunoresponsive glycolipids of bacterial origin. There is growing evidence that these glycolipids show similar structure-activity profiles to that of KRN7000 for CD1d restricted activation of NKT cells (20,21). For example, α -anomeric glycosphingolipids (e.g. GSL-1, 6) extracted from the gram negative bacterium *Sphingomonas paucimobilis* are able to stimulate both human and mouse NKT cells specifically and in a CD1d-dependent manner (22). Likewise, two major glycolipids have been isolated from *Borrelia burgdorferi*, the etiological agent of Lyme disease which is a multisystemic disorder that affects the skin, nervous system, heart and joints (23). The structures of these highly immunoreactive glycolipids have recently been elucidated as cholesteryl 6-O-acyl- β -D-galactopyranoside (BbGL1, 7) and 1,2-di-O-acyl-3-O- α -D-galactopyranosyl-*sn*-glycerol (BbGL2, 8, Figure 2); the major fatty acids were palmitate and oleate (24). Mouse and human NKT cells recognize α -linked O-glycosides BbGL2 but not the β -linked BbGL1. Recently, it was observed that NKT cell proliferation and cytokine production and the antigenic potency of BbGL2 was dependent on acyl chain length as well as degree of saturation (25). C-Glycoside analogs of BbGL2 (9) are therefore interesting targets (26).

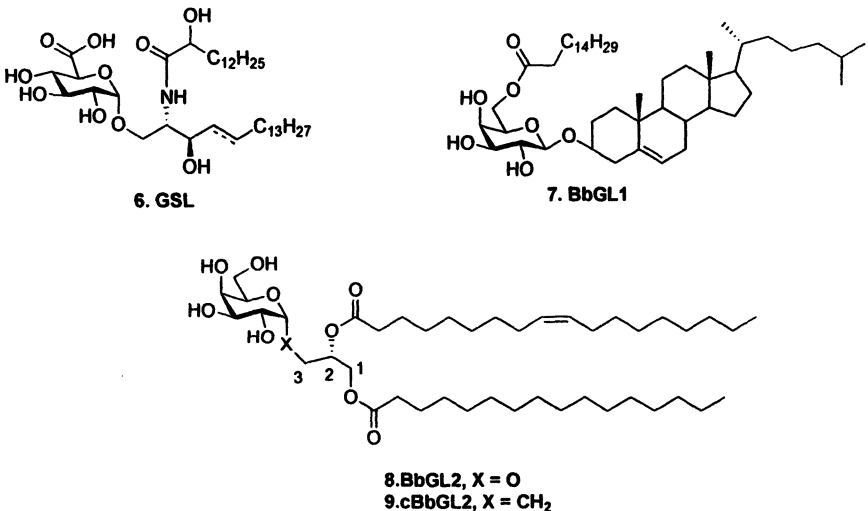


Figure 2. Bacterial glycolipids- NKT cells activators

Synthesis of α -Linked Glycolipids

2.1 Background

Numerous efforts have been invested in the synthesis of GalCer analogs to achieve acceptable purity and quantity for biological studies. However, stereoselective glycosylation of D-galactose derived donors with ceramide acceptors remains a central challenge. The glycosylation often proceeds in relatively low yields with poor α/β selectivity and purification of the desired α -isomer is arduous. Moreover, the preparation of acceptors and glycosyl donors involves multistep sequences lowering the overall synthetic efficiency. Commonly employed donors like fluorides (27-30) and trichloroacetimidates (29, 31-33) usually furnish α/β -mixtures with yields typically ranging from 30% to 60%. Other donors such as bromides (34), thiogalactosides (32) hemiacetals and phosphites (35) give similar results.

Over the past decade, our laboratory has been actively involved in developing methods for the preparation and applications of glycosyl iodide donors (36-49) including efficient α -stereoselective glycosylation protocols that work well with various sugars iodides (39-41). Recently, we have applied this methodology in the synthesis α -anomeric glycolipids.

2.2 α -Stereoselective glycosylation via 'In situ anomerization'

Glycosyl iodides were once believed to be too reactive and unstable to be useful for synthetic applications. Studies from our laboratory and others were instrumental in disproving this notion and have led to a revival in the chemistry of glycosyl iodides, over the past decade. Glycosyl iodides are not only 'useful' but also possess unique reactivity profiles, which can be exploited for the construction of challenging C-O and C-C linkages in a highly stereoselective manner. Installation of 1,2-*cis*-O-glycosidic linkages in D-Gal and D-Glc is a challenge as most glycosylations proceed through formation of an oxonium ion with concomitant nucleophilic attack of the acceptor in S_N1 fashion. This is true with most of the commonly used glycosyl donors. However, using glycosyl iodides, α -stereoselectivity can be achieved through Lemieux like in situ anomerization (50) of the preferentially generated α -iodide to a more reactive β -iodide, which subsequently undergoes S_N2 -like displacement to give the α -glycoside, exclusively (Scheme 1). These reactions are orders of magnitude faster than the corresponding glycosyl bromides, quantitatively affording only α -anomer. These results encouraged us to extrapolate the process to glycolipid syntheses.

2.3 First generation synthesis of α -GalCer analogs

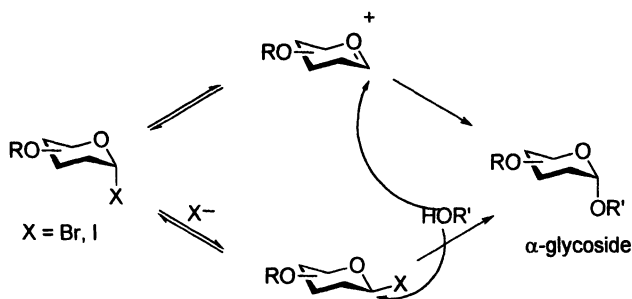
Our original approach to the synthesis of α -GalCer (*51*) involved a combination of highly reactive galactosyl iodide donors with electron rich ceramide acceptors. Accordingly, acceptor (*2S,3S*)-2-azido-3-*p*-methoxybenzyl sphingosine (**11**) was prepared from commercially available sphingosine **10** in four high yielding steps – amine to azide conversion, tritylation of the primary alcohol, masking the secondary alcohol with a *p*-methoxybenzyl (PMB) group, and deprotection of the trityl ether (Scheme 2). In a similar way (*2S,3S,4R*)-2-azido-3,4-*p*-methoxybenzyl phytosphingosine (**13**) was synthesized from **12**.

Strategic placement of the electron donating PMB group increased the reactivity of the acceptor and the conversion of amine to azide obviated the unwanted intramolecular hydrogen bonding interactions between the amine and primary alcohol. Coupling of acceptors **11** or **13** with *per-O*-benzylated galactosyl iodide **14** using tetrabutylammonium iodide as a promoter afforded only the α -*O*-glycosides **15** and **16** with yields over 90% (Scheme 3). Subsequent conversion of the azido group to an amine, followed by fatty acid coupling and hydrogenation afforded pure 4-desoxy-KRN7000 analog (**17**) and KRN7000 analog **18**, respectively (Scheme 4).

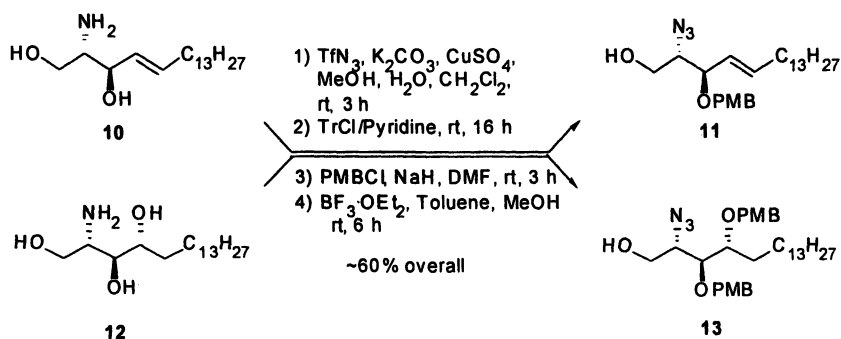
2.4 Streamlined synthesis of α -GalCer and BbGL2 analogs

Although our first generation synthesis offered advantages over existing technology it required several steps to ready the donor and acceptor for coupling. *Per-O*-benzylated iodide was generated from the corresponding anomeric acetate, which in turn required three steps from *D*-galactose; preparation of the acceptor required four steps, and subsequent deprotection of the coupled products needed an extra three steps. Furthermore, unsaturation in the lipid chain could not be retained under the hydrogenation conditions. We imagined that an attractive alternative would be to employ fully functionalized ceramide acceptors, but we were mindful of the fact that direct coupling of unprotected ceramide acceptors had been reported to be inefficient (*52*). Nevertheless, earlier studies in our laboratory indicated that silylated glycosyl iodide donors were more reactive than benzylated donors suggesting to us that increased efficiency and simplicity could be achieved by using *per-O*-trimethylsilylated (*O*-TMS) glycosyl iodides in a one-pot endeavor (Scheme 5).

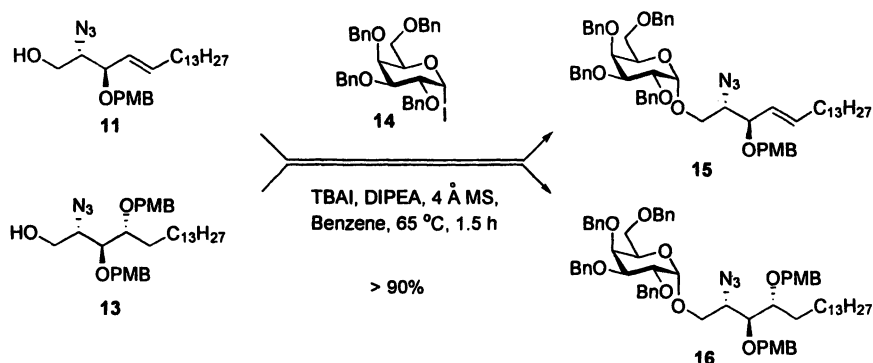
Indeed, under these conditions (Scheme 6), *per-O*-TMS galactosyl iodide **23**, generated in situ from *per-O*-TMS galactopyranoside **22**, underwent highly regioselective α -glycosidation in the presence of TBAI with fully functionalized glyceride **24-25** and ceramide acceptors **26-27** producing α -linked glycolipids



Scheme 1. In situ anomerization



Scheme 2. Preparation of ceramide acceptors



Scheme 3. α -Stereoselective glycosylation

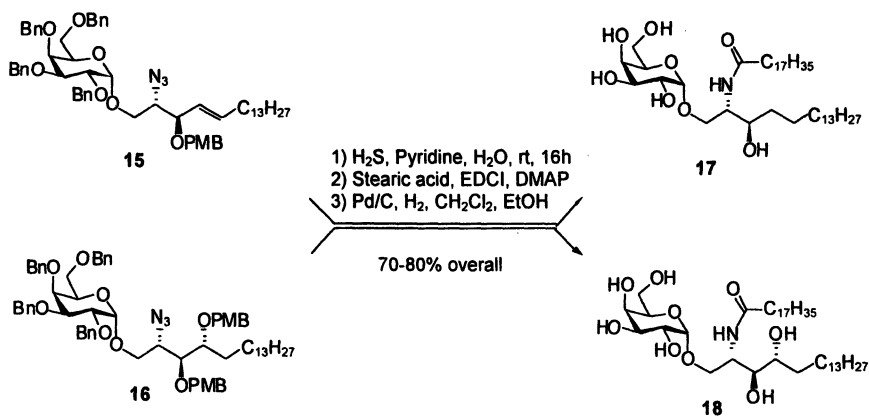
(53). Treatment of the crude product in the same pot with acidic resin in methanol afforded BbGL2 (**8**), its analog **29**, and GalCer analogs **28** and **18** in high yields and stereoselectivity (Figure 3). This mild one-pot protocol allowed the synthesis of pure α -anomeric glycolipids bearing sensitive functionalities including esters, amides and double bonds. Recent SAR studies on BbGL2 analogs indicate that **29** with a reversed arrangement of fatty acid chains exhibits the most potent immunogenic activity (26). Compound **30** is a saturated analog of **8** prepared by hydrogenation. Compounds **28** and **18** are analogs of KRN7000 with shorter ceramide lipids and differences in unsaturation. Microwave radiation has proven useful in these reactions when utilizing lipids with limited solubility such as acceptor **27**.

Synthesis of the C-glycoside analogue of BbGL2

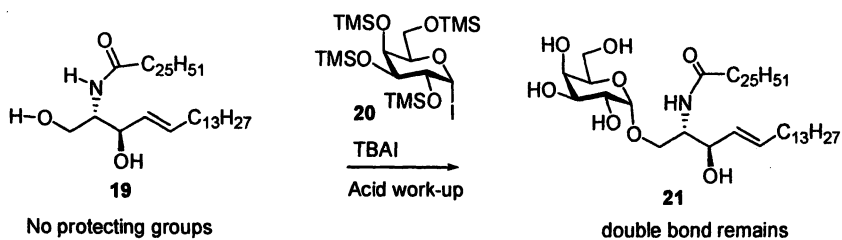
A useful extension of the in situ anomerization process involved employment of C-nucleophiles such as vinyl magnesium bromides (26). Grignard reactions of benzylated α -D-galactosyl iodides with vinyl magnesium bromide generated an α/β mixture slightly favoring the α -isomer (2.5:1). Stereoselectivity was significantly increased when the reaction was carried out under in situ anomerization conditions using TBAI in toluene at reflux, in which case the α -isomer formed in high yields (79%, $\alpha/\beta = 12/1$) (Scheme 7). The methodology proved useful in the first synthesis of an α -linked C-glycolipid corresponding to the immunoreactive bacterial glycolipid BbGL2 (Scheme 8).

Conclusions

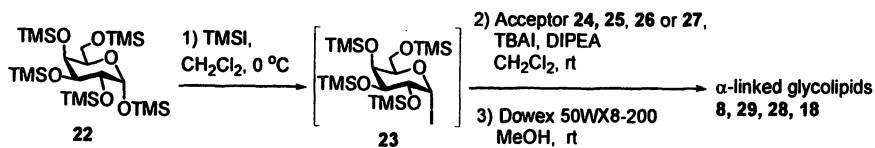
NKT cells, which are vital components of the immune system, recognize lipid antigen-presenting CD1d molecules. This recognition event leads to stimulation of NKT cells resulting in the release of various cytokines. Development of glycolipid antigens that can selectively induce either Th1 or Th2 responses is highly warranted. The availability of structurally diverse and chemically well-defined pure glycolipid analogs for biological assessment is a critical step toward understanding the innate immune system (54-57). Glycosyl iodides have proven to be uniquely suited for the synthesis of both O- and C-glycoconjugates. A one-pot synthesis utilizing fully functionalized acceptors offers an efficient process for analog development as demonstrated by the synthesis of biologically relevant natural products.



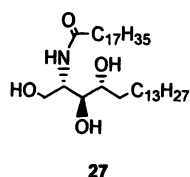
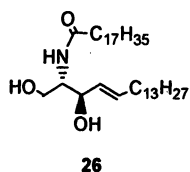
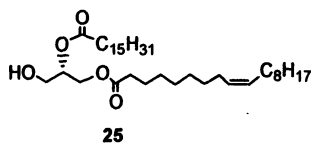
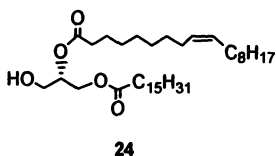
Scheme 4. Synthesis of KRN7000 analogs



Scheme 5. Streamlined synthesis of GalCer



Acceptors



Scheme 6. One-pot synthesis of α -anomeric glycolipids

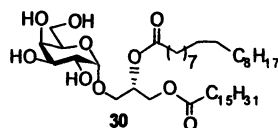
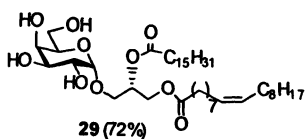
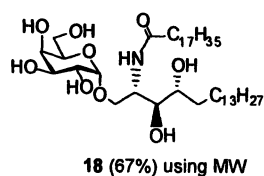
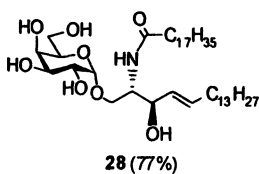
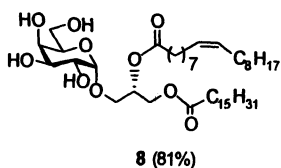
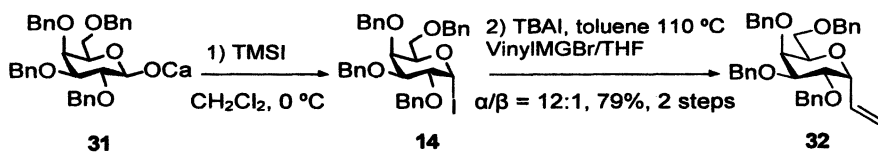
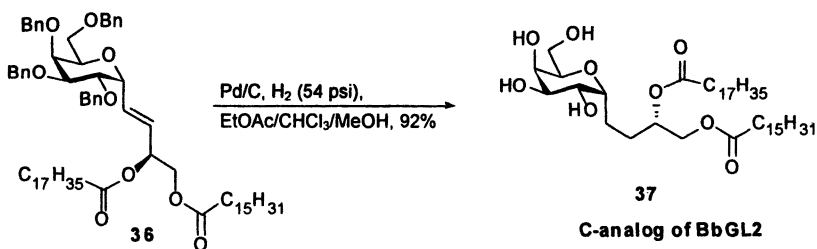
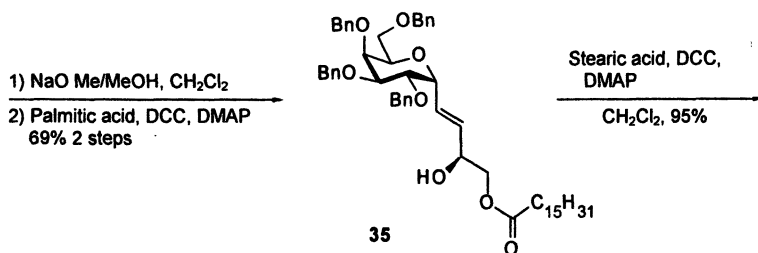
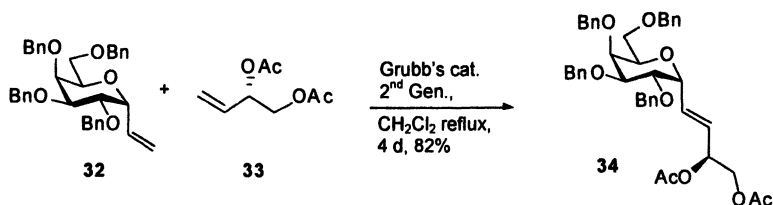


Figure 3. α -anomeric glycolipids



Scheme 7. Stereoselective vinyl Grignard reaction of galactosyl iodide



Scheme 8. Total synthesis of a C-analog of BbGL2

Acknowledgements

This work is supported by National Science Foundation CHE-0210807, NSF CRIF program (CHE-9808183), NSF Grant OSTI 97-24412, and NIH Grant RR11973 provided funding for the NMR spectrometers used on this project.

References

1. For a recent review on structure-activity relationships of glycolipids, see: Savage, P. B.; Teyton, L. Bendelac, A. *Chem. Soc. Rev.* **2006**, *35*, 771-779.
2. Natori, T.; Koezuka, Y.; Higa, T. *Tetrahedron Lett.* **1993**, *34*, 5591-5592.
3. Natori, T.; Morita, M.; Akimoto, K.; Koezuka, Y. *Tetrahedron* **1994**, *50*, 2771-2784.
4. Motoki, K.; Kobayashi, E.; Uchida, T.; Fukushima, H.; Koezuka, Y. *Bioorg. Med. Chem. Lett.* **1995**, *5*, 705-710.
5. Morita, M.; Motoki, K.; Akimoto, K.; Natori, T.; Sakai, T.; Sawa, E.; Yamaji, K.; Koezuka, Y.; Kobayashi, E.; Fukushima, H. *J. Med. Chem.* **1995**, *38*, 2176-2187.
6. Kobayashi, E.; Motoki, K.; Yamaguchi, Y.; Uchida, T.; Fukushima, H.; Koezuka, Y. *Bio. Med. Chem.* **1996**, *4*, 615-619.
7. Kakimi, K.; Guidotti, L. G.; Koezuka, Y.; Chisari, F. V. *J. Exp. Med.* **2000**, *192*, 921-930.
8. Gonzalez-Aseguinolaza, G.; Van Kaer, L.; Bergmann, C. C.; Wilson, J. M.; Schmieg, J.; Kronenberg, M.; Nakayama, T.; Taniguchi, M.; Koezuka, Y.; Tsuji, M. *J. Exp. Med.* **2002**, *195*, 617-624.
9. Duarte, N.; Stenstrom, M.; Campino, S.; Bergman, M. L.; Lundholm, M.; Holmberg, D.; Cardell, S. L. *J. Immunol.* **2004**, *173*, 3112-3118.
10. Kawano, T.; Cui, J.; Koezuka, Y.; Toura, I.; Kaneko, Y.; Motoki, K.; Ueno, H.; Nakagawa, R.; Sato, H.; Kondo, E.; Koseki, H.; Taniguchi, M. *Science* **1997**, *278*, 1626-1629.
11. Porcelli, S. A.; Modlin, R. L. *Annu. Rev. Immunol.* **1999**, *17*, 297-329.
12. Zeng, Z.; Castano, A. R.; Segelke, B. W.; Stura, E. A.; Peterson, P. A.; Wilson, I. A. *Science* **1997**, *277*, 339-345.
13. Pal, E.; Tabira, T.; Kawano, T.; Taniguchi, M.; Miyake, S.; Yamamura, T. *J. Immunol.* **2001**, *166*, 662-668.
14. Berkers, C. R.; Ovaa, H. *Trends. Pharmacol. Sci.* **2005**, *26*, 252-257.
15. Miyamoto, K.; Miyake, S.; Yamamura, T. *Nature* **2001**, *413*, 531-534.
16. Yang, G.; Schmieg, J.; Tsuji, M.; Franck, R. W. *Angew. Chem. Int. Ed.* **2004**, *43*, 3818-3822.

17. Chen, J.; Schmiege, J.; Tsuji, M.; Franck, R. W. *Org. Lett.* **2004**, *6*, 4077-4080.
18. Chen, G.; Chien, M.; Tsuji, M.; Franck, R. W. *ChemBioChem.* **2006**, *7*, 1017-1022.
19. Fujio, M.; Wu, D.; Garcia-Navarro, R.; Ho, D. D.; Tsuji, M.; Wong, C.-H. *J. Am. Chem. Soc.* **2006**, *128*, 9022-9023.
20. Wu, D.; Xing, G. W.; Poles, M. A.; Horowitz, A.; Kinjo, Y.; Sullivan, B.; Bodmer-Narkevitch, V.; Plettenburg, O.; Kronenberg, M.; Tsuji, M.; Ho, D. D.; Wong, C.-H. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1351-1356.
21. Kinjo, Y.; Wu, D.; Kim, G.; Xing, G. W.; Poles, M. A.; Ho, D. D.; Tsuji, M.; Kawahara, K.; Wong, C.-H.; Kronenberg, M. *Nature* **2005**, *434*, 520-525.
22. Sriram, V.; Du, W.; Gervay-Hague, J.; Brutkiewicz, R. *Eur. J. Immunol.* **2005**, *35*, 1692-1701.
23. Hossain, H.; Wellensiek, H.-J.; Geyer, R.; Lochnit, G. *Biochimie* **2001**, *83*, 683-692.
24. Ben-Menachem, G.; Kubler-Kielb, J.; Coxon, B.; Yergey, A.; Schneerson, R. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 7913-7918.
25. Kinjo, Y.; Tupin, E.; Wu, D.; Fujio, M.; Garcia-Navarro, R.; Benhia, M. R.; Zanonc, D. M.; Ben-Menachem, G.; Ainge, G. D.; Painter, G. F.; Khurana, A.; Hoebe, K.; Behar, S. M.; Beutler, B.; Wilson, I. A.; Tsuji, M.; Sellati, T. J.; Wong, C.-H.; Kronenberg, M. *Nat. Immunol.* **2006**, *7*, 978-986.
26. Kulkarni, S. S.; Gervay-Hague, J. *Org. Lett.* **2006**, *8*, 5765-5768.
27. Sakai, T.; Ueno, H.; Natori, T.; Uchimura, A.; Motoki, K.; Koezuka, Y. *J. Med. Chem.* **1998**, *41*, 650-652.
28. Ndongye, R. M.; Izmirian, D. P.; Dunn, M. F.; Yu, K. O.; Porcelli, S. A.; Khurana, A.; Kronenberg, M.; Richardson, S. K.; Howell, A. R. *J. Org. Chem.* **2005**, *70*, 10260-10270.
29. Kim, S.; Song, S.; Lee, T.; Jung, S.; Kim, D. *Synthesis* **2004**, 847-850.
30. Morita, M.; Motoki, K.; Akimoto, K.; Natori, T.; Sakai, T.; Sawa, E.; Yamaji, K.; Koezuka, Y.; Kobayashi, E.; Fukushima, H. *J. Med. Chem.* **1995**, *38*, 2176-2187.
31. Xia, C.; Yao, Q.; Schumann, J.; Rossy, E.; Chen, W.; Zhu, L.; Zhang, W.; De Libero, G.; Wang, P. G. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 2195-2199.
32. Plettenburg, O.; Bodmer-Narkevitch, V.; Wong, C.-H. *J. Org. Chem.* **2002**, *67*, 4559-4564.
33. Figueroa-Perez, S.; Schmidt, R. R. *Carbohydr. Res.* **2000**, *328*, 95-102.
34. Goff, R. D.; Gao, Y.; Mattner, J.; Zhou, D.; Yin, N.; Cantu, C.; Teyton, L.; Bendelac, A.; Savage, P. B. *J. Am. Chem. Soc.* **2004**, *126*, 13602-13603.
35. Luo, S.-Y.; Kulkarni, S. S.; Chou, C.-H.; Liao, W.-M.; Hung, S.-C. *J. Org. Chem.* **2006**, *71*, 1226-1229.

36. Gervay, J.; Nguyen, T. N.; Hadd, M. J. *Carbohydr. Res.* **1997**, *300*, 119-125.
37. Gervay, J.; Hadd, M. J. *J. Org. Chem.* **1997**, *62*, 6961-6967.
38. Gervay, J. *Organic Synthesis: Theory and Applications*; JAI Press: Greenwich **1998**, pp.121-153.
39. Hadd, M. J.; Gervay, J. *Carbohydr. Res.* **1999**, *320*, 61-69.
40. Lam, S. N.; Gervay-Hague, J. *Org. Lett.* **2002**, *4*, 2039-2042.
41. Lam, S. N.; Gervay-Hague, J. *Carbohydr. Res.* **2002**, *337*, 1953-1965.
42. Bhat, A. S.; Gervay-Hague, J. *Org. Lett.* **2001**, *3*, 2081-2084.
43. Ying, L.; Gervay-Hague, J. *Carbohydr. Res.* **2003**, *338*, 835-841.
44. Lam, S. N.; Gervay-Hague, J. *Org. Lett.* **2003**, *5*, 4219-4222.
45. Dabideen, D. R.; Gervay-Hague, J. *Org. Lett.* **2004**, *6*, 973-975.
46. Lam, S. N.; Gervay-Hague, J. *J. Org. Chem.* **2005**, *70*, 2387-2390.
47. Lam, S. N.; Gervay-Hague, J. *J. Org. Chem.* **2005**, *70*, 8772-8779.
48. El-Badry, M. H.; Gervay-Hague, J. *Tetrahedron Lett.* **2005**, *46*, 6727-6728.
49. El-Badri, M. H.; Willenbring, D.; Tantillo, D. J.; Gervay-Hague, J. *J. Org. Chem.* **2007**, *72*, 4663-4672.
50. Lemieux, R. U.; Hendriks, K. B.; Stick, R. V.; James, K. *J. Am. Chem. Soc.* **1975**, *97*, 4056-4062.
51. Du, W.; Gervay-Hague, J. *Org. Lett.* **2005**, *7*, 2063-2065.
52. Sakai, T.; Ueno, H.; Natori, T.; Uchimura, A.; Motoki, K.; Koezuka, Y. *J. Med. Chem.* **1998**, *41*, 650-652.
53. Du, W.; Kulkarni, S. S.; Gervay-Hague, J. *Chem Commun.* **2007**, 2336-2339.
54. Larkins, J.; Renukaradhya, G. J.; Sriram, V.; Du, W.; Gervay-Hague, J.; Brutkiewicz, R. *J. Immunol.* **2006**, *177*, 268-279.
55. Renukaradhya, G. J.; Roberts Webb, T. J.; Khan, M. A.; Lin, Y. L.; Du, W.; Gervay-Hague, J.; Brutkiewicz, R. *J. Immunol.* **2005**, *175*, 4301-4308.
56. Roberts Webb, T. J.; Litavec, R. A.; Khan, M. A.; Du, W.; Gervay-Hague, J.; Renukaradhya, G. J.; Brutkiewicz, R. *Eur. J. Immunol.* **2006**, *36*, 2595-2600.
57. Renukaradhya, G. J.; Sriram, V.; Du, W.; Gervay-Hague, J.; Kaer, L. V.; Brutkiewicz, R. *Int. J. Cancer* **2006**, *118*, 3045-3053.

Chapter 8

Chemical Glycobiology of Glycosphingolipids

Chengfeng Xia¹, Yalong Zhang¹, Wenpeng Zhang¹, Wenlan Chen¹,
Jing Song¹, Qingjia Yao¹, Yang Liu¹, Dapeng Zhou²,
Gennaro De Libero³, and Peng George Wang^{1,*}

¹The Departments of Biochemistry and Chemistry, The Ohio State University, 484 West 12 Avenue, Columbus, OH 43210

²Department of Melanoma Medical Oncology, University of Texas M. D. Anderson Cancer Center, Houston, TX 77030

³Experimental Immunology, University Hospital, 4031 Basel, Switzerland

The MHC-I like CD1d presents glycolipids to NKT cells, a subpopular T cells, for recognition while the classic MHC molecules present peptides to regulatory T cells. Upon stimulation, NKT cells produce cytokines and chemokines to regulate autoimmune responses. The marine sponge α -galactosylceramides, bacterial α -galacturonosylceramides and mammalian isoglobotrihexosylceramide are the well-known glycolipids which can be presented by CD1d to NKT cells. The former two have similar structures with an α -linkage between the sugar part and ceramide part, while the latter possesses a β -linkage. Structure activity relationship studies of the glycolipids were employed to demonstrate the interactions among CD1d, glycolipid and the T cell receptor (TCR) of NKT cells. Functionalization on the C3' and C4' of the KRN7000 showed that these analogues could selectively bind different species of NKT cells. By variation of the ceramide of iGb3, the HO-iGb3 which has phytosphingosine as the lipid part, was shown to possess greater potential in stimulating NKT cells to release cytokines. Four deoxy-iGb3 analogues of the terminal sugar were also prepared to illustrate the role of these hydroxy groups in communicating with NKT cells. A

metabolically stable iGb3 analogue, named *S*-iGb3, was synthesized to test whether it can improve activity.

Introduction

In contrast to the classic CD4⁺ and CD8⁺ T cells of the immune system that recognize specific peptide antigens bound to major histocompatibility complex (MHC) class II or class I antigen presenting proteins, respectively (1-3), lipids and glycolipid antigens are presented by MHC-I like CD1 molecules to non-MHC-restricted T lymphocytes (Figure 1) (4). CD1 proteins have evolved a unique hydrophobic binding groove that accommodates lipid antigens in both the secretory and endosomal compartments (5, 6). There are five classes of CD1 proteins in humans, comprised of group I CD1 (CD1a, -b, -c), group II CD1 (CD1d and CD1e) (7-11). However, a single class of CD1 protein (mCD1d) was found in mice, which is homologous to the human isoform CD1d. CD1d presents glycolipids to invariant natural killer T (NKT) cells, which co-express the natural killer marker NK1.1 and a T cell receptor (TCR). In mice, the CD1d-restricted T cells, called V α 14i NKT cells express a semi-invariant TCR formed by V α 14-J α 1 β and V β 8.2/V β 7/V β 2 chains (12, 13). In humans, NKT cells (V α 24i NKT cells) express the homologous V α 24-J α 1 β and V β 11 chains (14). NKT cells play a major role as a bridging system between innate and adaptive immunity (15). Like natural killer cells, NKT cells are activated in the first line of immune response within 2 to 6 hours upon stimulation to secrete proinflammatory T helper 1 (Th1) and immunomodulatory Th2 cytokines and chemokines. Secretion of these cytokines initiates the proliferation of lymphocytes for inflammation and immunoregulation activities. NKT cells activation propagate rapidly to other cell types, among which are NK cells, dendritic cells (DC), and subsets of B and conventional T cells. Thus, the quick and vigorous response of NKT cells to α -GalCer is closer to an innate rather than an adaptive immune response.

Agelasphins were a series of extracts from marine sponge discovered by Kirin Brewery Company in 1993 (16, 17). They showed a potent activity in prolonging the life span of mice when intraperitoneally inoculated with B16 mouse melanoma cells. The most striking structural feature of agelasphins is the α linkage between the sugar and ceramide, which is very uncommon in mammalian systems. The structure activity studies of the anti-tumor properties of glycolipids of marine origin revealed an optimized ligand for CD1d binding, termed KRN7000 (Figure 2), which possesses an 18-carbon phytosphingosine and 26-carbon acyl chain (18). This α -galactosylceramide (α -GalCer) can be readily loaded onto CD1d protein, and the CD1d/ α -GalCer complex on the surface of antigen presenting cells (APC, such as dendritic cells and

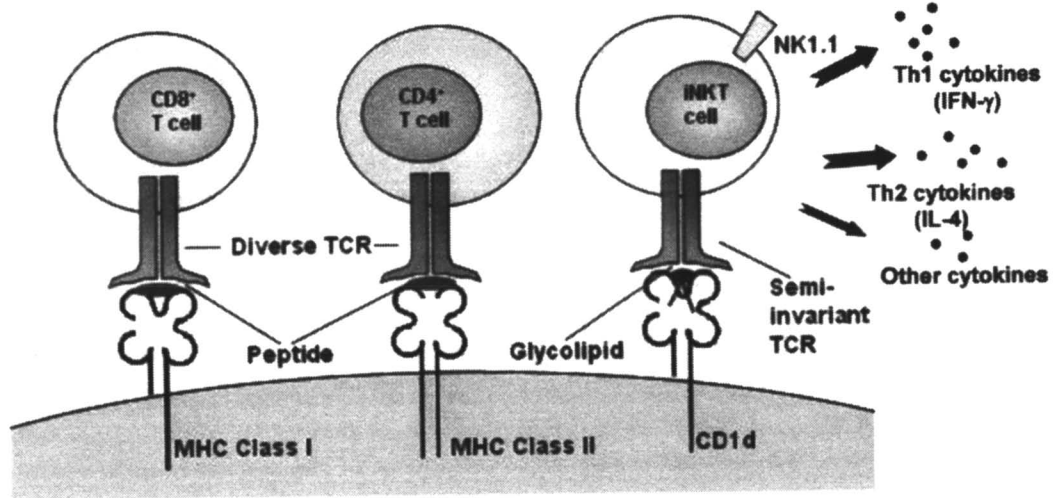


Figure 1. Presentation of peptide antigens and lipid/glycolipid antigens to T lymphocytes.

macrophages) will be recognized by the TCR of all types of NKT cells (19-21). The CD1d/ α -GalCer/TCR association triggers a rapid, transient, and massive response of NKT cells that accumulates Th1 and Th2 signaling peptides which includes the secretion of IFN- γ and IL-4.

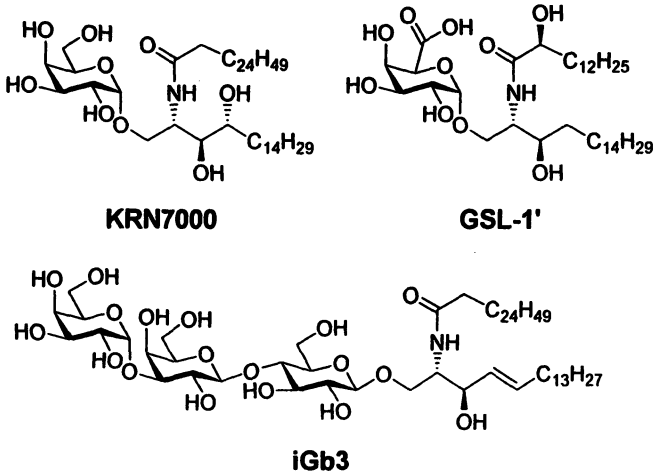


Figure 2. Structures of exogeneous and endogeneous glycolipid antigens of NKT cells.

α -GalCer has shown effective tumor suppressing properties in a variety of transplantable tumors, including melanomas, lymphomas, colon, prostate, lung, breast, renal and transplantable tumors. Besides the anticancer application, it also showed promise in treating a variety of autoimmune pathologies in animal models including multiple sclerosis (22), autoimmune type 1 diabetes (20, 23) and experimental encephalomyelitis (24, 25). Moreover, the promising effects of α -GalCer in organ transplants (26) and atherosclerosis (27) have been evidenced in animal models. The NKT cells involved in many autoimmune diseases have different roles. They have a suppressive role in diseases like type I diabetes, multiple sclerosis, rheumatoid arthritis, and Systemic Lupus Erythematosus (SLE) (14, 28). On the other hand, NKT cells can suppress protective immune responses or promote the development of atherosclerosis (29-31). The detailed mechanism involved in these phenomena still need to be uncovered. Recently, the role of α -GalCer and NKT cell in allergic asthma has been investigated. The latest results show that IFN- γ comes from α -GalCer stimulated NKT cells. Therefore, α -GalCer plays a positive role in inhibiting allergic asthma, while IL-4 may increase its severity (32-34).

Since α -GalCer comes from marine sponges and is not a normal product of mammalian cells, there has been a great effort to find the endogenous antigens for NKT cell development and activation. Based on several lines of evidence, Bendelac laboratory, in late 2004, discovered a lysosomal isoglobotrihexosylceramide (iGb3) (Figure 2) which is an endogenous ligand for NKT cells (35, 36). The iGb3 is a trisaccharide glycolipid whose sugar sequence is Gal α 1-3Gal β 1-4Glc β , and the ceramide consisting of 18-carbon *D*-erythro-sphingosine and 26-carbon acyl chain. Subsequently, several laboratories have confirmed that iGb3 stimulates both human V α 24i NKT cells and mouse V α 14i NKT cells, but its avidity is far less than α -GalCer. It is amazing that nature picks iGb3 from a number of similar glycolipids in the metabolism/catabolism pathway of glycosphingolipids as an internal agonist for NKT cells. Neither iGb4 (GalNAc β 1-3Gal α 1-3LacCer), Gb4 (GalNAc β 1-3Gal α 1-4LacCer), Gb3, β -LacCer, nor β -GlcCer are agonists for NKT cells even though they contain part of the carbohydrate sequence of iGb3 (37, 38). However, the processing of iGb4 into iGb3 in the lysosome was found to be necessary for NKT cell recognition (35). This indicated that the terminal galactose in iGb3 is the most critical recognition moiety in the CD1d/glycolipid/TCR complex. In addition, it was also found that trafficking of CD1d to lysosomal compartments and the presence of lipid transport proteins such as saposins were essential in presenting the iGb3 antigen (39).

However, the mammalian cells have less chances for contacting with the marine sponge, and the iGb3 is so weak and seemed existing more likely for the development of NKT cells from the mainstream T cells precursor than as an endogeneous antigen to trigger the innate immune system. So what is the purpose of mammalian cells to generate NKT cells? Many microbes produce lipopolysaccharide which can bind with CD14 and then be recognized by macrophages or other antigen presenting cells to evoke the inflammatory response. However, multiple Gram-negative bacteria do not produce lipopolysaccharide (40). How does the innate immune system detect the presence of these microbes? The discovery of glycosphingolipids from the cell wall of *Sphingomonas*, which also can be recognized by mouse and human NKT cells, suggested that NKT cells provided an innate immune response to certain microorganisms through recognition of their antigen receptor (41, 42). These might be useful in providing protection from bacteria that can't be detected by pattern recognition receptors such as Toll-like receptor. The glycosphingolipids separated from *Sphingomonas* species include α -galacturonosylceramide (GSL-1') and α -glucuronosylceramide (GSL-1) (Figure 2). They possess a similar structure to α -GalCer except for the carboxyl group on 6'-position of the sugar. When they were added to a CD1d-coated plate, they could stimulate a V α 14i NKT cell hybridoma to release IL-2, indicating that they were the antigen for activation of NKT cells. None of these compounds were as active as the

synthetic KRN7000, however, and the galacturonic-acid containing GSL-1' was more reactive than the glucuronic-acid containing GSL-1.

Modification of α -Galactosylceramide on 3'- and 4'- Positions

Even though KRN7000 was found to be the best glycosyl ceramide for NKT cell activation, its high immunostimulation property has led to the synthesis of many analogs to find improved results. A variety of α -GalCer analogues with truncated sphingosine and acyl chains have been synthesized and screened for NKT cell stimulating potencies. The results showed that the structure of glycolipid affects not only the magnitude of the NKT cell stimulation, but also the nature of the stimulation. Interestingly one of the α -GalCer analogs with a 9-carbon sphingosine chain and a 24-carbon acyl chain, OCH, was able to elicit an altered cytokine production profile of NKT cells, with a suppressed IFN- γ response but retained IL-4 response (25). Moreover, incorporation of unsaturated bonds into the acyl chain of α -GalCer also polarized the secretion of cytokines. The glycolipid with *cis*-double bonds in the acyl chain such as C20:2 caused a Th2 bias in NKT cells responses similar to the lipid chain-truncated glycolipids (43). A C-glycoside analog of KRN7000 was synthesized and was found to activate NKT cells at very low concentrations and promote Th1 responses *in vivo* (44). The C-glycoside was much more stable *in vivo*, and was thus effective with intervals up to four days between drug dosage and challenge, whereas the O-glycoside was effective with only a one-day interval (45).

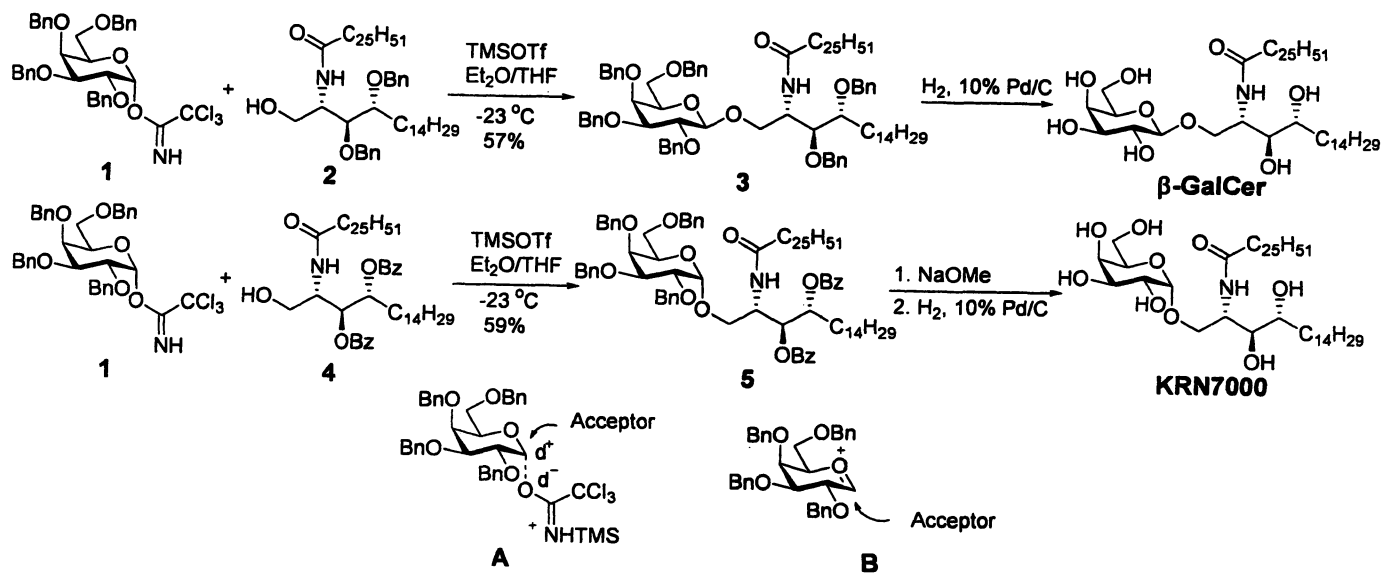
Some extensive results were published on the role of the sugar part in the stimulation of NKT cells. Replacement of the galactose sugar with glucose results in a weaker agonist than α -GalCer (46). The 2-hydroxyl of the sugar was found to be very critical for CD1d binding (47, 48). When it was substituted by hydrogen, fluoro, or acetylamide, the associated biological activities were dramatically decreased (49). Incorporation of a fluorophore or biotin group at C6' of the galactose of α -GalCer, did not show significant stimulatory differences between these compounds and KRN7000 (50). These results suggest that small modifications at the C6' position of the galactose do not interfere with the stimulatory potency of glycolipid. These discoveries are in agreement with the experiments done by Taniguchi (21), Kronenberg (51) and Zhou (35). When the C2, C3 or C4 of the galactose was capped with another sugar, the resulting diglycosylceramide lost its activity if the additional sugar cap was not removed by glycosidase. In contrast, the Gal α 1,6GalCer still stimulated NKT cells without truncation to the monoglycosylceramide.

The KRN7000 was synthesized from the galactosyl donor and the amide-type acceptor (52). In our initial synthesis of α -GalCer, we used the active "armed" trichloroacetimidate donor 1 and acceptor 2 under TMSOTf as a

catalyst for the glycosylation reaction (Scheme 1). We observed the formation of β -galactosyl ceramide derivative **3** under these reaction conditions. We then switched the protecting groups on the ceramide acceptor from benzyl to benzoyl to reduce the activity of the acceptor. Under the same reaction conditions, glycosylation of **1** with **4**, compound **5** was isolated as the sole product with the desired α -configuration. Our hypothesis for the formation of β -GalCer in case of benzyl protected acceptor is that it results from the high reactivity of **2**. It is well known that the glycosylation outcome depends on the delicate balance of donor/acceptor reactivity. The formation of β -anomeric product was hypothesized to have resulted from the high reactivity of the acceptor. While the donor was activated by TMSOTf, the armed acceptor attacked the donor in an SN2 fashion in transition state A before the full development of oxacarbenium intermediate. The result is production of β -configuration. It was expected that a less reactive acceptor favored α -glycosylation through a fully developed oxacarbenium intermediate.

Till date, most of the research has been focused on variation of the lipid part, sphingosine (18, 25, 33) moiety or acyl chain (43, 54), of the α -GalCer. Only very recently a few papers reported the replacement of the 2'-hydroxy group with fluoride, acetamide, or hydrogen moieties. However, in all these analogues the biological activity was completely abolished (49). Since the 3'-OH and 4'-OH of galactose are also assumed to be crucial to activate NKT cells no further research was carried out to demonstrate the importance of 3' or 4'-OH modified α -GalCer analogues. It was of our interest to synthesize and check the biological activities of α -GalCer analogues at these positions.

The 3'-azido-substituted galactosyl donor was prepared from 1,2:5,6-di-*O*-isopropylidene- α -gulofuranose **6** which is accessible from diacetone-D-glucose (55). The azido group was introduced by first activating the 3-hydroxy group with Tf₂O followed by substitution with NaN₃ (Scheme 2). The two isopropylidene groups were removed by treatment with 90% TFA. Then it was acetylated by using Ac₂O in pyridine to give compound **8**. It was subject to glycosylation with thiophenol and the acetyl protecting group was removed with freshly prepared NaOMe. The C4 and C6 hydroxy group were selectively protected with benzylidene and the C2 hydroxy with PMB to give compound **10**. Since both of the stereoselectivity and the yield of the glycosylation reaction were unsatisfactory when using thiophenyl with ceramide, the thio-donor **10** was converted to trichloroacetimide donor **11** by Schmidt's method. The glycosylation of the 3'-substituted galactosyl donor with acceptor **12** was performed exactly according to the protocol for preparation of KRN7000 giving a yield up to 88%. Here the short acyl chain (8-carbon instead of 26-carbon) was employed to improve the solubility of glycolipid in aqueous during bioassay. The two benzoyl groups were saponified by treatment with freshly prepared



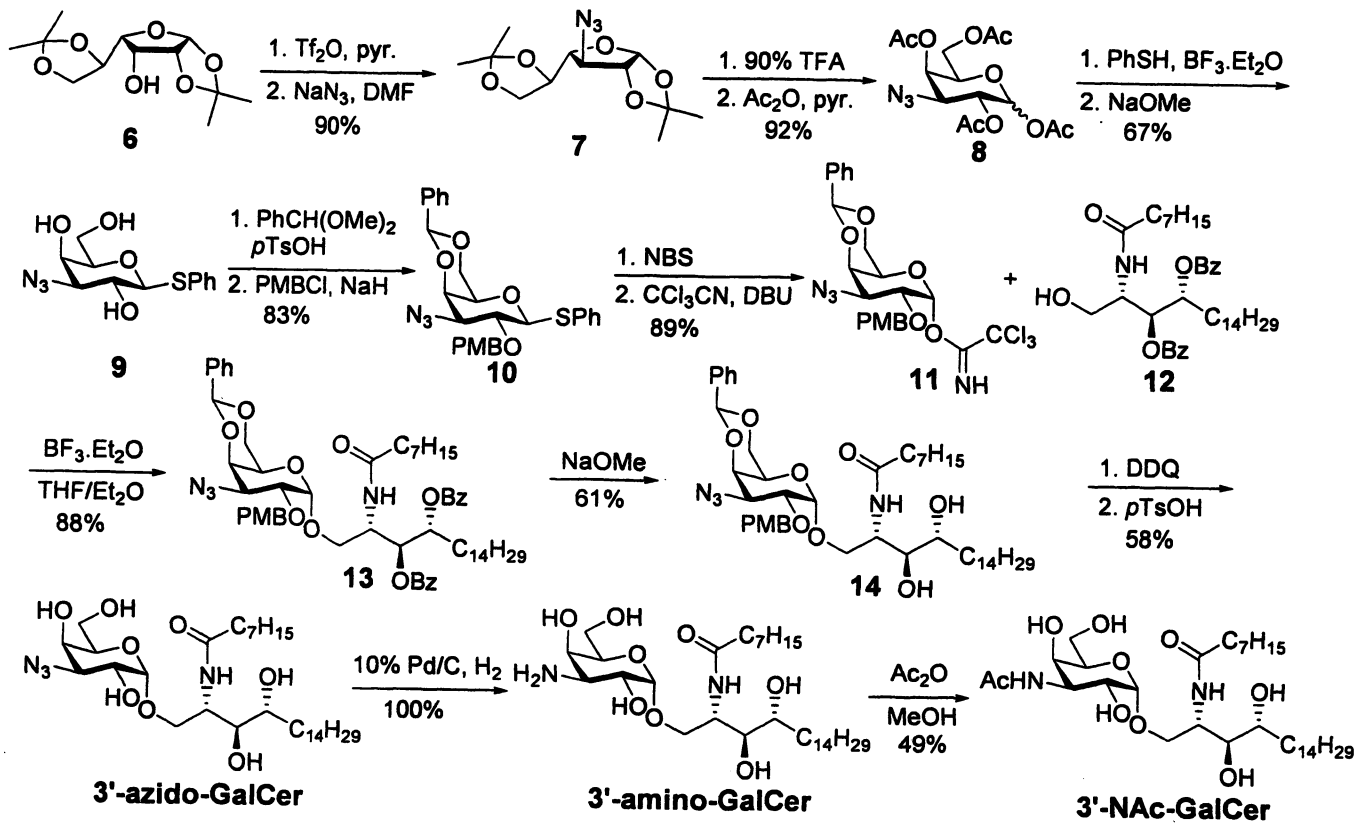
Scheme 1. Synthesis of KRN7000 using the "armed" and "disarmed" concept.

NaOMe in dry MeOH. After removal the PMB group with DDQ and the benzylidene group in acidic condition, the 3'-azido-GalCer was obtained in good yield. Reduction of the azido group to the amino group afforded the 3'-amino-GalCer. Selectively acetylation of the amino group in MeOH afforded the 3'-NAc-GalCer.

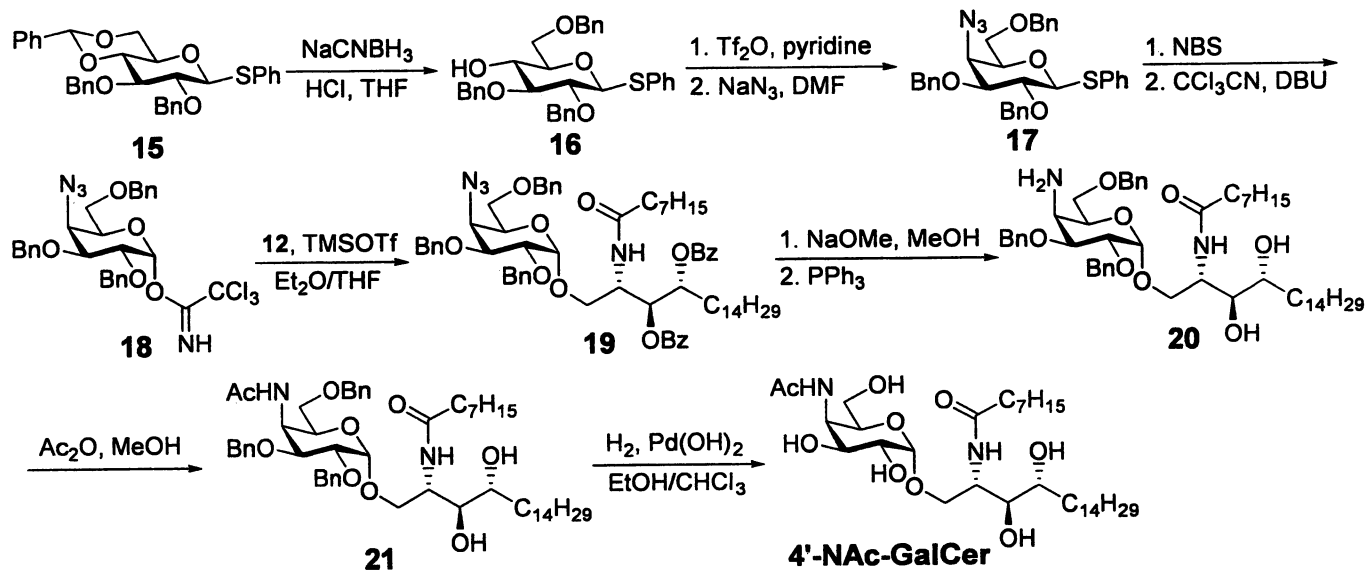
The 4'-substituted α -GalCer analogues were also designed and some of them were synthesized (Scheme 3). The benzylidene of the glucose **15** was regioselectively reduced in acidic condition by NaCNBH₃ to give **16**. The azido group was then installed using a SN2 reaction to give the 4-azido-galactose **17**. The trichloroacetimide donor **18** was subsequently generated under reaction conditions similar to those used to prepare the donor **11**. Glycosylation via catalysis with TMSOTf afforded the glycolipid **19**. At this stage the two benzoyl groups were removed under basic condition. However, when it was subjected to hydrogenation, several by-products were formed with no product being detected. To overcome this obstacle, the azido group was reduced to the amino group by treatment with triphenylphosphine. Reaction with acetic anhydride then gave compound **21**, which was subjected to hydrogenation again, this time resulting in a smoother reaction to afford the product 4'-NAc-GalCer in excellent yield.

In Figure 3A, a mouse NKT cell hybridoma, which is a selected clone that produces high amount of IL-2 cytokine, was used to measure the stimulatory activities of glycolipid antigens. The glycolipid antigens are presented by mouse bone marrow derived dendritic cells, which are the most potent cell type in terms of presenting glycolipid antigens. The mouse NKT hybridoma cells secrete IL-2 when stimulated by glycolipids presented by dendritic cells. Figure 3A indicate that the 3' and 4' substitutions cause complete loss of stimulatory activity.

In Figure 3B, a direct TCR binding assay was used to measure the affinity of glycolipids to the invariant T cell receptor of NKT cells. This assay was based on the tetramer technology as staining T cell receptors. The glycolipids were loaded to biotinylated CD1d protein. A multimer of the glycolipid-CD1d complex was then formed by streptavidin, which was conjugated to a fluorochrome. The multimeric glycolipid-CD1d can directly stain the NKT cells, and the affinity can be reflected by mean fluorescence intensity. The NKT cells are often stained by combination with another color for T cell receptor (double color staining). In Figure 3B, the X axis represents the staining signal for the V α 24 NKT cell receptor (used by human and rhesus NKT cells but not mouse NKT cells). The Y axis represents the NKT cells which are simultaneously stained by the glycolipid-CD1d tetramers (with a different color). We used this method to stain the human, rhesus and mouse NKT cell lines. The human NKT cell line is expanded NKT cells from a healthy human donor, which contains 50% of V α 24i NKT cells. The rhesus NKT cell line is expanded NKT cells from a healthy rhesus monkey, that contains 90% of V α 24 NKT cells. The mouse NKT cell line is a mixture of V α 14i NKT cell (recognizing α -GalCer), and a



Scheme 2. Synthesis of 3'-substituted α -GalCer analogues.



Scheme 3. Synthesis of the 4'-N-Ac-GalCer.

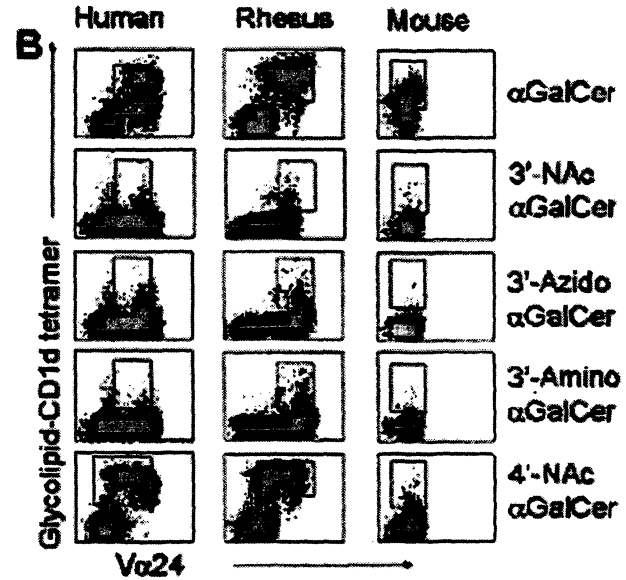
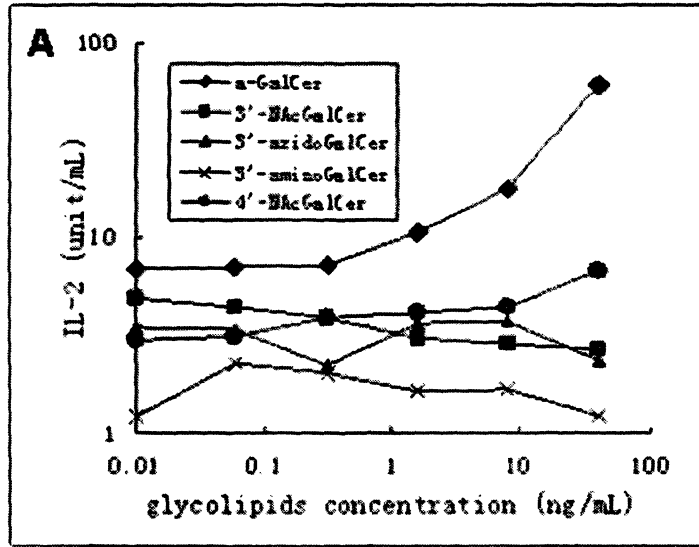


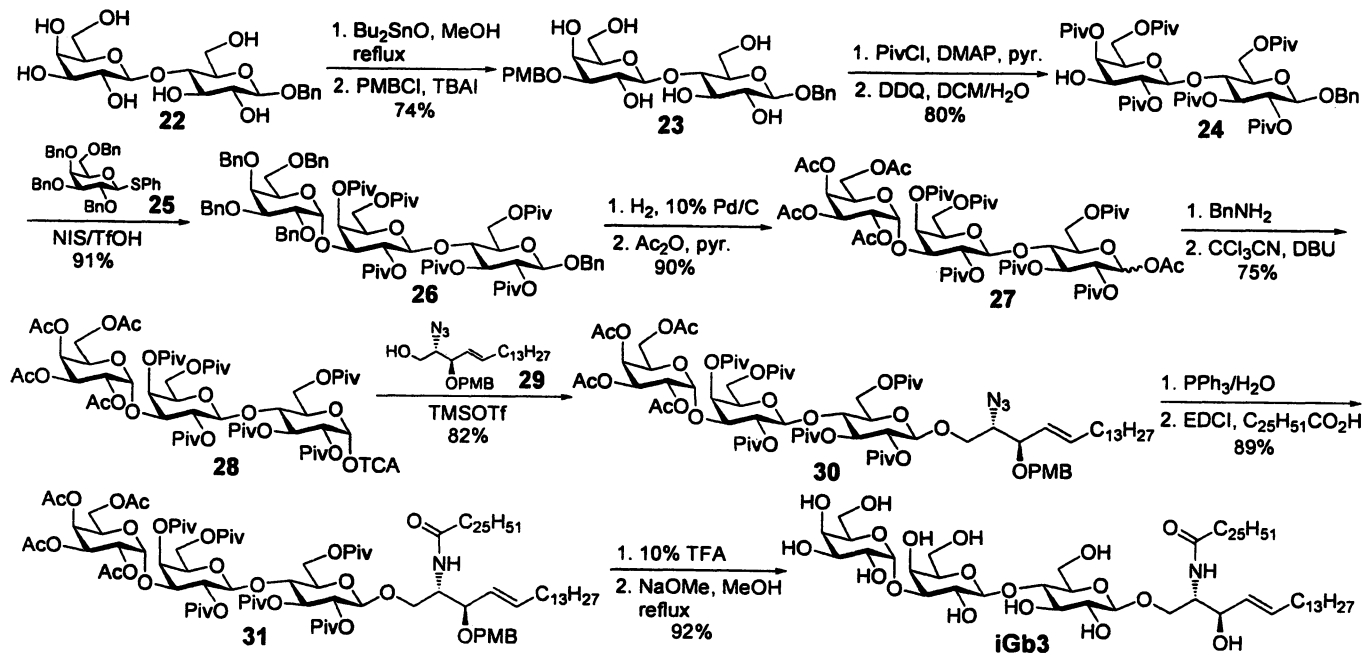
Figure 3. Stimulation of hybridomas NKT cells by α -GalCer analogues and measurement of the binding with NKT cells

non-V α 14i NKT cells (no binding to α -GalCer), at a 50/50 percentage. **Figure 3B** indicates that all of the 3' and 4' substituted α -GalCer lost binding to mouse NKT cells (V α 24 negative because mouse use V α 14 as the invariant TCR). However, for both human and rhesus NKT cells (V α 24 positive), we observed a significant percentage of tetramer positive cells (3 to 90 percentage among the V α 24 positive NKT cells). The results thus showed that a significant percentage of human and rhesus NKT cells bind to 3' and 4'-substituted α -GalCer, especially in the case of 4'-NAc-GalCer.

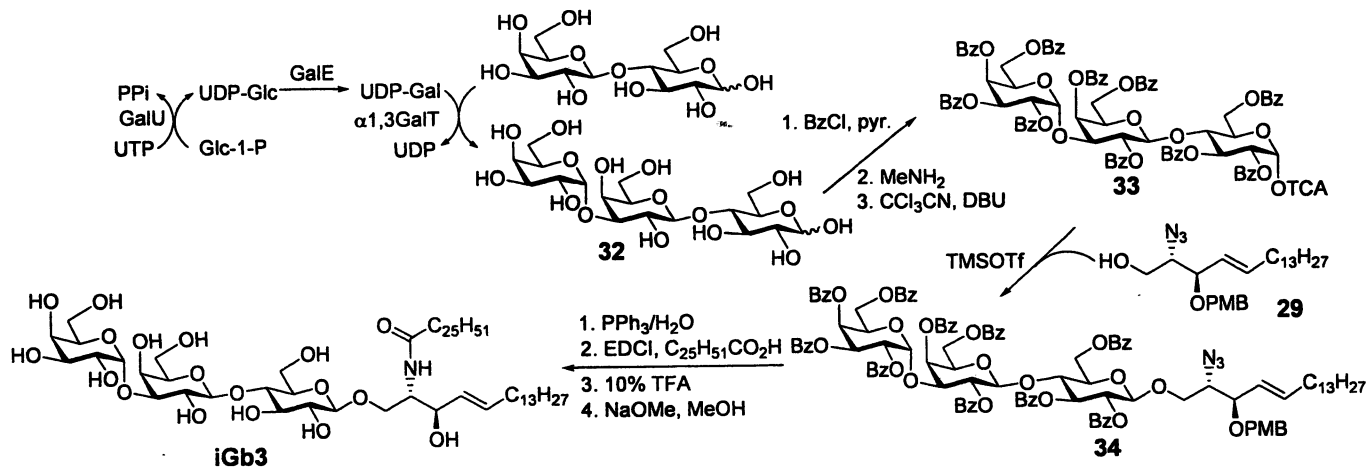
Chemo and Chemo-Enzymatic Synthesis of iGb3

The chemical synthesis of the trisaccharide Gal α 1-3Gal β 1-4Glc has been well developed in the past decade (56, 57). The regioselective monoalkylation of the C3' hydroxy group of the benzyl lactose **22** with *p*-methoxybenzyl chloride (PMBCl) was achieved through activating this hydroxy group with dibutylstannylen acetal (**Scheme 4**). The bulky pivaloyl group, which could significantly restrain the formation of the orthoester, was employed to protect the remaining hydroxy groups. Oxidative cleavage of the PMB group by DDQ released the 3'-OH to afford the disaccharide acceptor **24**. The glycosylation of acceptor **24** with perbenzyl phenyl thiogalactoside donor **25** was carried out under activation by NIS/TfOH to afford the protected trisaccharide **26** in a yield of 91%. At this juncture, the benzyl groups were switched to acetyl groups to avoid reduction of the double bond on the lipid during hydrogenation. Selective exposure of the anomeric hydroxy group was achieved by treatment with benzylamine in THF. It was then converted to trichloroacetimidate **28** under basic conditions with trichloroacetinitrile.

When an amide compound was used as the acceptor for glycosylation, the yield is around 45% (35). It is well known that the amide can cause low reactivity on the glycosylation due to hydrogen binding. As a result, we used azide **29** (21) instead of the amide for the acceptor. Furthermore, it was found that if the 3-OH on sphingosine was protected with an ester, the acyl group would shift to the amine during reduction of the azide (58). PMB is an excellent protecting group which can be cleaved by oxidation or under acidic conditions, and allowed this problem to be overcome. The glycosylation of the Schmidt's type donor **28** with lipid **29** was performed at -20 °C for 2 h in 82% yield. The azide was reduced to an amine by triphenylphosphine with stoichiometric amount of water, and then subjected to amide formation with cerotic acid and EDCI to afford protected iGb3 **31**. Cleavage of the PMB proceeded smoothly in 10% trifluoroacetic acid in dichloromethane. Finally, all the ester groups were removed by reaction with sodium methoxide in refluxing methanol for 24 h to yield iGb3 in 92% yield.



Scheme 4. Synthesis of iGb3 by the chemo method.



Scheme 5. Chemoenzymatic synthesis of *i*Gb3 using the three-enzyme system.

Generally, enzymatic glycosylation is one of the most practical methods for oligosaccharide synthesis. The isoglobotrihexose, which is the trisaccharide for the synthesis of iGb3, can be easily formed by catalysis with a single galactotransferase called α -1,3-GalT using UDP-galactose (UDP-Gal) and lactose. However, the high cost of UDP-Gal limits application of the synthesis to the large-scale. Multiple-enzyme sugar nucleotide regeneration systems have been extensively developed to avoid using the costly stoichiometric amount of sugar nucleotides (59, 60). In these systems, UDP-Gal was generated through recycling of UTP by multiple enzymes as follows: UDP-Gal reacts with lactose to produce trisaccharide and the by-product UDP. The UDP then is recycled by the enzyme PykF to regenerate UTP. The UTP is further converted to UDP-Gal by the enzymes GalU and GalPUT (Scheme 1). Nowadays, UTP is commercially available in large quantities at a low price. Therefore, the direct use of UTP in stoichiometric amounts can dramatically simplify the multiple-enzyme UDP-Gal generation system. Previously, we reported the enzymatic syntheses of globotrihexose and isoglobotrihexose using superbead or superbug techniques (61, 62). Herein we have developed a simpler 3-enzyme UDP-Gal generation system to produce trisaccharides (Scheme 5) (63). The His6-tagged enzymes GalE, GalU and α -1,3-GalT were immobilized on a Ni²⁺ resin column, after which a solution of Glu-1-P, UTP and lactose containing 0.01 M Tri-HCl and 0.01 M MnCl₂ was circulated through the column by a pump at room temperature for two days. After the reaction was complete (monitored by TLC), the solution was pumped out for purification and the column was ready for the next batch of reactions. Using this methodology, the isoglobotrihexose was isolated in 78% yield.

The isoglobotrihexose was subsequently perbenzoylated to ensure the formation of β -configuration in later glycosylation. The anomeric Bz was removed by treatment with MeNH₂ and was then converted to trichloroacetimide donor as in the above chemo-synthesis shown in Scheme 4. The glycosylation was performed as described and the acyl group was introduced by reduction following acylation. Removal of all the protecting groups afforded the iGb3. This chemoenzymatic protocol was much more efficient and easier to perform.

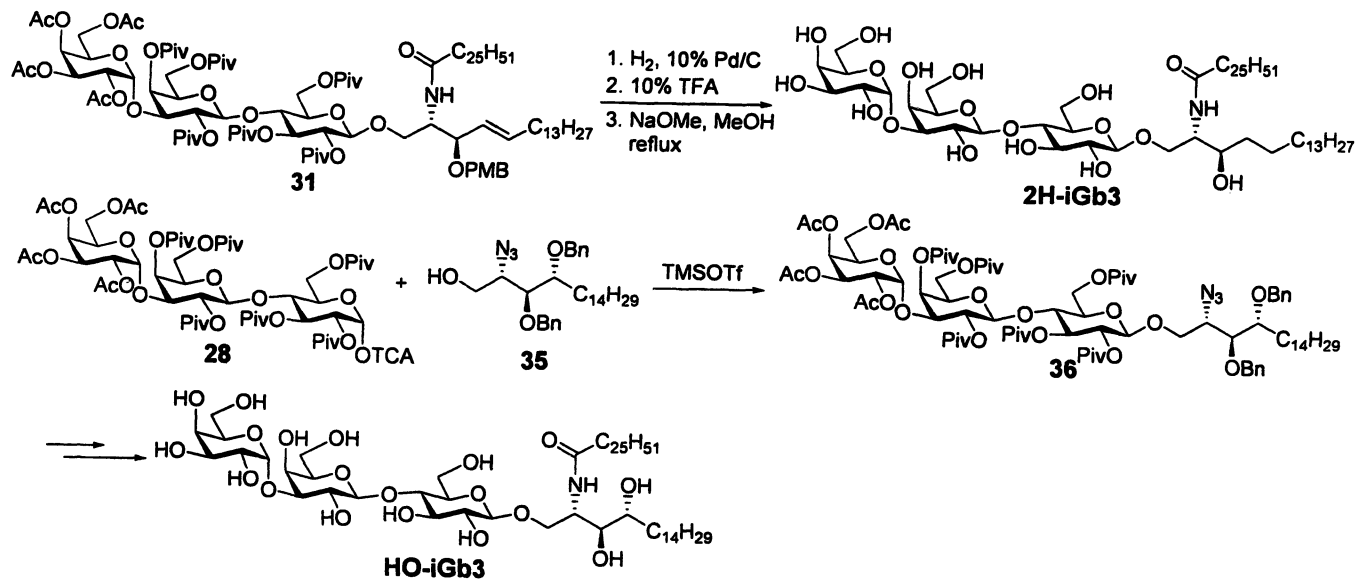
Structure Activity Relationship Studies of iGb3

Using the protocols of SAR studies of α -GalCer on the ceramide part, several iGb3 analogues with different ceramide part were also under investigation. The *D*-erythro-sphingosine of iGb3 contains a double-bond between C4 and C5, while the phytosphingosine of α -GalCer lacks it. Moreover, the phytosphingosine has an extra hydroxy group on C4. The differences between the phytosphingosine of α -GalCer and the *D*-erythro-sphingosine of

iGb3 may affect the ability of iGb3 to stimulate the NKT cells. For example, the hydroxy group may form hydrogen-bonds with CD1d and thereby increase the stability of the glycolipid-CD1d complex. Alternatively, It may effect the orientation of the sugar head on the surface of CD1d for recognition by the T cell receptor (TCR) of NKT cells. To examine these hypothesis, the 2H-iGb3 was prepared by hydrogenation reduction of the double bond of compound **31** and removal of the protecting groups. On the other hand, the HO-iGb3 was synthesized by glycosylation of trisaccharide donor **28** with phytosphingosine acceptor **35**, instead of erythro-sphingosine **29**, to give glycolipid **36**. Using similar procedures to those previously mentioned it was finally converted to HO-iGb3 (**Scheme 6**).

iGb3 and its two analogues were incubated with hCD1d-transfected human B lymphoblastoid (C1R). Two hours later, CD1d-restricted NKT cells clone JS7 were added and were co-cultured for 36 hours. IFN- γ and IL-4 were quantified in cell culture supernatants by ELISA. The results showed that the 2H-iGb3 was slightly more active in stimulation of iNKT cells in low concentration relative to iGb3. At higher concentrations (e.g. 20 $\mu\text{g/mL}$), very efficient stimulation of iNKT cells was observed (**Figure 4**). These indicated that reduction of the double bond on the lipid to the saturated bond increased the flexibility of the glycolipids, which made 2H-iGb3 easier to adopt the appropriate confirmation to enter the F' pocket of CD1d. An impressive stimulation was observed when using HO-iGb3 to activate iNKT cells. Both Th1 and Th2 cytokines were released when human iNKT cells were stimulated with HO-iGb3 at a concentration of 10 ng/mL (100-fold lower concentration than iGb3). This may result from the greater stability of the HO-iGb3 and CD1d complex because the 4-OH of the phytosphingosine can form an additional hydrogen-bond with the F' pocket. The results are in agreement with the observation by Kronenberg, et al (64). They examined the interactions between the solubilized form of the TCR and CD1d-glycolipd complex by surface plasmon resonance. The equilibrium dissociation constants (K_D), based on TCR binding at equilibrium, are 0.35 μM for KRN7000 (same ceramide as HO-iGb3) and 1.12 μM for 4-deoxy-GalCer (same ceramide as 2H-iGb3).

The 2' and 3' hydroxy group of α -GalCer are essential for its activity in stimulation of NKT cells. X-ray structural studies showed that both of these two hydroxy groups were responsible for binding with Asp 151 of CD1d (65, 66). Since the flexibility of the terminal sugar of iGb3, no X-ray results were obtained to demonstrate the role of the hydroxy groups of iGb3 in maintaining its activity. The deoxy analogues are a powerful tool to illustrate the relationship between the hydroxy groups and the receptor. Four deoxy galactosyl donors **37-40** were used to synthesize the deoxy iGb3 analogues using the procedures in **Scheme 7**. Preliminary bioassay results showed that although the hydroxy groups of the terminal sugar of iGb3 have some effects in stimulating NKT cells, they were not as crucial as the hydroxy groups in the α -GalCer.



Scheme 6. Modification of the ceramide part of iGb3.

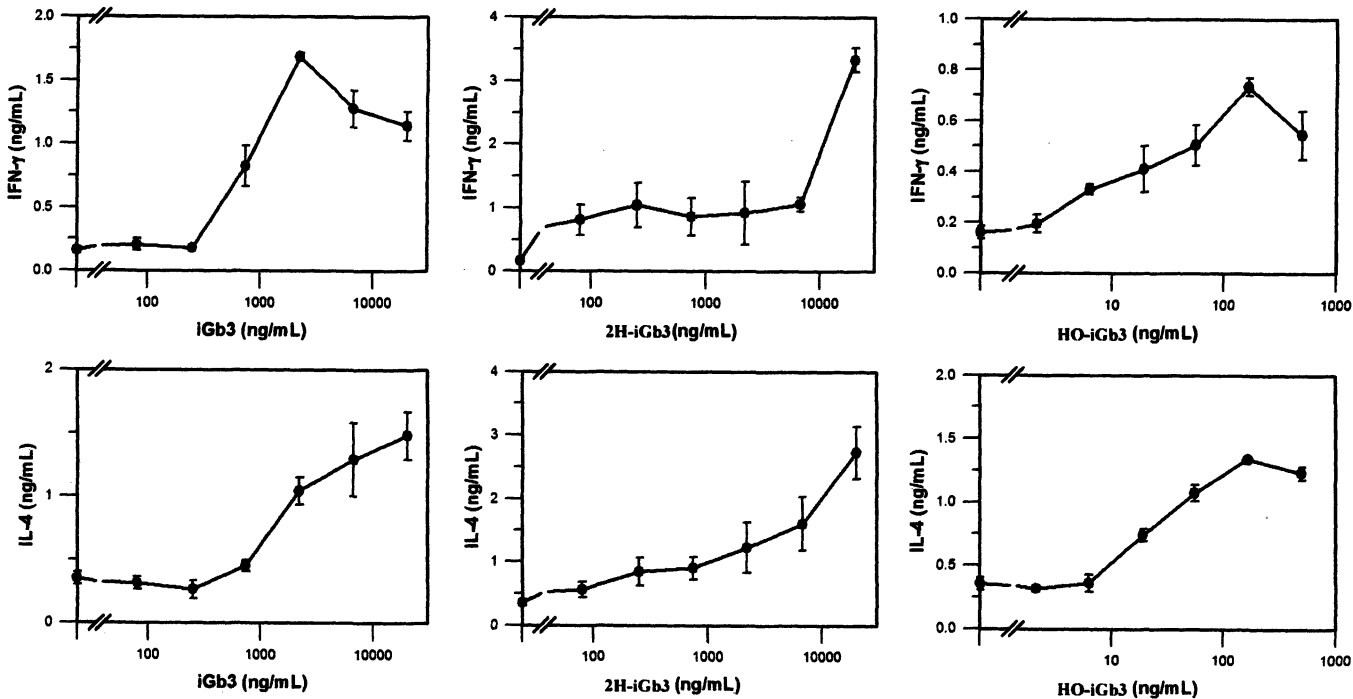
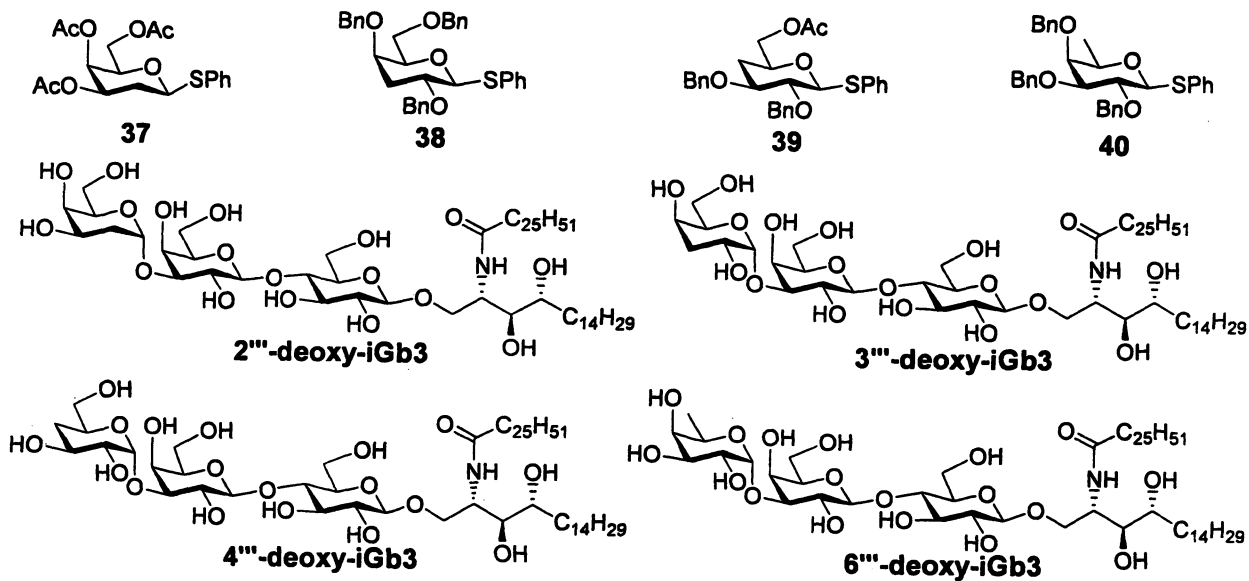


Figure 4. Activation of NKT cells by iGb3 analogues



Scheme 7. Deoxy galactosyl donors and terminal deoxy *i*Gb3 analogues

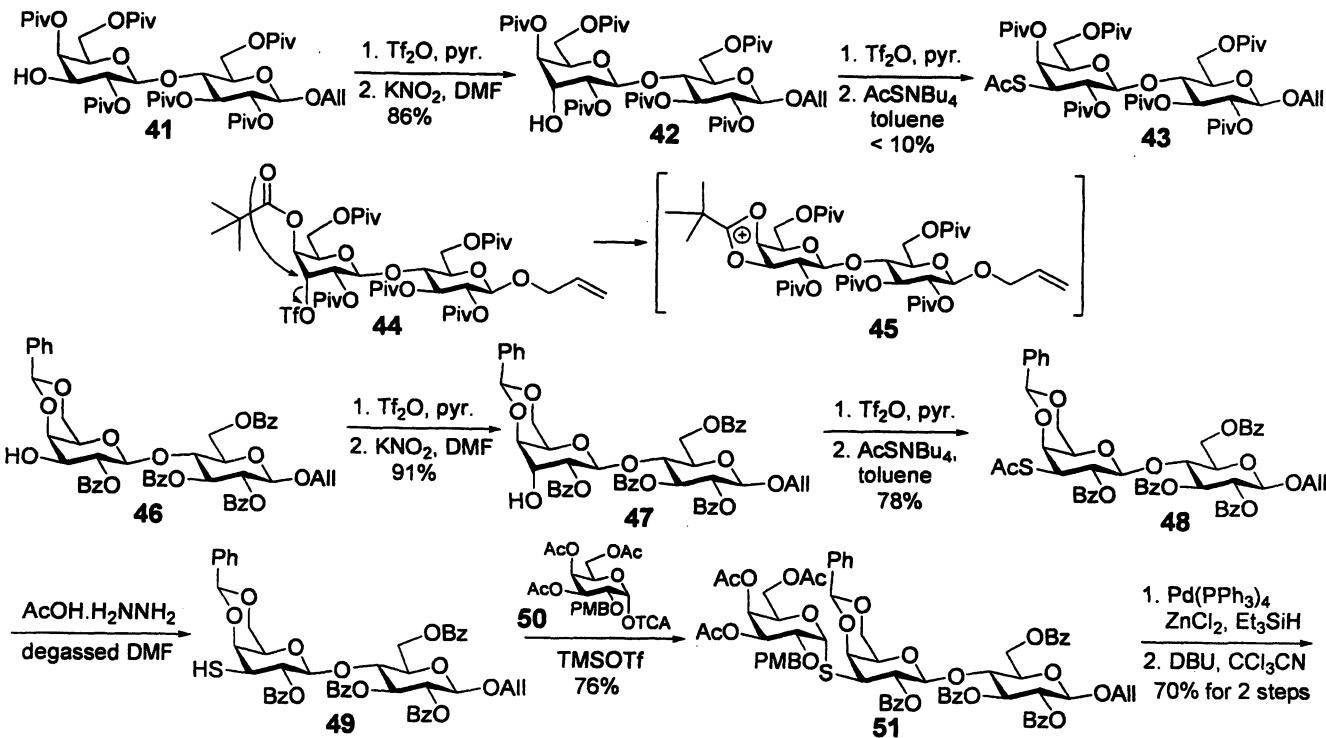
Metabolic Stability of iGb3 Analogues

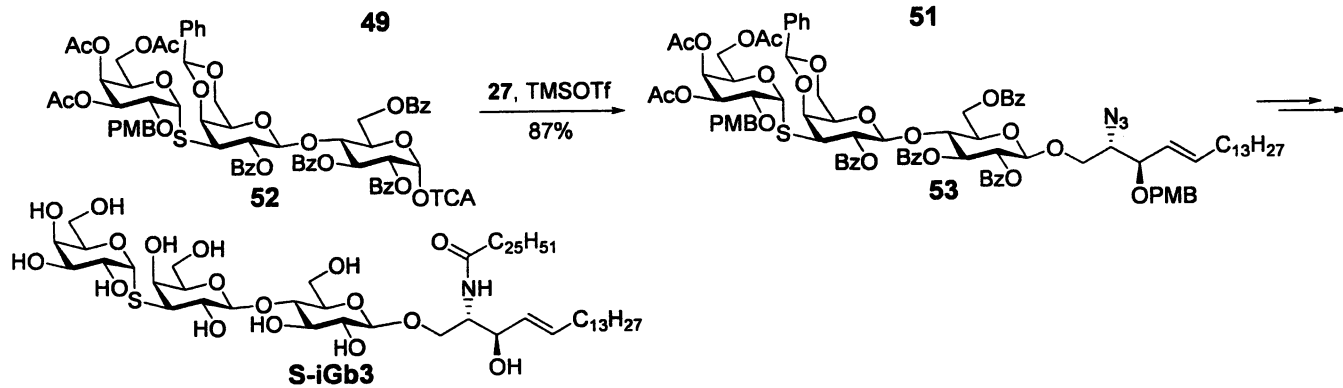
In the lysosome, the terminal GalNAc of iGb4 was hydrolyzed by β -hexosaminidase to afford iGb3 which can serve as an endogenous antigen for NKT cells. However, the iGb3 can be further hydrolyzed by α -galactosidase to give the β -lactosylceramide, which is not active in stimulating NKT cells. There is a competition between the iGb3 be loaded by CD1d and hydrolyzed by α -galactosidase. The metabolic stable iGb3 analogue will be resistant to α -galactosidase and therefore will increase its activity. Furthermore, it is well known that thioglycosides are resistant to glycosidase (67). Conformational studies of *S*-linked oligosaccharides (thio-oligosaccharides) have suggested that the flexibility about the anomeric linkage is increased, but that they still adopt a very similar conformational space as their *O*-linked counterparts (68). The *S*-iGb3 was prepared as described in Scheme 8 (69).

In order to get the original stereochemistry of the thio group, inversion of 3'-OH stereochemistry was required before the second inversion in the introduction of the thio group. Therefore, the equatorial 3'-OH of **41** was converted to an axial configuration by treatment with triflic anhydride followed by KNO_2 . Compound **42** was triflated again and 3'-*O*-triflate was substituted with AcSNBu_4 to give *S*-acetate **43** in an $\text{S}_{\text{N}}2$ manner, thus reforming the original stereochemistry (70). However, the yield was very low (<10%) and varied in every trial. We assumed that the C-4' hydroxyl derivative was formed because of the neighboring participation of the 4'-*O*-ester in compound **44** via the cyclic intermediate **45** (71).

To prevent this neighboring group participation, benzylidene was selected as the protecting group to protect the C-4' and C-6' hydroxyl groups. Following the same procedure as above, the thioester **48** was obtained in high yield. The acetyl was selectively removed by treatment with hydrazine acetate in degassed DMF to afford the thio-lactose derivative **49**. The glycosylation of thio **49** with donor **50** using TMSOTf as catalyst gave *S*-trisaccharide **51** in good yield (Scheme 8). The allyl group was removed by Pd $(\text{PPh}_3)_4$ in the presence of ZnCl_2 and Et_3SiH (72) followed by treatment with CCl_3CN and DBU to afford the donor **52** (73). Following similar procedures as in the preparation of iGb3, the lipid **29** was glycosylated and the acyl chain was introduced. Removal of the protecting groups afforded the *S*-iGb3.

The bioassay using mouse (C57BL6) bone marrow derived dendritic cells showed that the *S*-iGb3, which is similar to iGb3, stimulates NKT cell lines at a concentration above 10 ng/mL. A 10 fold increase was observed for *S*-iGb3 compared to iGb3 (Figure 5). The results indicated that the conformational change of *S*-iGb3 does not abolish the NKT cell recognition. The thioglycoside is resistant to α -galactosidases in the antigen presenting cells to be hydrolyzed,





Scheme 8. Synthesis of S-iGb3 through reinversion protocol.

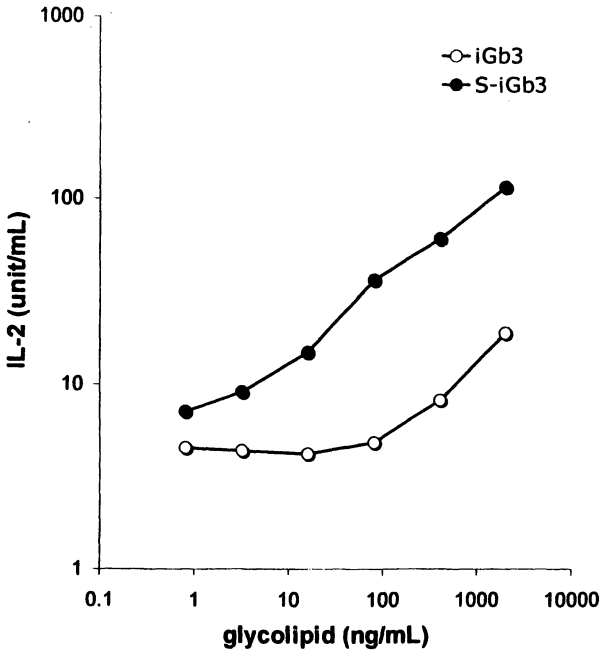


Figure 5. *S-iGb3* and *iGb3* were loaded by dendritic cells to stimulate DN32.D3 hybridoma cells

thus the *S-iGb3* has a longer retaining time in the lysosome before being loaded to CD1d. We have assumed that the degradation of *iGb3* requires the physical interaction of glycolipids with saposins, and *iGb3*/CD1d will be inaccessible to glycosidase cleavage.

Conclusion

The interactions among the CD1d/glycolipid/TCR complexes determine the profile of cytokines released by NKT cells. Different analogs of α -GalCer and *iGb3* were prepared by modification of both the sugar and lipid moieties. By introducing an additional hydroxyl group onto the lipid chain, the HO-*iGb3* can increase the affinity between CD1d and hence improve its activity. Substitution of the hydroxyl groups on the terminal sugar alters the recognition of TCR with glycolipids, thus resulting in the different profiles of cytokines produced. Studies are still undergoing to reveal the MHC-I like immuno response.

Acknowledgements

These projects are supported by The Ohio State University to Peng George Wang, M.D. Anderson Cancer Center to Dapeng Zhou and SNF (grant 3100A0-109918) to Gennaro De Libero.

References

1. Krummel, M. F.; Davis, M. M. *Cur. Opin. Immunol.* **2002**, *14*, 66.
2. Bromley, S. K.; Burack, W. R.; Johnson, K. G.; Somersalo, K.; Sims, T. N.; Sumen, C.; Davis, M. M.; Shaw, A. S.; Allen, P. M.; Dustin, M. L. *Annu. Rev. Immunol.* **2001**, *19*, 375.
3. Gao, G. F.; Jakobsen, B. K. *Immunol. Today* **2000**, *21*, 630.
4. Van Kaer, L. *Nat. Rev. Immunol.* **2005**, *5*, 31.
5. Batuwangala, T.; Shepherd, D.; Gadola Stephan, D.; Gibson Kevin, J. C.; Zaccai Nathan, R.; Fersht Alan, R.; Besra Gurdyal, S.; Cerundolo, V.; Jones, E. Y. *J. Immunol.* **2004**, *172*, 2382.
6. Zajonc, D. M.; Elsliger, M. A.; Teyton, L.; Wilson, I. A. *Nat. Immunol.* **2003**, *4*, 808.
7. Gadola, S. D.; Zaccai, N. R.; Harlos, K.; Shepherd, D.; Castro-Palomino, J. C.; Ritter, G.; Schmidt, R. R.; Jones, E. Y.; Cerundolo, V. *Nat. Immunol.* **2002**, *3*, 721.
8. Martin, L. H.; Calabi, F.; Lefebvre, F. A.; Bilsland, C. A. G.; Milstein, C. *Proc. Nat. Acad. Sci. USA* **1987**, *84*, 9189.
9. Martin, L. H.; Calabi, F.; Milstein, C. *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 9154.
10. Calabi, F.; Milstein, C. *Nature* **1986**, *323*, 540.
11. Calabi, F.; Jarvis, J. M.; Martin, L.; Milstein, C. *Eur. J. Immunol.* **1989**, *19*, 285.
12. Parekh, V. V.; Wilson, M. T.; Van Kaer, L. *Rev. Immunol.* **2005**, *25*, 183.
13. Bendelac, A.; Rivera, M. N.; Park, S.-H.; Roark, J. H. *Ann. Rev. Immunol.* **1997**, *15*, 535.
14. Kronenberg, M.; Gapin, L. *Nat. Rev. Immunol.* **2002**, *2*, 557.
15. Sharif, S.; Arreaza, G. A.; Zucker, P.; Mi, Q.-S.; Delovitch, T. L. *J. Mol. Med.* **2002**, *80*, 290.
16. Natori, T.; Koezuka, Y.; Higa, T. *Tetrahedron Lett.* **1993**, *34*, 5591.
17. Natori, T.; Morita, M.; Akimoto, K.; Koezuka, Y. *Tetrahedron* **1994**, *50*, 2771.
18. Morita, M.; Motoki, K.; Akimoto, K.; Natori, T.; Sakai, T.; Sawa, E.; Yamaji, K.; Koezuka, Y.; Kobayashi, E.; Fukushima, H. *J. Med. Chem.* **1995**, *38*, 2176.

19. Shimosaka, A. *Int. J. Hematol.* **2002**, 76 Suppl 1, 277.
20. Naumov, Y. N.; Bahjat, K. S.; Gausling, R.; Abraham, R.; Exley, M. A.; Koezuka, Y.; Balk, S. B.; Strominger, J. L.; Clare-Salzer, M.; Wilson, S. B. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, 98, 13838.
21. Taniguchi, M.; Seino, K.-i.; Nakayama, T. *Nature Immunol.* **2003**, 4, 1164.
22. Illes, Z.; Kondo, T.; Newcombe, J.; Oka, N.; Tabira, T.; Yamamura, T. *J. Immunol.* **2000**, 164, 4375.
23. Sharif, S.; Arreaza, G. A.; Zucker, P.; Mi, Q.-S.; Sondhi, J.; Naidenko, O. V.; Kronenberg, M.; Koezuka, Y.; DeLovitch, T. L.; Gombert, J.-M.; Leite-de-Moraes, M.; Gouarin, C.; Zhu, B.; Hameg, A.; Nakayama, T.; Taniguchi, M.; LePault, F.; Lehuen, A.; Bach, J.-F.; Herbelin, A. *Nat. Med.* **2001**, 7, 1057.
24. Jahng, A. W.; Maricic, I.; Pedersen, B.; Burdin, N.; Naidenko, O.; Kronenberg, M.; Koezuka, Y.; Kumar, V. *J. Exp. Med.* **2001**, 194, 1789.
25. Miyamoto, K.; Miyake, S.; Yamamura, T. *Nature* **2001**, 413, 531.
26. Seino, K.-I.; Fukao, K.; Muramoto, K.; Yanagisawa, K.; Takada, Y.; Kakuta, S.; Iwakura, Y.; Van Kaer, L.; Takeda, K.; Nakayama, T.; Taniguchi, M.; Bashuda, H.; Yagita, H.; Okumura, K. *Proc. Natl. Acad. Sci. USA* **2001**, 98, 2577.
27. Ostos, M. A.; Recalde, D.; Zakin, M. M.; Scott-Algara, D. *FEBS Lett.* **2002**, 519, 23.
28. Hammond, K. J. L.; Godfrey, D. I. *Tissue Antigens* **2002**, 59, 353.
29. Smyth, M. J.; Crowe, N. Y.; Hayakawa, Y.; Takeda, K.; Yagita, H.; Godfrey, D. I. *Cur. Opin. Immunol.* **2002**, 14, 165.
30. Smyth, M. J.; Godfrey, D. I. *Nat. Immunol.* **2000**, 1, 459.
31. Swann, J.; Crowe, N. Y.; Hayakawa, Y.; Godfrey, D. I.; Smyth, M. J. *Immunol. Cell Bio.* **2004**, 82, 323.
32. Hachem, P.; Lisbonne, M.; Michel, M.-L.; Diem, S.; Roongapinun, S.; Lefort, J.; Marchal, G.; Herbelin, A.; Askenase, P. W.; Dy, M.; Leite-de-Moraes, M. C. *Eur. J. Immunol.* **2005**, 35, 2793.
33. Lack, G.; Bradley, K. L.; Hamelmann, E.; Renz, H.; Loader, J.; Leung, D. Y. M.; Larsen, G.; Gelfand, E. W. *J. Immunol.* **1996**, 157, 1432.
34. Walter, D. M.; Wong, C. P.; DeKruyff, R. H.; Berry, G. J.; Levy, S.; Umetsu, D. T. *J. Immunol.* **2001**, 166, 6392.
35. Zhou, D.; Mattner, J.; Cantu, C., 3rd; Schrantz, N.; Yin, N.; Gao, Y.; Sagiv, Y.; Hudspeth, K.; Wu, Y.-P.; Yamashita, T.; Teneberg, S.; Wang, D.; Proia Richard, L.; Levery Steven, B.; Savage Paul, B.; Teyton, L.; Bendelac, A., *Science* **2004**, 306, 1786.
36. Sandrin, M. S. **2002**, UA2002081688.
37. Conzelmann, E.; Sandhoff, K. *Adv. Exp. Med. Bio.* **1980**, 125, 295.
38. Keusch, J. J.; Manzella, S. M.; Nyame, K. A.; Cummings, R. D.; Baenziger, J. U. *J. Bio. Chem.* **2000**, 275, 25315.

39. Zhou, D.; Cantu, C., III; Sagiv, Y.; Schrantz, N.; Kulkarni, A. B.; Qi, X.; Mahuran, D. J.; Morales, C. R.; Grabowski, G. A.; Benlagha, K.; Savage, P.; Bendelac, A.; Teyton, L. *Science* **2004**, *303*, 523.
40. Lin, M.; Rikihisa, Y. *Infect. Immunol.* **2003**, *71*, 5324.
41. Mattner, J.; DeBord, K. L.; Ismail, N.; Goff, R. D.; Cantu, C.; Zhou, D.; Saint-Mezard, P.; Wang, V.; Gao, Y.; Yin, N.; Hoebe, K.; Schneewind, O.; Walker, D.; Beutler, B.; Teyton, L.; Savage, P. B.; Bendelac, A. *Nature* **2005**, *434*, 525.
42. Kinjo, Y.; Wu, D.; Kim, G.; Xing, G.-W.; Poles, M. A.; Ho, D. D.; Tsuji, M.; Kawahara, K.; Wong, C.-H.; Kronenberg, M. *Nature* **2005**, *434*, 520.
43. Yu, K. O. A.; Im, J. S.; Molano, A.; Dutronc, Y.; Illarionov, P. A.; Forestier, C.; Fujiwara, N.; Arias, I.; Miyake, S.; Yamamura, T.; Chang, Y.-T.; Besra, G. S.; Porcelli, S. A. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 3383.
44. Schmieg, J.; Yang, G.; Franck, R. W.; Tsuji, M. *J. Exp. Med.* **2003**, *198*, 1631.
45. Yang, G.; Schmieg, J.; Tsuji, M.; Franck, R. W. *Angew. Chem., Int. Ed.* **2004**, *43*, 3818.
46. Kawano, T.; Cui, J.; Koezuka, Y.; Toura, I.; Kaneko, Y.; Motoki, K.; Ueno, H.; Nakagawa, R.; Sato, H.; Kondo, E.; Koseki, H. *Science* **1997**, *278*, 1626.
47. Iijima, H.; Kimura, K.; Sakai, T.; Uchimura, A.; Shimizu, T.; Ueno, H.; Natori, T.; Koezuka, Y. *Bio. Med. Chem.* **1998**, *6*, 1905.
48. Zeng, Z. H.; Castano, A. R.; Segelke, B. W.; Stura, E. A.; Peterson, P. A.; Wilson, I. A. *Science* **1997**, *277*, 339.
49. Wu, D.; Xing, G.-W.; Poles, M. A.; Horowitz, A.; Kinjo, Y.; Sullivan, B.; Bodmer-Narkevitch, V.; Plettenburg, O.; Kronenberg, M.; Tsuji, M.; Ho, D. D.; Wong, C.-H. *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 1351.
50. Zhou, X.-T.; Forestier, C.; Goff, R. D.; Li, C.; Teyton, L.; Bendelac, A.; Savage, P. B. *Org. Lett.* **2002**, *4*, 1267.
51. Prigozy, T. I.; Naidenko, O.; Qasba, P.; Elewaut, D.; Brossay, L.; Khurana, A.; Natori, T.; Koezuka, Y.; Kulkarni, A.; Kronenberg, M. *Science* **2001**, *291*, 664.
52. Xia, C.; Yao, Q.; Schuemann, J.; Rossy, E.; Chen, W.; Zhu, L.; Zhang, W.; De Libero, G.; Wang, P. G. *Bio. Med. Chem. Lett.* **2006**, *16*, 2195.
53. Ndonye, R. M.; Izmirian, D. P.; Dunn, M. F.; Yu, K. O. A.; Porcelli, S. A.; Khurana, A.; Kronenberg, M.; Richardson, S. K.; Howell, A. R. *J. Org. Chem.* **2005**, *70*, 10260.
54. Goff, R. D.; Gao, Y.; Mattner, J.; Zhou, D.; Yin, N.; Cantu, C., III; Teyton, L.; Bendelac, A.; Savage, P. B. *J. Am. Chem. Soc.* **2004**, *126*, 13602.
55. Lowary, T. L.; Hindsgaul, O. *Carbohydr. Res.* **1994**, *251*, 33.
56. Janczuk, A. J.; Zhang, W.; Andreana, P. R.; Warrick, J.; Wang, P. G. *Carbohydr. Res.* **2002**, *337*, 1247.

57. Zhang, W.; Wang, J.; Li, J.; Yu, L.; Wang, P. G. *J. Carbohydr. Chem.* **1999**, *18*, 1009.
58. Rai, A. N.; Basu, A. *Org. Lett.* **2004**, *6*, 2861.
59. Ichikawa, Y.; Liu, J. L. C.; Shen, G. J.; Wong, C. H. *J. Am. Chem. Soc.* **1991**, *113*, 6300.
60. Wang, P.; Shen, G. J.; Wang, Y. F.; Ichikawa, Y.; Wong, C. H. *J. Org. Chem.* **1993**, *58*, 3985.
61. Chen, X.; Fang, J.; Zhang, J.; Liu, Z.; Shao, J.; Kowal, P.; Andreana, P.; Wang, P. G. *J. Am. Chem. Soc.* **2001**, *123*, 2081.
62. Zhang, J.; Kowal, P.; Chen, X.; Wang, P. G. *Org. Biomol. Chem.* **2003**, *1*, 3048.
63. Yao, Q.; Song, J.; Xia, C.; Zhang, W.; Wang, P. G. *Org. Lett.* **2006**, *8*, 911.
64. Sidobre, S.; Hammond, K. J. L.; Benazet-Sidobre, L.; Maltsev, S. D.; Richardson, S. K.; Ndonge, R. M.; Howell, A. R.; Sakai, T.; Besra, G. S.; Porcelli, S. A.; Kronenberg, M. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 12254-12259.
65. Koch, M.; Stronge, V. S.; Shepherd, D.; Gadola, S. D.; Mathew, B.; Ritter, G.; Fersht, A. R.; Besra, G. S.; Schmidt, R. R.; Jones, E. Y.; Cerundolo, V. *Nat. Immunol.* **2005**, *6*, 819.
66. Zajonc, D. M.; Cantu, C.; Mattner, J.; Zhou, D.; Savage, P. B.; Bendelac, A.; Wilson, I. A.; Teyton, L. *Nat. Immunol.* **2005**, *6*, (8), 810-818.
67. Albrecht, B.; Puetz, U.; Schwarzmann, G. *Carb. Res.* **1995**, *276*, 289.
68. Geyer, A.; Hummel, G.; Eisele, T.; Stefan, R.; Schjmidt, R. R. *Chem.-Eur. J.* **1996**, *2*, 981.
69. Xia, C.; Zhou, D.; Liu, C.; Lou, Y.; Yao, Q.; Zhang, W.; Wang, P. G. *Org. Lett.* **2006**, *8*, 5493.
70. Rich, J. R.; Szpacenko, A.; Palcic, M. M.; Bundle, D. R. *Angew. Chem., Int. Ed.* **2004**, *43*, 613.
71. Liakatos, A.; Kiefel, M. J.; Von Itzstein, M. *Org. Lett.* **2003**, *5*, 4365.
72. Eisele, T.; Windmuller, R.; Schmidt, R. R. *Carbohydr. Res.* **1998**, *306*, 81.
73. Maruyama, M.; Takeda, T.; Shimizu, N.; Hada, N.; Yamada, H. *Carbohydr. Res.* **2000**, *325*, 83.

Chapter 9

Clinical Outcome of Breast and Ovarian Cancer Patients Treated after High-Dose Chemotherapy and Autologous Stem Cell Rescue with Theratope (STn-KLH) Cancer Vaccine

Theratope (STn-KLH) Cancer Vaccine following Autologous Transplant

Leona A. Holmberg^{1,2}, Katherine A. Guthrie^{1,3},
and Brenda M. Sandmaier^{1,2}

¹Clinical Research Division, Fred Hutchinson Cancer Research Center,
1100 Fairview Avenue North, D5-390, Seattle, WA 98109-1024

²Department of Medicine, University of Washington School of Medicine,
Seattle, WA 98195

³Department of Biostatistics, School of Public Health and Community
Medicine, University of Washington, Seattle, WA 98195

The success of autologous peripheral blood stem cell transplantation (ASCT) as treatment for breast and ovarian cancer is limited by high relapse rates. Following ASCT, patients have lower tumor burden, may have decreased number of regulatory cells that can suppress ability to mount immune response and have lower amounts of circulating immunosuppressive tumor-associated molecules like MUC1 that cause anergy. As a consequence, after ASCT patients are more likely to respond immunologically to a cancer vaccine. ASCT thus provides a platform for delivering effective immunotherapy. We speculated that the successful delivery of immunotherapy after ASCT might decrease relapse rates and prolong survival. Theratope (Sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine incorporates a synthetic STn antigen that mimics the unique tumor-associated STn carbohydrate on MUC1 found on breast and ovarian cancer

cells. From 1995 through 2000, 70 patients (16 with stage II/III breast cancer, 17 with stage III/IV ovarian cancer and 37 with stage IV breast cancer) were treated after ASCT with two different formulations of Theratope given with Detox B. Theratope was generally well tolerated with minimal toxicity. Both humoral and cellular responses were able to be measured in the laboratory in the vaccinated patients. In this chapter, we report on the long term outcome of the Theratope vaccinated ASCT patients in correlation with measurable immune responses.

Introduction

The success of ASCT as treatment for breast and ovarian cancer is limited by high relapse rates (1-6). Relapses result from infusion of tumor contaminated stem cell products or from incomplete eradication of endogenous tumor cells. Higher doses of chemotherapy only increase morbidity and mortality. Thus, studying the addition of alternative therapy options such as immunotherapy after ASCT seems warranted. Vaccination against unique tumor associated molecule is a novel way to treat cancer. In a non-transplant setting, vaccination strategies, though, are problematic because of immune incompetence. We speculated that one way to overcome this immune incompetence was to vaccinate patients early after ASCT. Following ASCT, patients have lower tumor burden and thus have lower amounts of circulating immunosuppressive molecules that can cause immune anergy. In addition, ASCT patients may have lower levels of active suppressor regulatory cells, especially as most patients receive Cyclophosphamide to mobilize their peripheral blood stem cell product. We felt that patients would respond to a vaccine early after ASCT as they were not that immunologically inactive. In fact early after ASCT, our group has previously shown immunity is reestablishing itself as manifested by breaking of tolerance seen with the development of pseudo autologous GVHD syndrome (7, 8).

Sialyl-Tn (STn) is a unique carbohydrate expressed tumor antigen on the core region of the epithelial MUC1 that is expressed on up to eighty percent of breast and ovarian cancer cells (9-11). STn is defined by the structure NANA alpha (2-6) GaINAC. Expression of STn is associated with poor prognosis and is important in growth and metastasis (12, 13). Animal studies have shown that immunization with STn vaccine slows tumor growth and prolongs survival (14, 15). Consequently, STn appeared to be an ideal candidate to incorporate into a vaccine to deliver immunotherapy to ASCT patients. Theratope (STn-KLH) cancer vaccine consists of a chemically synthesized carbohydrate that emulates the known unique tumor-associated carbohydrate determinant STn on human adenocarcinoma cells.

When we started our studies with vaccination of ASCT patients with Theratope there was data from a non-transplant setting that the Theratope vaccine was well tolerated with minimal side effects (16). But, it was controversial as to whether ASCT patients could mount an immune response against the vaccine. The initial immunological responses in ASCT patients has been previously published (17, 18). Both humoral and cellular immunity were seen as determined by antibody to OSM (ovine submaxillary mucin, a natural STn) and to STn, T cell proliferation to STn, gamma interferon production and lysis of tumor cell lines including the STn- bearing OVCAR cells. In this chapter, we update the outcome of the seventy ASCT breast and ovarian cancer patients vaccinated with Theratope and correlate their long term outcome with in vitro immunological responses.

Methods

Study Design

Between September 1995 and November 2000, 70 patients were vaccinated after autologous peripheral blood stem cell transplantation (ASCT) with one of two formulations of Theratope (Biomira, Edmonton, Alberta, Canada) coupled with Detox B SE (Detox B) (Corixa, Hamilton, MT). The difference between the two formulations was that the later one has an increased ratio of STn conjugated to KLH compared to the first formulation. The vaccination schedule for the two different formulations was previously published as well as medical treatment history of ASCT patients (18).

For patients' characteristics, see Table I. Patients were vaccinated if they had Karnofsky performance score $\geq 70\%$, had recovered from acute toxicity of ASCT, had $ANC > 1.5 \times 10^9/L$ and platelet count (untransfused) $> 20 \times 10^9/L$, had total bilirubin or SGOT $< 2 \times$ upper limit of normal (ULN), alkaline phosphatase $< 3 \times$ ULN and creatinine ≤ 2 mg/dl. Patients were ineligible for vaccine if they were receiving steroid therapy or had evidence of an active autoimmune disease. Patients who were allergic to shellfish or had history of anaphylaxis to any drug or history of splenectomy were excluded. After obtaining informed consent, patients were treated per protocol previously approved by FHCRC Institutional Review Board under IND from US Food and Drug Administration. Simultaneous hormone therapy, surgery and radiation therapy were allowed.

All patients were monitored for toxicity and grading per WHO toxicity criteria. Overall, the vaccine was well tolerated and toxicity has been previously published (18, 19). Most common toxicities seen were local reaction at site of injection including mainly erythema and induration and flu like symptoms. The median time to being vaccinated after ASCT with first formulation of Theratope tested was 127.5 days; the median time for the second formulation of Theratope was 103 days.

Table I. Patient characteristics by treatment group in breast and ovarian cancer ASCT patients

	Cohort		
	Vaccine 1	Vaccine 2	Control
All patients			
Total N	40	30	99
Relapse risk: N (%)			
Low	10 (25)	5 (17)	57 (58)
Intermediate	7 (17)	17 (57)	35 (35)
High	23 (58)	8 (27)	7 (7)
Diagnosed with SMN: N (%)	2 (5)	1 (3)	2 (2)
Follow-up in survivors by relapse risk (years) : median			
Low	9.2	7.0	7.6
Intermediate	9.1	6.0	7.2
High	8.8	6.7	5.9
Breast cancer ASCT patients			
Total N	33	20	92
Relapse risk: N (%)			
Low	10 (30)	5 (25)	57 (62)
Intermediate	5 (15)	8 (40)	30 (33)
High	18 (55)	7 (35)	5 (5)
Hormone receptors: N (%)			
E-/P-	10 (30)	5 (25)	27 (29)
E+/P+	16 (49)	12 (60)	44 (48)
E-/P+	2 (6)	1 (5)	8 (9)
E+/P-	2 (6)	0	10 (11)
E+/unknown	1 (3)	0	1 (1)
Unknown	2 (6)	2 (10)	2 (2)
Follow-up in survivors by relapse risk (years) :median			
Low	9.2	7.0	7.6
Intermediate	9.1	6.0	7.1
High	8.5	6.7	5.9

NOTE: SMN is secondary malignancy.

Because there were a number of drugs to study in combination with ASCT, the outcome of a control group of patients were also tracked. The control group were a group of patients that underwent ASCT during the same time period as the vaccinated patients and were eligible to receive the vaccine but did not get treated with it. They were not vaccinated at their own request, had no insurance coverage or were not referred by their oncologist for participation in the study. On this study, tumor cells from vaccinated and control patients were not stained for the expression of STn.

Immune Responses In Vitro

Antibody Testing Against OSM or STn

Blood was collected from patients and clarified. Sera were stored at -80 °C. IgM and IgG titers to OSM or STn were evaluated in serum samples. These antigens were used as the target in solid phase enzyme-linked immunoabsorbant assays before and after treatment with Theratope as previously described (17).

Proliferation Assay

Peripheral blood mononuclear cells were obtained by density gradient separation through Ficol-Hypaque technology (17). Thymidine incorporation proliferation assays were set up as previously described using human serum albumin (HSA) or STn-HSA as stimulant (17). A stimulation index of > 2 SD compared to normal donors mononuclear cells was used as evidence of positive response. Fifty seven percent of tested vaccinated patients had evidence of specific T cell proliferation against STn.

Cytotoxic Assay

OVCAR is a human ovarian cell line that expresses STn on the cell surface. OVCAR was obtained from American Type Culture Collection (Bethesda, MD), grown and then placed in standard ⁵¹chromium release assay (17). Mononuclear cells from patients were incubated in the presence or absence of Interleukin 2 (IL2) (Genzyme, Cambridge, MA) at a concentration of 10⁶ cells/ml with 100 units of IL2 for 48 hours -72 hours. Cells were washed and used as effector cells in a standard four hour killing assay as previously described (17). Fifty three

percent of the tested vaccinated patients had evidence of > 20% increase in killing against STn-positive cells compared to STn-negative cells lines.

Definitions

Patients were grouped for comparison into three classes: low, intermediate and high risk for relapse. The low risk group included high risk stage II/III (non-inflammatory breast cancer). The intermediate risk group included inflammatory breast cancer stage IIIB, platinum sensitive stage III/IV ovarian and stage IV chemoresponsive breast cancer. The high risk group included chemorefractory stage IV breast cancer and platinum refractory stage III/IV ovarian cancer. Patients who relapsed or progressed after six months from platinum based regimens were considered to have platinum sensitive disease. Patients were designated as having chemoresponsive stage IV breast cancer if they were placed into a complete remission (CR) or partial remission (PR) disease status after first therapy given for metastatic disease. Progression of disease was defined as a > 25 % increase in measurable disease or development of new lesions. Patients who achieved a PR after ASCT were considered to relapse at first documentation of disease progression.

Statistical Considerations

The distributions of immune function parameters were compared across treatment groups via t-test. Kaplan Meier estimates were used to summarize overall survival (OS) and event free survival (EFS). Survival was calculated from PBSC infusion until death or date of last contact. EFS is defined as the absence of death, relapse, disease progression and de novo secondary cancer. Cox regression models were fitted for the outcomes of OS and EFS. P values associated with regression models were derived from the Wald test. The last possible day of contact was February 21, 2006.

All treatment group comparisons were performed first in all patients and then in breast cancer patients only. We also considered the subsets of breast cancer patients defined by low relapse risk and estrogen receptor status, although the sample sizes of these groups precluded adjusted analyses.

The associations of immune function parameters with EFS and OS in vaccinated breast cancer patients were estimated via Cox regression models. The immune response results were split in groups by the median values, with the medians determined separately for each vaccine group, and the resulting categories combined. The potential for differential effects of immune response and vaccine formulation across relapse risk groups was tested using interaction terms.

Results

Outcome of All Patients

Event free survival (EFS) and overall survival (OS) of all patients is described by treatment group and relapse risk (Figures 1 and 2, respectively). For these analyses, patients vaccinated with either of the two different formulations of Theratope have been combined. For the whole cohort, there was no statistically significant difference in EFS between control and vaccinated patients after adjusting for risk group.

Since the majority of ASCT patients vaccinated had breast cancer, we focused our evaluation on those patients. Among the breast cancers only, the vaccinated patients had a marginally statistically significant reduced risk of death compared to control patients after adjusting for risk of relapse (HR =0.5, 95% CI 0.3-1.0, $p=.05$). The vaccinated breast cancer patients with low risk for relapse and estrogen receptor (ER) negative disease had the highest EFS rate relative to control groups of breast cancer patients ($p=.02$) (Figures 3a and b).

Immune Function Testing

For a listing of the results of immune testing of all vaccinated and breast cancer alone patients, see Table IIa and b.

For all patients, the IgG antibody titer against STn was statistically significantly higher after vaccination with the second formulation compared to the first formulation ($p=.004$). The IgG antibody titer against OSM and OVCAR killing without IL2 were marginally statistically significantly higher after vaccination with the first formulation than compared to the second formulation ($p=.05$ and $p=.06$, respectively).

Among breast cancer patients alone, the IgG antibody titer to STn was statistically significantly higher after vaccination with the second formulation compared to the first formulation ($p=.009$). The IgG antibody titer to OSM was marginally statistically significant higher after vaccination with the second formulation compared to the first formulation ($p=.08$) and were thus used to evaluate outcome.

Factors associated with EFS in Vaccinated Breast Cancer Patients

Since breast cancer patients were the largest group of patients treated with the two different formulations of Theratope vaccine, we focused again on evaluating the outcome after vaccination as a function of in vitro immunological responses. Fifty-three patients were available for this analysis.

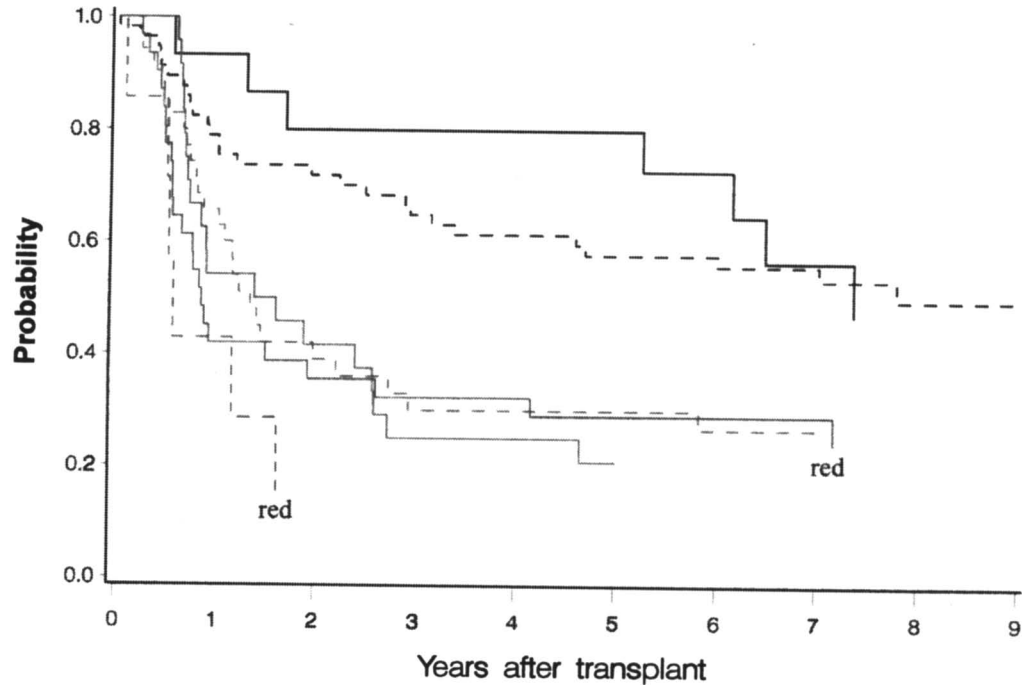
Event-free survival by relapse risk and treatment group in breast and ovarian cancer patients

Figure 1. Solid lines represent the vaccinated group, dashed lines represent the control group. Low risk patients are shown in black, intermediate risk in green and high risk in red.

Overall survival by relapse risk and treatment group in breast and ovarian cancer patients

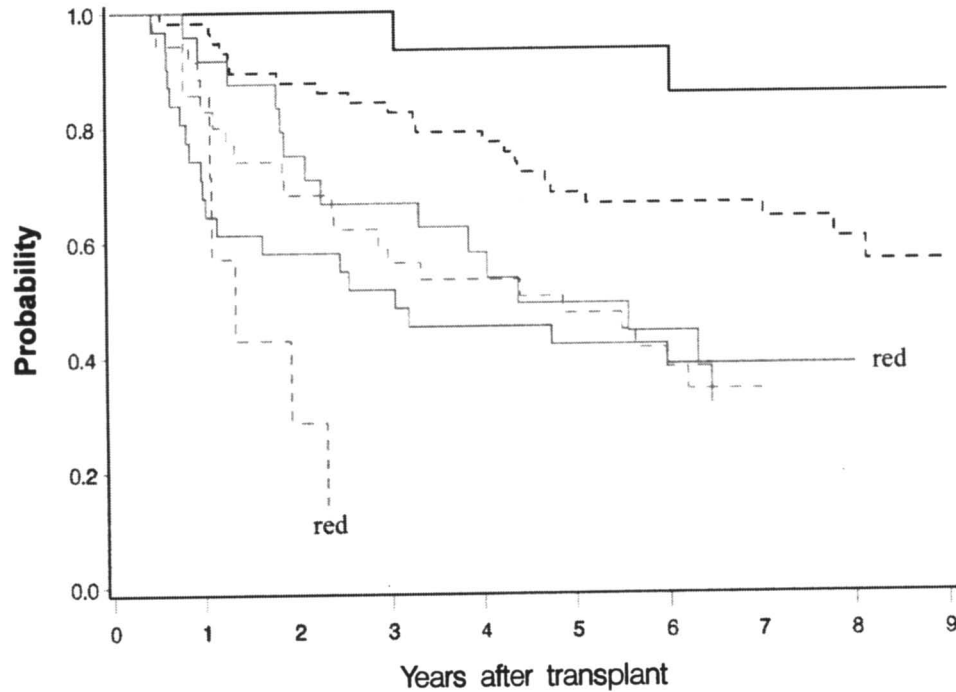


Figure 2. Solid lines represent the vaccinated group, dashed lines represent the control group. Low risk patients are shown in black, intermediate risk in green and high risk in red.

Event-free survival by estrogen receptor status and treatment group in low risk, breast cancer patients only

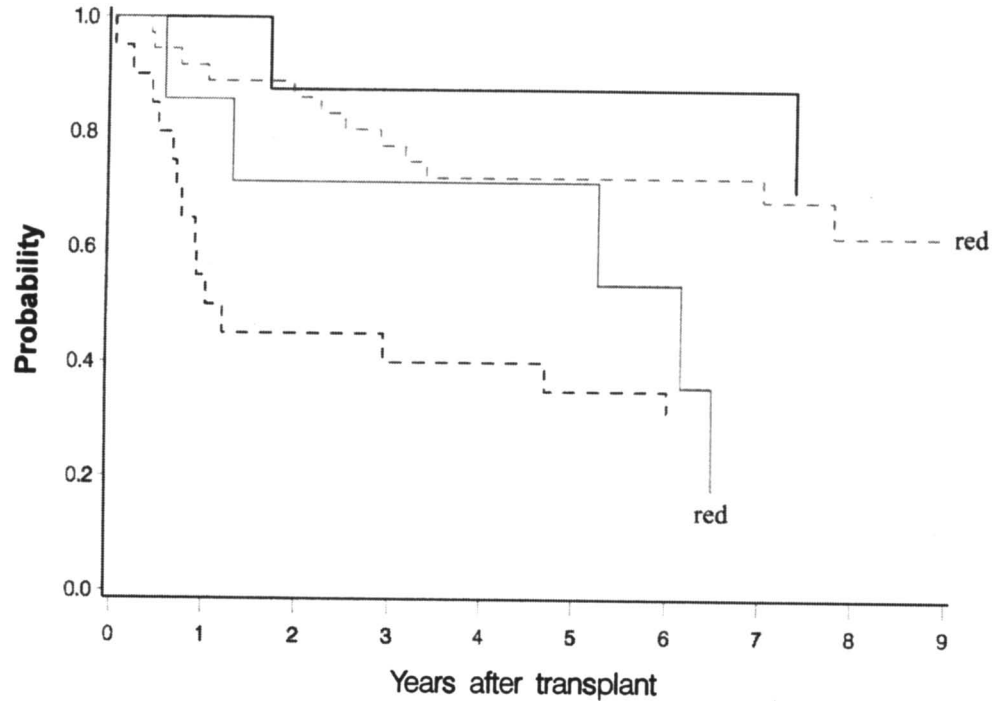


Figure 3a. Solid lines represent the vaccinated group, dashed lines represent the control group. Low risk ER negative patients are shown in black and low risk ER positive patients in red.

Overall survival by estrogen receptor status and treatment group in low risk, breast cancer patients only

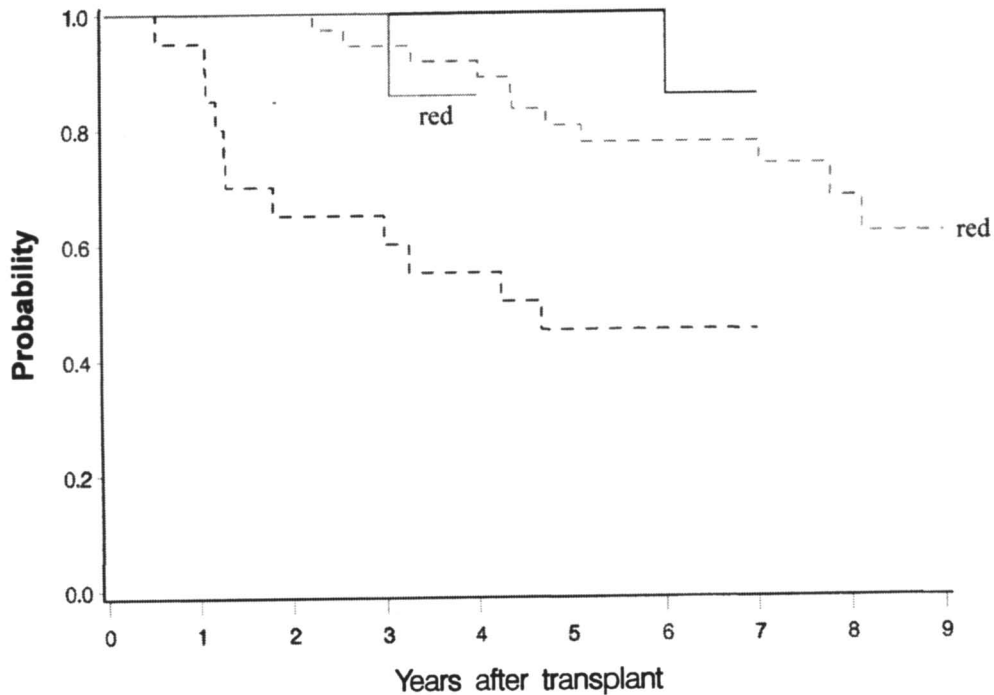


Figure 3b. Solid lines represent the vaccinated group, dashed lines represent the control group. Low risk ER negative patients are shown in black and low risk ER positive patients in red.

Table II.a Laboratory values by treatment group in breast and ovarian cancer ASCT patients

Laboratory value	Cohort	
	Vaccine 1	Vaccine 2
OSM IgG levels: N	36	29
median (range)	120 (0 – 1280)	640 (0 – 20480)
STN IgG levels: N	36	29
median (range)	640 (0 – 5120)	5120 (80 – 81920)
Proliferation index: N	36	30
median (range)	2.5 (0.9 – 25.0)	2.3 (1.5 – 34.0)
OVCAR killing with IL2: N	32	30
median (range)	46.5 (11.0 – 98.0)	48.0 (6.4 – 95.0)
OVCAR killing without IL2: N	32	30
median (range)	5.6 (0 – 70.0)	16.3 (4.0 – 68.0)

Table II.b Laboratory values by treatment group in breast cancer ASCT patients only

Laboratory value	Cohort	
	Vaccine 1	Vaccine 2
OSM IgG levels: N	31	19
median (range)	80 (0 – 1280)	640 (10 – 20480)
STN IgG levels: N	31	19
median (range)	640 (0 – 5120)	5120 (320 – 81920)
Proliferation index: N	31	20
median (range)	2.6 (0.9 – 25.0)	2.3 (1.5 – 23.3)
OVCAR killing with IL2: N	28	20
median (range)	45.6 (11.0 – 92.0)	48.0 (12.0 – 95.0)
OVCAR killing without IL2: N	28	20
median (range)	5.6 (0 – 70.0)	16.1 (4.0 – 68.0)

In an unadjusted analysis of EFS, patients with intermediate and high risk for relapse had a nearly three- fold increase risk of failure compared to patients in low risk group. Higher levels of IgG antibody to OSM and higher proliferation index to STn were also associated with reduced risks for failure compared to levels lower than the medians.

Table III describes the risk of failure by vaccine formulation and OSM IgG antibody levels and OVACAR killing immune response, separately for each relapse risk category. The combinations shown are those that resulted in interaction terms with $p < 0.15$. In most cases, the sample sizes are too small to produce significant subgroup results, but all comparisons suggest differential effects of factors across relapse risk levels. For example, the protective effect of OSM IgG antibody titer level on the risk of failure is found only in those patients with intermediate or high relapse risk. In low risk patients, OSM IgG antibody titer level is not associated with EFS.

In a multivariable model restricted to intermediate and high risk patients, the highest levels of OSM IgG antibody titer and OVACAR killing after IL2 stimulation were both significantly associated with superior EFS rates (HR=0.2, 95% CI 0.1-0.6 and HR=0.3, 95% CI 0.2-0.8, respectively).

Factors Associated with Survival in Vaccinated Breast Cancer Patients

As outlined in Table IV, in a multivariable analysis, patients with intermediate and high risk for relapse had greatly increased rates of death compared to lower risk patients. Recipients of the second formulation of Theratope vaccine had a lower risk of death than recipients of the first formulation of vaccine. Higher IgG titers of antibody to OSM was associated with lower rates of death. Proliferative index against STn and OVACAR killing with IL2 or without IL2 stimulation were not statically significantly different.

Discussion

After ASCT, patients were able to mount effective humoral and cellular immunity to the Theratope cancer vaccine. There were slightly different amplitudes of responses with the second formulation having the highest IgG antibody response to STn ($p = .009$) in breast cancer patients. In all the subgroups of patients treated, the patients with estrogen receptor negative stage II or III non-inflammatory breast cancer had the most benefit from Theratope vaccination compared to control group. The EFS rate was statistically significantly higher in the vaccinated than in the control group, $p = .02$ (Figure 3a). Since all patients with estrogen receptor positive disease after ASCT also received additional hormonal therapy this may have impacted on the ability to see any benefit in

Table III. Failure counts and percentages, with unadjusted hazard ratios (95% CI) for EFS in vaccinated breast cancer patients only (n=53)

<i>Characteristic</i>	<i>Relapse risk</i>					
	<i>Low</i>		<i>Intermediate</i>		<i>High</i>	
	<i>N (%)</i>	<i>HR (95% CI)</i>	<i>N (%)</i>	<i>HR (95% CI)</i>	<i>N (%)</i>	<i>HR (95% CI)</i>
Vaccine						
Phase II	5 (50)	1.0	4 (80)	1.0	15 (83)	1.0
Phase III	2 (40)	1.1 (0.2 – 5.9)	6 (75)	1.7 (0.4 – 6.9)	4 (57)	0.5 (0.2 – 1.5)
OSM IgG levels						
≤ median	2 (40)	1.0	7 (100)	1.0	13 (93)	1.0
> median	4 (44)	1.3 (0.2 – 6.9)	3 (50)	0.2 (0.1 – 1.1)	4 (44)	0.3 (0.1 – 0.8)
OVCAR with IL2						
≤ median	1 (14)	1.0	6 (75)	1.0	9 (100)	1.0
> median	4 (80)	6.4 (0.7 – 57.1)	4 (80)	0.5 (0.1 – 2.2)	9 (64)	0.3 (0.1 – 0.9)
OVCAR without IL2						
≤ median	2 (29)	1.0	5 (83)	1.0	10 (91)	1.0
> median	3 (60)	2.9 (0.5 – 18.0)	5 (71)	0.4 (0.1 – 1.6)	8 (67)	0.7 (0.3 – 1.8)

Table IV. Multivariable hazard ratios for OS with 95% CI in vaccinated breast cancer patients only (n=53)

<i>Characteristic</i>	<i>Death: N (%)</i>	<i>HR (95% CI)</i>	<i>p-value</i>
All patients	26 (49)		
Relapse risk			
Low	3 (20)	1.0	
Intermediate	7 (54)	9.0 (2.1 – 38.8)	0.003
High	16 (64)	6.9 (1.8 – 26.5)	0.005
Vaccine			
Phase II	21 (64)	1.0	
Phase III	5 (25)	0.2 (0.1 – 0.6)	0.007
OSM IgG level			
≤ median	17 (65)	1.0	
> median	7 (29)	0.1 (0.02 – 0.3)*	<0.001
Unknown	2 (67)		

* Patients with unknown values are excluded from analysis.

patients with stage II/III ER positive breast cancer. This comparison though is limited by variability of patients treated with two formulations of Theratope, small sample size, different lengths of follow-up and non-random assignment of patients to groups. Only an adequately randomized study will address these issues. Given the current climate with respect to performing ASCT for breast or ovarian cancer, thus study is likely to never be done.

Many studies have been published in the interim about immune reconstitution after ASCT. Early after ASCT in metastatic breast cancer patients, the recovery of lymphocytes has been reported to be associated with superior outcomes. The absolute number of lymphocytes seen at day +15 post ASCT correlates with higher level of freedom from relapse ($p=0.007$) and increased overall survival ($p=0.04$) (20). Compared to autologous bone marrow transplant recipients, peripheral blood ASCT patients have an improved response to recall proteins. The response to recall antigens, such as haemophilus type b and tetanus toxoid, tends to reoccur faster after ASCT than the response to neo-antigens (21). Innate immunity cells such as NK cells are one of the first immune cells to recover early after ASCT and may reach normal levels within one month post infusion of stem cells. Early after ASCT, there are also normal to elevated numbers of CD8+ cells and CD3+ cells seen while CD4+ cells though remaining low for a prolong time after ASCT (22). CD45+ RO+ memory cells exceed CD45+RA+ naïve T cells during the first year post ASCT (23, 24). Thus, T cell recovery early post ASCT represents the expansion of mature T cells with a limited T cell repertoire. Following autologous PBSC stem cell transplantation in absence of B cells depleting drugs such as Rituximab, B cells reconstitute to

normal levels within 2-4 months (23, 24). Finally, dendritic cells recovery has also been shown to occur early after ASCT (25).

When the outcome of breast cancer patients vaccinated with Theratope was compared with regards the ability to mount an immune response, the following observations were made. In multivariable model restricted to intermediate and high risk for relapse the highest IgG antibody titer to the naturally occurring STn bearing molecule OSM and killing of OVCAR after stimulation of mononuclear cells with IL2 both were associated with significant superior EFS rates (HR 0.2, 95% CI 0.1-0.6 and HR=0.3, 95% CI 0.2-0.08, respectively). Survival outcome was associated in multivariable analysis with risk for relapse, formulation of Theratope vaccine used, and IgG antibody titer against OSM.

Since the response to protein recall antigens tend to recover faster than the response to neo-antigen protein post transplant, it is possible that much of the immunological responses seen in our vaccinated ASCT patients was due to reactivation by Theratope of memory immune cells. Previously, there has been published in non-transplant setting similar findings that the patients with the highest antibody titer had the best overall outcome (26). Our data uphold that even after ASCT the IgG antibody titer to OSM can be used as a surrogate marker for patients most likely to benefit from the Theratope cancer vaccine.

ADCC cells include IL2 activated mononuclear cells and are part of innate immunity system. ADCC killing is dependent on the presence of IgG antibody coating the tumor cells targets to allow the ADCC effector cells to recognize the target and bind to it and thus deliver lytic hits. One can speculate that some of the biologically relevance impact of Theratope vaccination on disease control that impact on EFS rates is mediated through ADCC activity as there is an association of OSM antibody IgG response and killing of OVCAR after IL2 stimulation with disease control. Our data though is limited by the lack of measurement of ADCC activity and T cells and non-T cell specific killing of autologous cancer cells after vaccination with Theratope. In the future though there may be a way to increase the immune response following ASCT with the addition of non-specific immune stimulators such as IL2 to the Theratope cancer vaccine. The clinical benefit of this amplification of immune responses remains to be proven by newly designed clinical trials.

Acknowledgements

Supported in part by grants CA67112 and CA66186 from the National Cancer Institutes, National Institutes of Health, DHSS, Bethesda, MD, United States of America and Biomira, Inc, Edmonton, Alberta, Canada. We want to express out gratitude to study nurse K. Lilleby and Laboratory personnel D. Oparin and E. Santos for their assistance, to the Autologous Transplant Team at FHCRC and to the patients and their families.

References

1. Kroger N.; Frick M.; Glisz O. et al. Randomized trial of single compared with tandem high-dose chemotherapy followed by autologous stem-cell transplantation in patients with chemotherapy-sensitive metastatic breast cancer. *J Clin Oncol* **2006**, *24*, 3919-3926.
2. Hanrahan E.; Broglio K.; Frye D. et al. Randomized trial of high-dose chemotherapy and autologous hematopoietic stem cell support for high-risk primary breast carcinoma: follow-up at 12 years. *Cancer* **2006**, *106*, 2327-2336.
3. Bensinger W. I.; Schiffman K. S.; Holmberg L. et al. High dose busulfan, melphalan, thiotepa and peripheral blood stem cell infusion for the treatment of metastatic breast cancer. *Bone Marrow Transplant* **1997**, *19*, 1183-1189.
4. Gutierrez-Delgado F.; Holmberg L. A.; Hooper H. et al. High-dose busulfan, melphalan and thiotepa as consolidation for non-inflammatory high-risk breast cancer. *Bone Marrow Transplant* **2000**, *26*, 51-59.
5. Stiff P. J. ; Bayer R.; Kreger C. et al. High-dose chemotherapy with autologous transplantation for persistent /relapsed ovarian cancer: a multivariate analysis of survival for 100 consecutively treated patients. *J Clin Oncol* **1997**, *15*, 1309-1317.
6. Holmberg L. A. ; Demirer T.; Rowley S. et al. High-dose busulfan, melphalan and thiotepa followed by autologous peripheral blood stem cell (PBSC) rescue in patients with advanced stage III/IV ovarian cancer. *Bone Marrow Transplant* **1998**, *22*, 651-659.
7. Holmberg L. A. ; Kikuchi K. ; Gooley T. et al. Gastrointestinal graft versus host disease in recipients of autologous hematopoietic stem cells: incidence, risk factors and outcome. *Biol of Blood Marrow Transpl* **2006**, *12*, 226-234.
8. Adams K. M.; Holmberg L. A.; Leisenrig W. et al. Risk factors for syngeneic graft-versus-host disease after adult hematopoietic cell transplantation. *Blood* **2004**, *104*, 1894-1897.
9. Thor A.; Ohuchi N.; Szpak C. A. et al. Distribution of oncofetal antigen tumor-associated glycoprotein-72 defined by monoclonal antibody B72.3. *Cancer Res* **1986**, *46*, 3118-3124.
10. Reddish M. A.; Jackson L.; Kogantry R. R. et al. Specificities of anti-sialyl-Tn and anti-Tn monoclonal antibodies generated using novel clustered synthetic glycopeptides epitopes. *Glycoconj J* **1997**, *14*, 549-560.
11. Girling A.; Bartkova J.; Burchell J. et al. A core protein epitope of the polymorphic epithelial mucin detected by the monoclonal antibody SM-3 is selectively exposed in a range of primary carcinomas. *Int J Cancer* **1989**, *43*, 1072-1076.
12. Kinney A. Y.; Sahin A.; Vernon S. W. et al. The prognostic significance of sialyl-Tn antigen in women treated with breast carcinoma treated with adjuvant chemotherapy. *Cancer* **1997**, *80*, 2240-2249.

13. Kobayashi H.; Terao T.; Kawashima Y. et al. Serum sialyl Tn as an independent predictor of poor prognosis in patients with epithelial ovarian cancer. *J Clin Oncol* **1992**, *10*, 95-101.
14. Fung P. Y. S.; Madej M.; Koganty R. R.; Longenecker B. M. Active specific immunotherapy of a murine mammary adenocarcinoma using a synthetic tumor-associated glycoconjugate. *Cancer Res* **1990**, *50*, 4308-4314.
15. Singhal A.; Fohn M.; Hakomari S. I. Induction of alpha N-acetylgalactosamine-O-serine/threonine (Tn) antigen-mediated cellular immune response for active immunotherapy in mice. *Cancer Res* **1991**, *51*, 406-411.
16. MacLean G. D.; Reddish M. A.; Kogarty R. R. et al. Immunization of breast cancer patients using a synthetic sialyl-Tn glycoconjugate plus Detox[™] adjuvant. *Cancer Immunol Immunother* **1993**, *36*, 215-222.
17. Sandmaier B. M.; Oparin D. V.; Holmberg L. A. et al. Evidence of a cellular immune response against sialyl-Tn in breast and ovarian cancer patients after high-dose chemotherapy, stem cell rescue and immunization with Theratope STn-KLH cancer vaccine. *J Immunotherapy* **1999**, *22*, 54-66.
18. Holmberg L. A.; Oparin D. V.; Gooley T.; Sandmaier B. The role of cancer vaccines following autologous stem cell rescue in breast and ovarian cancer patients: experience with the STn-KLH vaccine (Theratope). *Clinical Breast Cancer (Supple)* **2003**, S144-150.
19. Holmberg L. A.; Oparin D. V.; Gooley T. et al. Clinical outcome of breast and ovarian cancer patients treated with high dose chemotherapy, autologous stem cell rescue and Theratope vaccine STn-KLH cancer vaccine. *Bone Marrow Transpl* **2000**, *25*, 1233-1241.
20. Nieto Y.; Shpall E. J.; McNiece I. K. et al. Prognostic analysis of early lymphocyte recovery in patients with advanced breast cancer receiving high-dose chemotherapy with an autologous hematopoietic progenitor cell transplant. *Clinic Cancer Res* **2004**, *10*, 5076-86.
21. Chan C. Y.; Molrine D. C.; Antin J, H. et al. Antibody responses to tetanus toxoid and Haemophilus influenza type b conjugate vaccine following autologous peripheral blood stem cell transplantation (PBSCT). *Bone Marrow Transplant* **1997**, *20*, 33-38.
22. Rosillo M. C.; Ortuno F.; Moraleda J. M. et al. Immune recovery after autologous or rh-G-CSF primed PBSC transplantation. *Euro J Haematol* **1996**, *56*, 301-307.
23. Steingrimsdottir H.; Gruber A.; Bjorkholm M. et al. Immune reconstitution after autologous hematopoietic stem cell transplantation in relation to underlying disease, type of high-dose therapy and infectious complications. *Haematologica* **2000**, *85*, 832-838.

24. Hernandez M. D.; delCanizo M. C.; Gonzalez M. et al. Immune reconstitution after autologous progenitor hemopoietic cell transplantation. A study comparing autologous bone marrow and autologous peripheral blood transplantation. *Medicina Clinica* **1998**, *110*, 768-773.
25. Avigan D.; Wu J.; Joyce R. et al. Immune reconstitution following high-dose chemotherapy with stem cell rescue in patients with advanced breast cancer. *Bone Marrow Transplant* **2000**, *26*, 169-176.
26. MacLean G. D.; Reddish M. A.; Longenecker B. M. Antibodies against mucin-associated sialyl-Tn epitopes correlate with survival of metastatic adenocarcinoma patients undergoing active specific immunotherapy with synthetic STn vaccine. *J Immunotherapy* **1996**, *19*, 9-68.

Chapter 10

Synthesis and Evaluation of Anticancer Vaccine Candidates, C-Glycoside Analogs of sTn and PSA

Michel Weïwer¹, Fei Huang¹, Chi-Chang Chen¹, Xuejun Yuan¹, Kazuo Tokuzaki², Hiroshi Tomiyama², and Robert J. Linhardt^{1,*}

¹Department of Chemistry and Chemical Biology, Rensselaer Polytechnic Institute, Troy, NY 12180

²Kotobuki Pharmaceutical Company, LTD, Oaza Sakaki, Sakaki-machi, Hanishina-gun, Nagano, 389-0697 Japan

Neuraminic acids are biologically important and occupy the terminal position of the glycoconjugate glycans on the outside of cells. Sialyl-Tn (sTn) is found on tumor-associated glycoprotein antigens present on the surface of cancer cells, including those associated with carcinomas of the breast, prostate, pancreas, colon, ovary, lung, and stomach. The sTn antigen is well known as a prognostic indicator and has proven to be an effective target for cancer therapy. Conjugate vaccines of sTn-KLH show remarkable immunogenicity, resulting in the production of both IgM and IgG type antibodies. The sTn C-glycoside was designed and synthesized and its KLH-conjugate was evaluated for immunogenicity in mice. Neuraminic acid C-glycosides such as sTn and polysialic acid (PSA) might be useful in preparing immunogens for active immunization against neuraminic acid containing glycoconjugates in the design and preparation of anti-cancer vaccines.

Introduction

Ulosonic acids, or sialic acids (Figure 1), constitute a unique family of complex monosaccharides that are involved in many important biological functions (1-4). Their biological roles include facilitating cell-cell interactions (5), aggregation (6, 7) and development (8). Ulosonic acids are 8- and 9-carbon monosaccharides that contain an anomeric carboxylate, a deoxygenated methylene C-3 ring carbon, an hydroxylated 2- or 3-carbon side chain at C-6 and are differentially functionalized at C-5. Neuraminic acid (*5-amino-3,5-dideoxy-D-glycero-D-galacto-non-2-ulosonic acid*) does not occur naturally, but many of its derivatives (Neu5Ac, KDN, KDO, Figure 1) are found widely distributed in animal tissues and in bacteria, especially in glycoproteins and gangliosides (2, 3). These sugars can be found in a wide variety of glycosidic linkages, mainly $\alpha(2,3)$ and $\alpha(2,6)$ to galactose or lactose but also as $\alpha(2,8)$ and $\alpha(2,9)$ in polysialic acids (2, 5). The most common sialic acid, *N*-acetylneuraminic acid (Neu5Ac), is a constituent of a large number of glycoconjugates (glycoproteins, glycopeptides, glycolipids, etc.) and is always found at the non-reducing termini of oligosaccharide chains. Neu5Ac represents an important biological receptor domain that interacts with enzymes, hormones, toxins, microbes, viruses and cells (9-11). For example, cell-cell recognition between circulating leukocytes in blood vessels and endothelial cells is believed to occur through the interaction between mammalian lectins (selectins) and Neu5Ac containing oligosaccharides ligands (10). Moreover, many pathogens employ Neu5Ac or other sialic acids to promote infection. Some viruses use hemagglutinin, a sialic acid binding lectin, others utilize neuraminidases, hydrolase-type enzymes that cleaves sialic acid glycosidic linkage, to gain entry into the cells they infect (12-15). *Neisseria meningitidis*, a pathogenic bacteria, biosynthesizes an extracellular capsule composed of Neu5Ac homopolymers as camouflage to escape host immune response (13, 14).

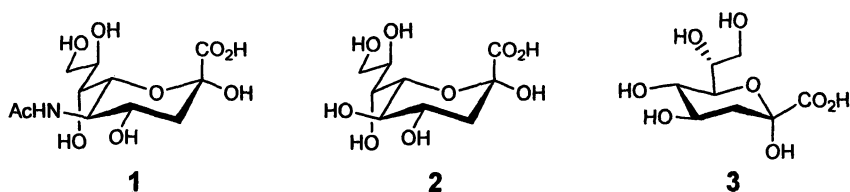


Figure 1. The α -anomeric forms of some sialic acids, Neu5Ac (1), KDN (2) and KDO (3).

The linkage of neuraminic acid to glycoconjugates is among the most labile glycosidic linkages and can be cleaved *in vitro* under mildly acidic conditions

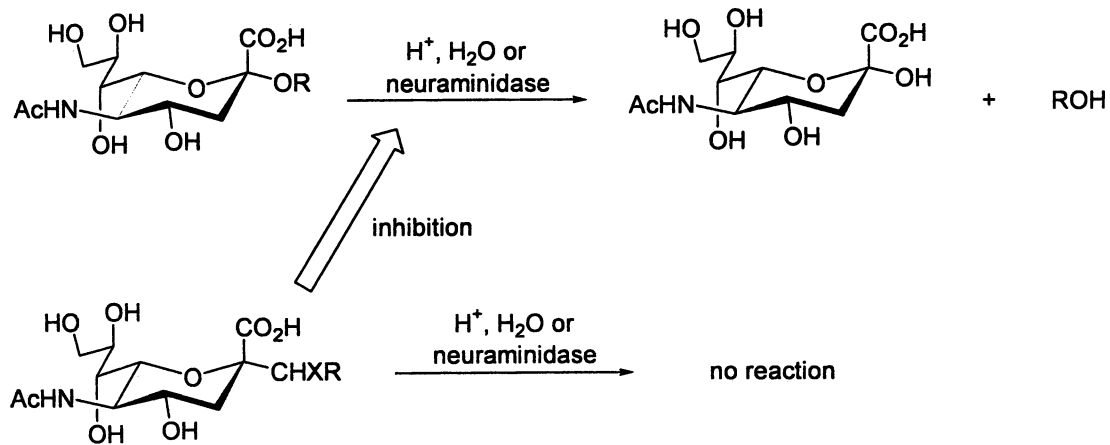


Figure 2. O-glycosides vs C-glycosides stability.

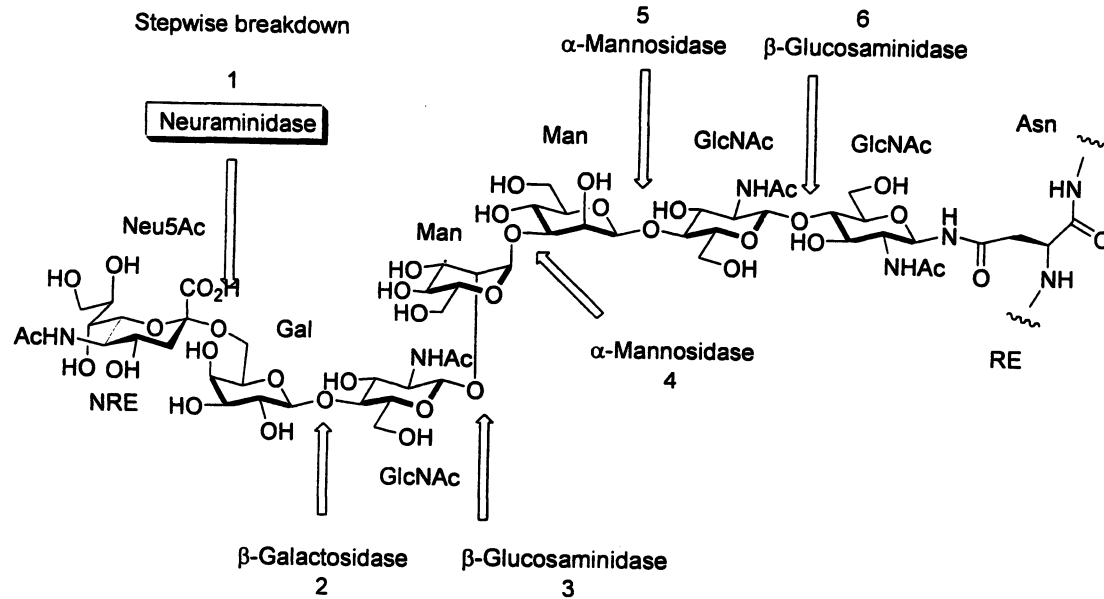


Figure 3. Glycoprotein glycan catabolism

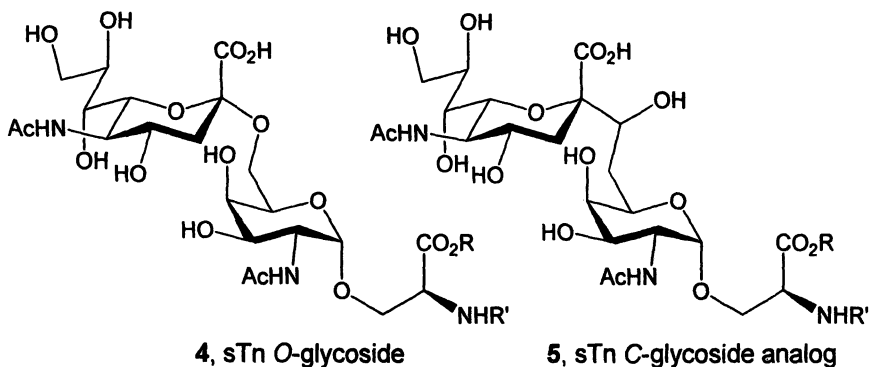
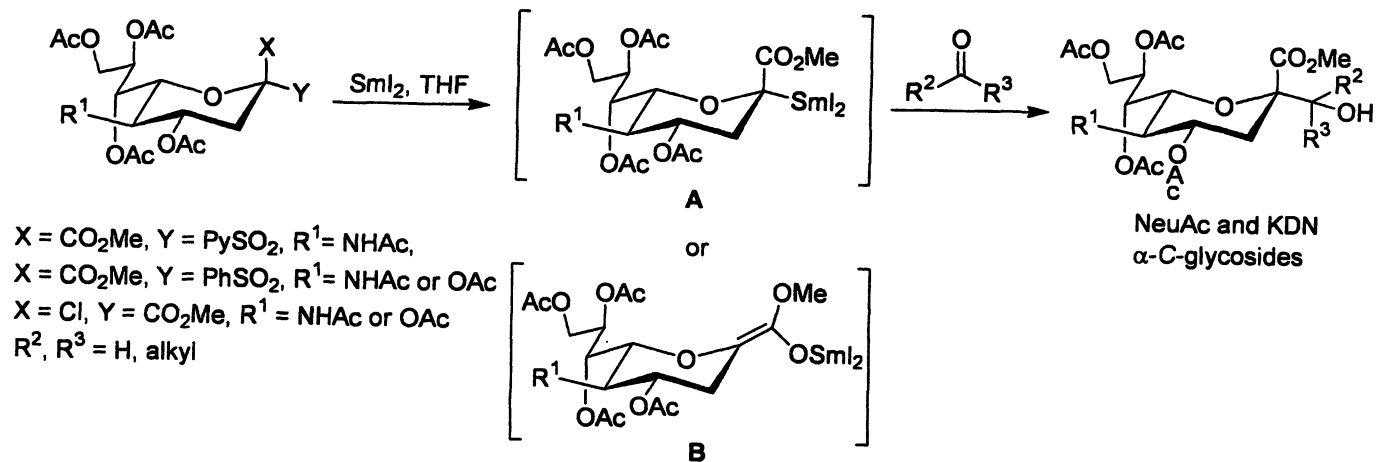


Figure 4. Tumor antigen sTn O-glycoside 4 and its C-glycoside analog 5.

in nature. Thus, it can react with electrophilic type acceptors like ketones or aldehydes to generate the corresponding C-glycoside. Our laboratory has recently used this strategy to complete the synthesis of important C-glycosides targets (37, 43, 44).

Synthesis of the C-Glycoside Analog of sTn

Fully protected sTn C-glycoside analog 6 was prepared by C-glycosylation of the neuraminic acid sulfone donor 7 with the aldehyde acceptor 8 (Scheme 2). The sulfone donor 7 was easily prepared from neuraminic acid in four steps as previously described (42, 45). The critical intermediate, aldehyde acceptor 8, was prepared in 14 steps, from the commercially available diisopropylidene galactose derivative 9. The one carbon extension step was accomplished using a sequential iodination/cyanation early in the synthetic scheme because it is often a problematic step. Iodination of alcohol 9 gave the 6-iodo derivative 10 that was transformed to the corresponding cyano derivative 11 *via* a nucleophilic displacement with KCN in a modest yield (30 %) probably due to unfavorable electronic and steric effects arising from the ring oxygen atom and the axially oriented oxygen at C-4 respectively. Reduction of the 6-cyano derivative 11 using DIBAL-H afforded aldehyde 12. Quantitative reduction of aldehyde 12 with NaBH₄ in MeOH afforded the corresponding alcohol 13. Deisopropylideneation of 13 was then accomplished by treatment with amberlite IR-120 (H⁺) resin in water at 80 °C for 3 h and provided the 6-deoxy-D-galactohexopyranose 14. The one carbon extended galactal 15 was obtained in good yield from 14 using a one pot, three step procedure. First, peracetylation was accomplished using acetic anhydride and catalytic HBr/HOAc. Then, the



Scheme 1. Samarium mediated C-glycosylations.

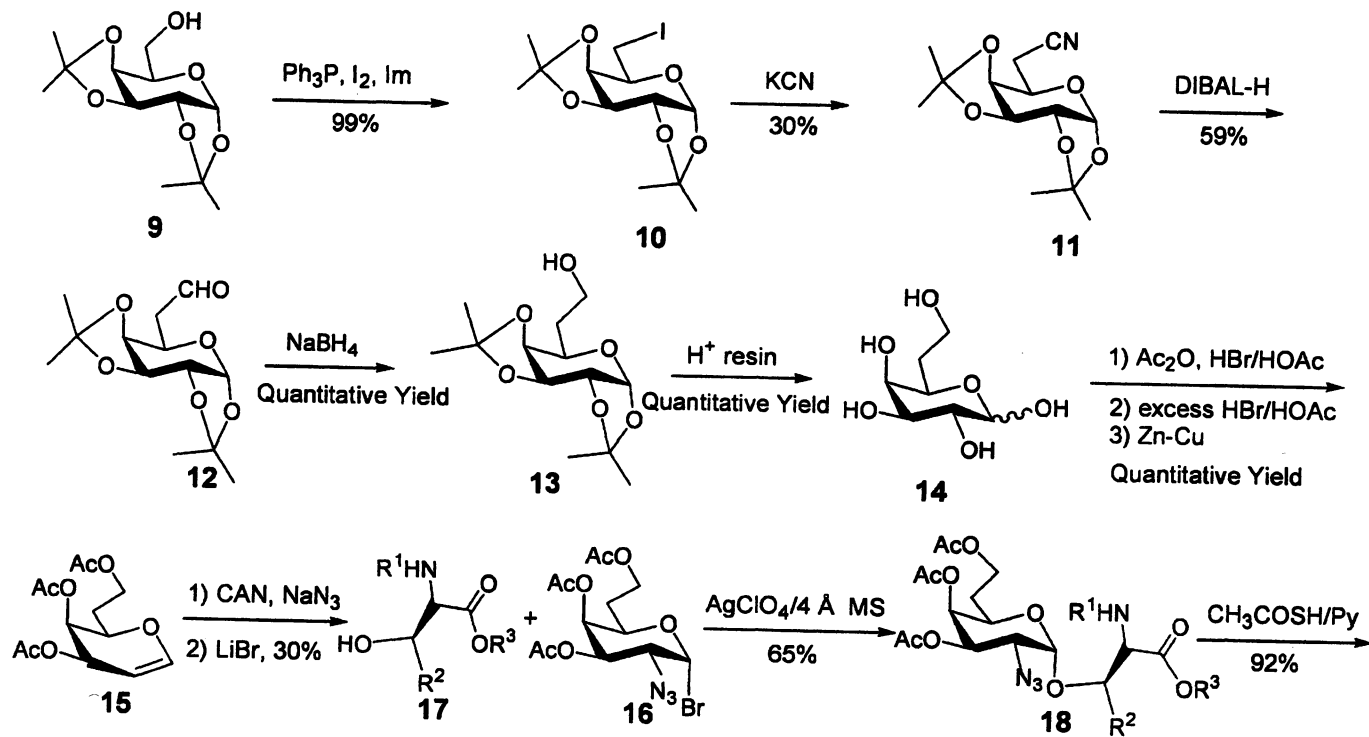
anomeric acetate was converted to the corresponding bromide by treatment with excess HBr/HOAc. Finally, reductive elimination of the 1-bromo and 2-acetoxy groups using Zn/Cu afforded **15** (46). Azidonitration (47) of **15** with ceric ammonium nitrate (CAN) and sodium azide in dry acetonitrile afforded mainly the 2-azido-1-nitrate addition product having the desired galacto configuration as confirmed by ^1H NMR spectroscopy. Treatment of the crude product with lithium bromide in dry acetonitrile under ionic conditions afforded **16** in moderate yield. Glycosylation of **16** with the N^α -benzyloxy carbonyl protected OBN ester of L-Serine **17** (48), in the presence of silver perchlorate afforded the glycopeptide **18** in good yield and stereoselectivity. Conversion of **18** by treatment with thioacetic acid in pyridine (49), afforded the desired glycopeptide product **19** in 92% yield. Selective enzymatic deacetylation of the C-7 primary acetyl group using an esterase from *Rhodospiridium toruloides* (50, 51) at pH 5 afforded **20** in quantitative yield. The site of the enzymatic deacetylation was unambiguously established as C-7 using ^1H NMR spectra. The desired aldehyde acceptor **8** was then obtained *via* Swern oxidation (52).

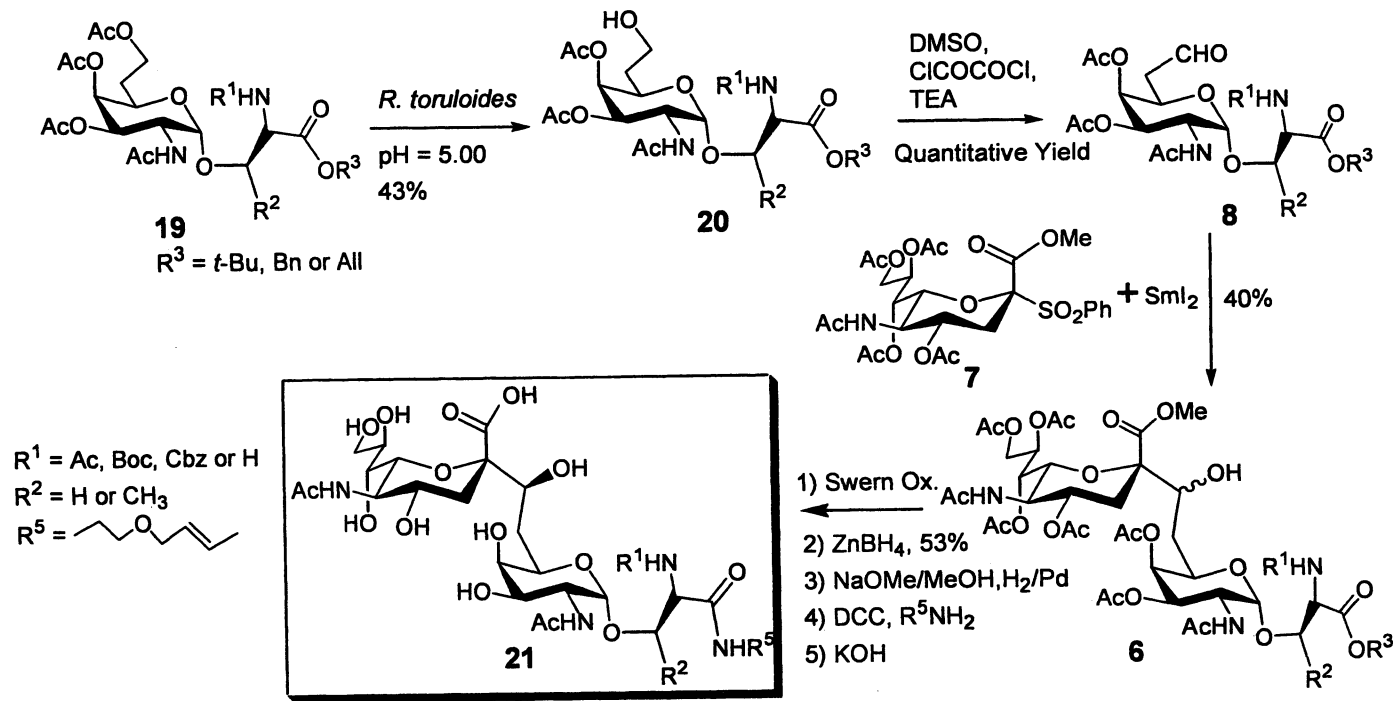
The crucial C-glycosylation step was then accomplished using samarium chemistry developed in our laboratory (53). The neuraminic acid sulfone donor **7** was coupled to aldehyde acceptor **8** in the presence of freshly prepared SmI_2 to afford the fully protected sTn α -C-glycoside **6** as a diastereomeric mixture. Efforts to deoxygenate the bridge hydroxymethylene group by Barton deoxygenation failed. Chemical resolution was achieved by oxidizing the bridge hydroxyl group to ketone which was then stereoselectively reduced in the presence of $\text{Zn}(\text{BH}_4)_2$. After deacetylation, hydrogenolysis, amidation and saponification, **21** was afforded in > 90% *de* (37).

Synthesis of a Double C-Glycoside Analog of sTn

We then embarked in the synthesis of a double C-glycoside analog of sTn **31** (Schemes 3 and 4) (43). This target was designed to reduce the number of synthetic steps required to prepare **21**, to facilitate its conjugation to KLH (Keyhole Limpet Hemocyanin) carrier protein, to further increase biological half-life and to enhance immunological response.

The α -C-glycoside derivative **27** was first synthesized in six steps from *N*-acetylglucosamine as described by Cipolla *et al* (Scheme 3) (54). The 6-hydroxyl group of compound **27** was then protected as the 6-TBDPS silyl ether followed by 3,4-isopropylideneation to afford compound **28**. Based on our previous experience, we decided to double protect the acetamido group at C-2 position in order to avoid cyclic hemiaminal formation later in the scheme when the C-6 position will be oxidized to the corresponding aldehyde. Thus, **28** was treated



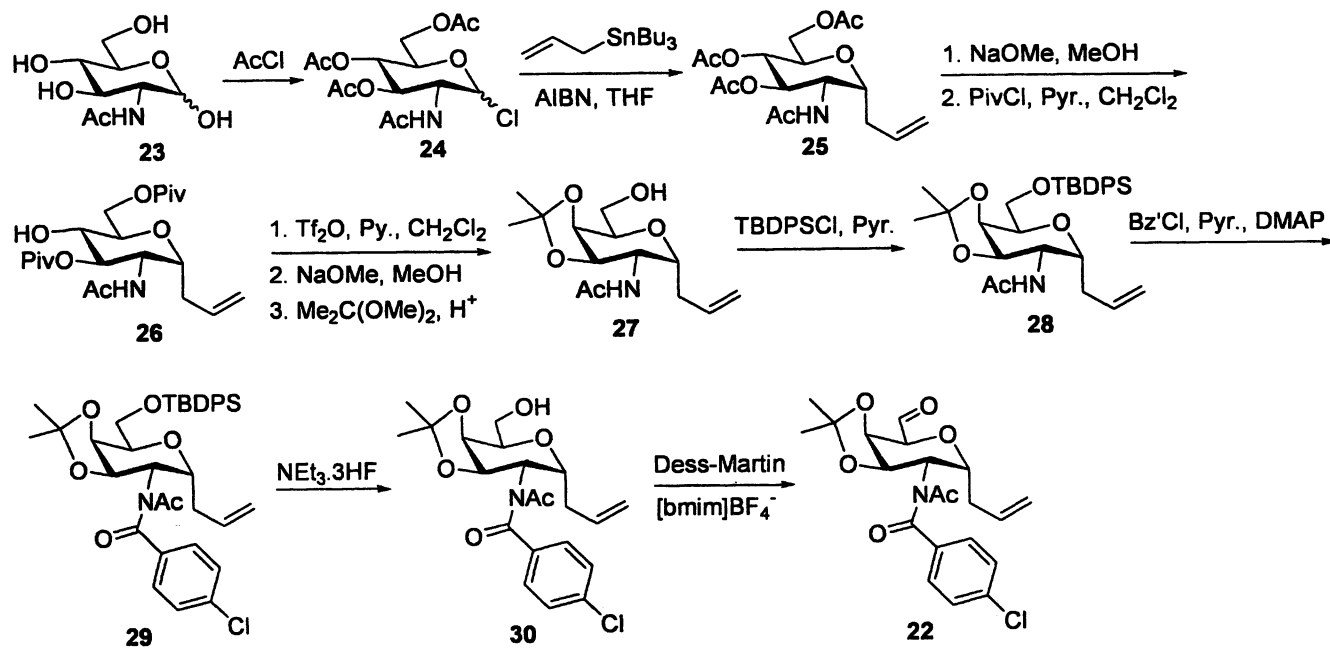


Scheme 2. Synthesis of the C-glycoside analog of *sTn*.

with *p*-chlorobenzoyl chloride in a mixture of pyridine and dichloromethane to give **29** in excellent yield. Chemoselective removal of the TBDPS protection at C-6 position using trihydrofluoride triethylamine afforded the corresponding alcohol **30** in quantitative yield (TBAF/AcOH system proved to be inefficient for this transformation). We then decided to explore very mild conditions for the oxidation of C-6 hydroxyl group to the corresponding aldehyde in order to prevent C-5 epimerization. Hypervalent iodine reagents in ionic liquid have been previously reported as mild chemoselective and in some cases regioselective oxidative process (55). Thus, compound **30** was dissolved in 1-butyl-3-methylimidazolium tetrafluoroborate ($[\text{bmim}]^+\text{BF}_4^-$) and treated with 1.5 equivalent of Dess-Martin periodinane for 3 hours. Extraction of the product from the ionic liquid was performed using diethyl ether and resulted in the isolation of a single product which was identified with HRMS and ^1H NMR as the desired aldehyde **22**. Because of its propensity for epimerization, the aldehyde was directly used in the C-glycosylation reaction.

Neu5Ac sulfone donor **7** was reacted with aldehyde **22** in the presence of 6 equivalents of SmI_2 in THF (Scheme 4) to afford the desired double C-glycoside product **32** in a modest 25% yield but with complete diastereocontrol in favor of the *S* isomer, as demonstrated by molecular modeling and NOESY experiments. Complete deprotection was next accomplished in three steps. Deisopropylidenation using acetic acid at 80°C for 1h, followed by deacylation with NaOMe in MeOH and saponification of the methyl ester using 0.1M KOH afforded **33** in quantitative yield. After oxidative cleavage of the terminal olefin of **33**, reductive amination was accomplished by treatment of the corresponding aldehyde with NaCNBH_3 in the presence of the KLH hapten to give target compound **31**.

O-Linked conjugate vaccines of sTn-KLH have shown good immunogenicity, resulting in the production of both IgM and IgG type antibodies (56, 57). The KLH-conjugates of sTn C-glycosides were evaluated for immunogenicity in mice (Figure 5, Table 1). Female BALB/c mice, 5 weeks old, were immunized with *O*-linked sTn-KLH, *C*-linked sTn-KLH or double-*C*-linked sTn-KLH conjugate (1 or 10 μg) 3 times at two weeks interval. Seven days after the last immunization, mice were anesthetized with ether and blood samples were collected from abdominal vein. Antisera were separated by centrifugation. IgG and IgM antibody titers of the antisera against sTn antigen were assayed by ELISA. The titer was defined as highest dilution yielding an absorbance of greater over that of normal sera. Both KLH-conjugate of sTn C-glycosides lead to the production of IgM and IgG type antibodies. The C-glycosides were found to be more immunologically active than the natural *O*-linked sTn antigen (Table 1). Furthermore, the C-glycosides of sTn enhanced the production of the tightly binding IgG class of antibodies.



Scheme 3. Synthesis of the C-glycoside acceptor.

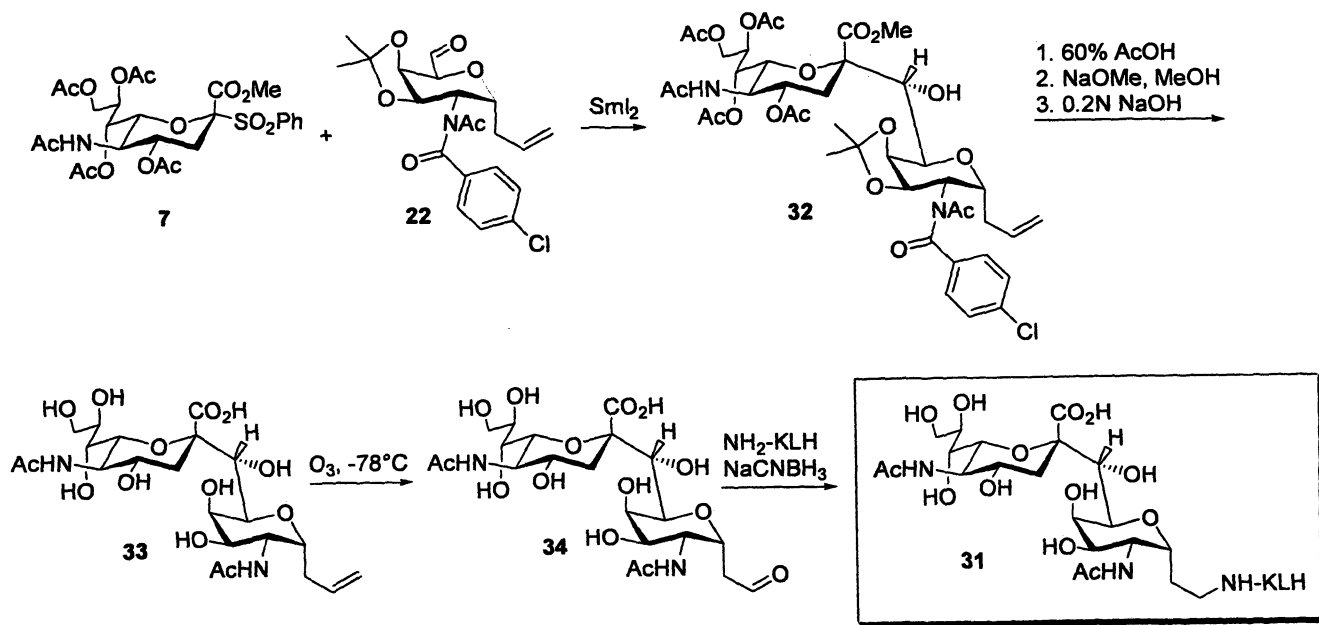
Synthesis of PSA C-disaccharide

Tumor related antigen often contains *O*-linked Neu5Ac- α (2,8)-Neu5Ac disaccharide units (58). The *O*-linkage being catabolically unstable, we have been interested in the synthesis of a *C*-linked Neu5Ac disaccharide (44). Indeed, vaccination with a catabolically stable sialic acid *C*-glycoside analog might enhance immunogenicity. Ultimately, the synthesis of *C*-linked polysialic acids (PSAs) would be of particular interest.

Polysialic acids are naturally occurring helical, linear homopolymers composed entirely of negatively charged sialic acid residues joined by α (2 \rightarrow 8), α (2 \rightarrow 9), or α (2 \rightarrow 8)/ α (2 \rightarrow 9) alternating ketosidic linkages and are commonly found N-linked to a neural cell adhesion molecule (NCAM) (59). The precise function of PSA has not yet been established but a well-demonstrated property of PSAs is in cell-cell interaction and adhesion. It is postulated that alteration of PSA glycans in NCAM reduces cell adhesion and may be involved in invasive metastasis (60, 61). The unusual lability of PSA, their participation in developmental biology and their reappearance in various tumors make their *C*-glycosides analogs ideal targets for a wide array of experimental, biological and possibly therapeutic applications.

Retrosynthetic analysis (Scheme 5) suggested that samarium mediated *C*-glycosylation of aldehyde acceptor 36 with Neu5Ac phenyl sulfone donor 7 would afford the nor-*C*-disaccharide target 35 (44). The success of this synthetic route relies on the design of aldehyde acceptor 36, which we envision could be synthesized using the previously reported sialic acid thiophenyl glycoside 39 (62).

We first considered the use of *t*-butyldiphenylsilyl (TBDPS) protection in order to differentiate the C-9 primary hydroxyl group. Unfortunately, after successful isopropylidene protection of 7,8-hydroxyl groups we encountered decomposition problems during the deprotection step of the 9-OTBDPS group using TBAF. We then considered another strategy involving a selective ring opening of 8,9-*p*-methoxybenzylidene ring (63). First, regioselective protection of 8,9-diol gave 40 in 94% yield, which was then regioselectively opened using aluminium (III) chloride and $\text{BH}_3\cdot\text{NMe}_3$ to afford the 9-*p*-methoxybenzyl (PMB) ether 41 in 60% yield (Scheme 6). Once the differentiation of the C-9 position was accomplished, we focused our attention on the protecting groups at 4,7,8 positions. Modeling suggested that allyl protection should allow efficient *C*-glycosylation with the bulky nucleophile derivative of 7. The ease of introduction, electron-donating properties, orthogonality, and ease of removal made OAll an ideal protecting group for synthesis of desired aldehyde acceptor 36 (64). Perallylation of 41 under standard conditions resulted in unexpected



Scheme 4. Synthesis of the double C-glycoside analog of *sTn*.

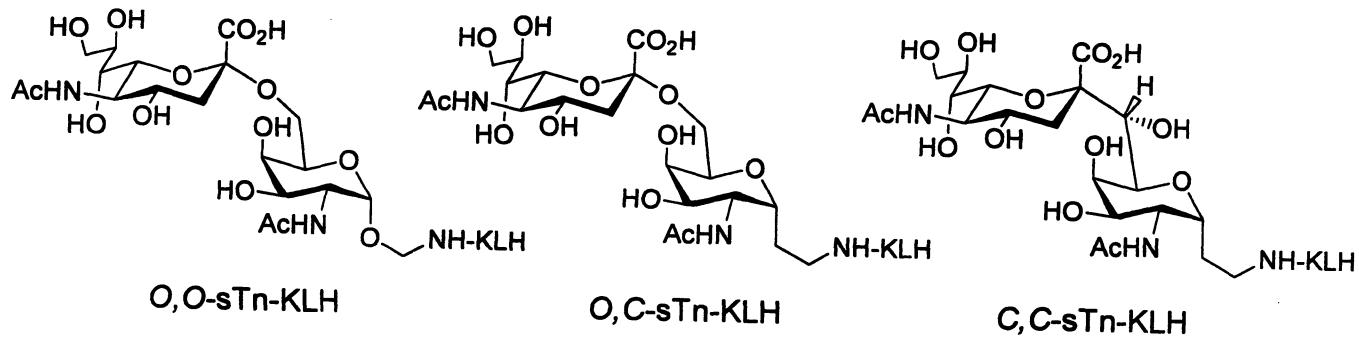
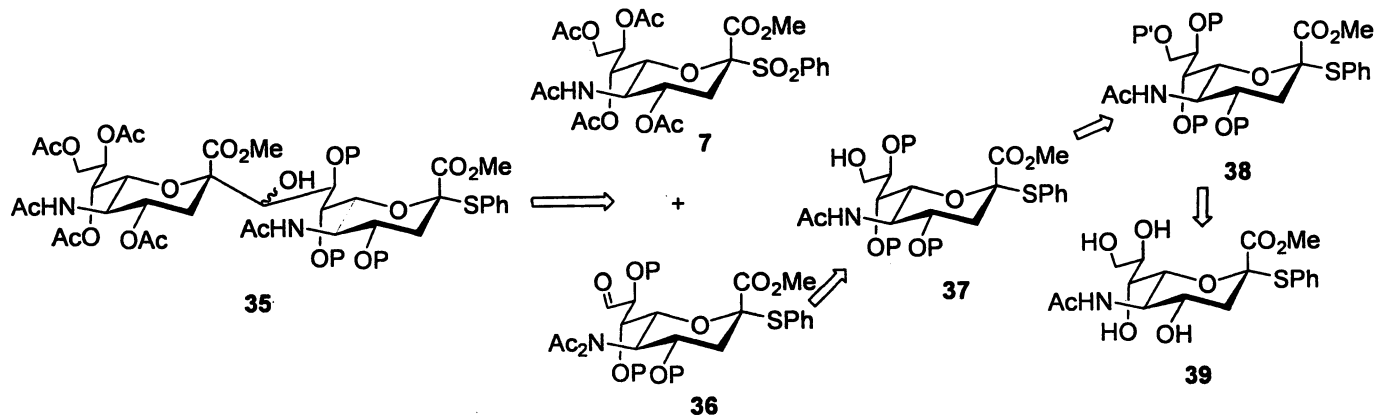
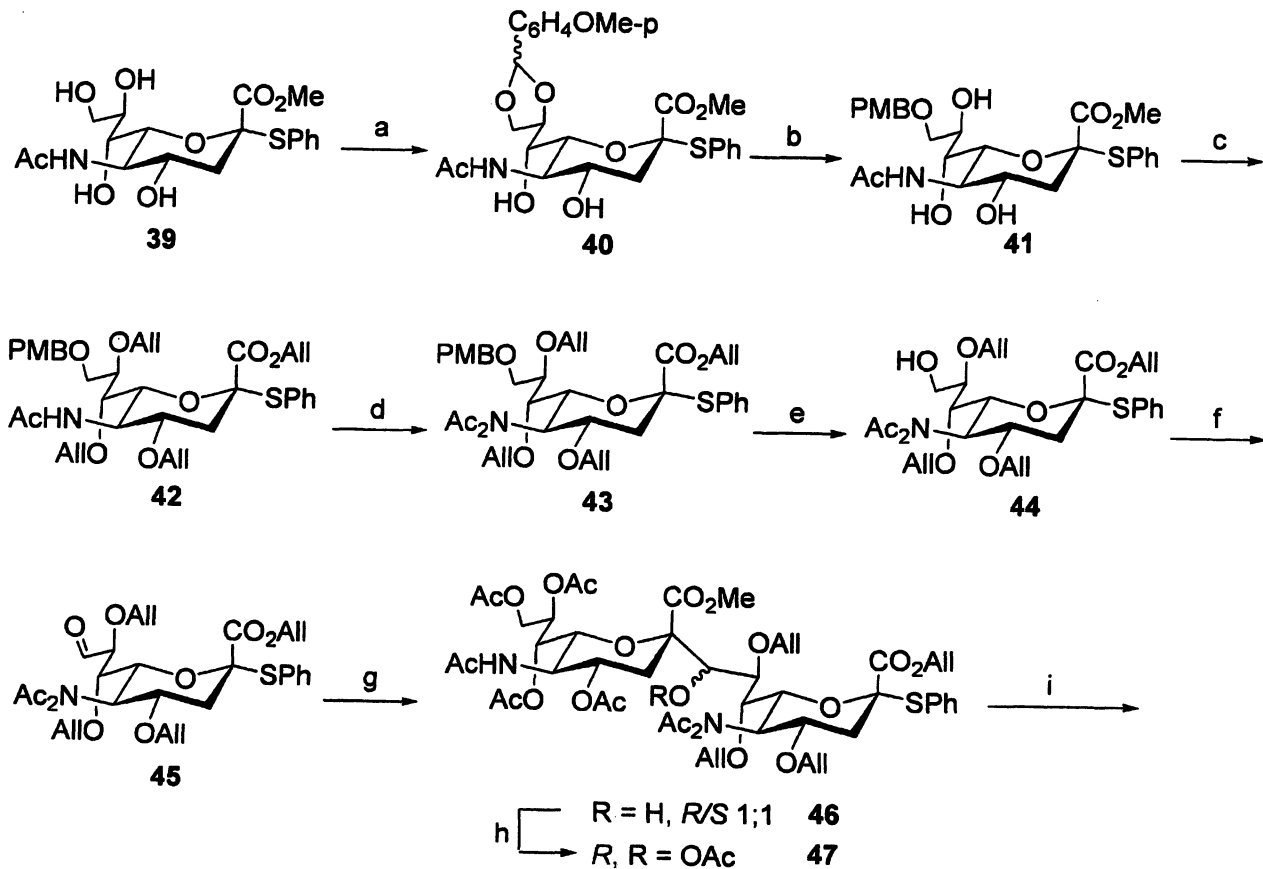
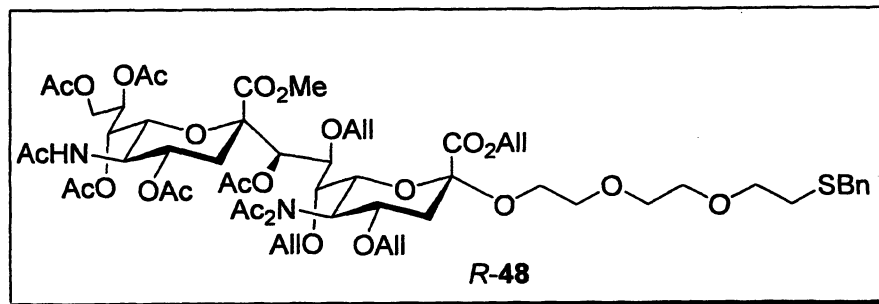


Figure 5. KLH conjugates of sTn O- and C-glycosides.



Scheme 5. Retrosynthetic scheme for the synthesis of PSA C-disaccharide 35.





Scheme 6. Synthesis of C-Neu5Ac disaccharide. a) $p\text{-MeOPhCH(OMe)}_2$, CSA, MeCN, 94%; b) $\text{BH}_3 \cdot \text{NMe}_3$, AlCl_3 , MS 4A 60%; c) AllBr, BaO/Ba(OH) $_2$ ·8H $_2$ O, DMF, 75%; d) $\text{CH}_2(\text{CH}_3)\text{COOCCH}_3$, TsOH, 65 °C, 80%; e) CAN, 74%; f) Dess-Martin reagent, 70%; g) 7, 5 equiv 0.1N SmI $_2$, THF, 35%; h) Ac $_2$ O, pyr.; i) 2-[2-(2-benzylthioethoxy)ethoxy]ethanol. TfOMe. MS 4A.

Table I. Immunological Evaluation of *O*- and *C*-Glycosides of Tn and sTn

	IgG	IgM
<i>O,O</i> -sTn-KLH	1860 ± 1278	480 ± 981
<i>O,C</i> -sTn-KLH	11520 ± 2862	2560 ± 2147
<i>C,C</i> -sTn-KLH	7360 ± 4005	2330 ± 1198

Female BALB/c mice were immunized three times with 1 µg/mouse (*O,C*) or 10 µg/mouse (*C,C* and *O,O*) of KLH conjugates. Seven days after the third immunization, IgG and IgM titers were measured by ELISA. Data are expressed as the mean ± SD of 5 mice.

lactamization. Alternatively, a mixture of barium oxide and barium hydroxide gave the desired, albeit transesterified, allylated compound **42** in 75% yield (65). Then, the *N*-acetyl group was double protected in order to avoid the formation of an hemiaminal ring later in the scheme (during oxidation of the 9-hydroxyl to the corresponding aldehyde). *N*-acetylation of **42** afforded **43** in 80% yield (66). Deprotection of the PMB was accomplished using cerium ammonium nitrate and gave the 9-hydroxyl derivative **44**, which was further oxidized to the corresponding aldehyde acceptor **45** in 70% yield (67). *C*-Glycosylation of the aldehyde **45** with Neu5Ac sulfone donor **7** in the presence of freshly prepared SmI₂ (68) gave the desired nor- α (2,8)-*C*-neuraminic acid disaccharide **46** in 35% yield, as a diastereomeric mixture (*R/S* 1:1) at the bridge hydroxymethylene group. The two diastereomers were separated by flash chromatography and the *R* isomer was acetylated at the bridge hydroxyl group to give **47**. Hapten conjugation was then accomplished by *O*-glycosylation of 2-[2-(2-benzylthioethoxy)ethoxy]ethanol using methyl triflate as promoter.

Conclusion and Perspectives

Samarium mediated *C*-glycosylation has proven to be an effective method for the synthesis of α -*C*-glycosides of Neu5Ac. We have successfully used this strategy to synthesize α -*C*-glycosides analogs of sTn antigen and polysialic acids. These catabolically stable, carbohydrate based, potential anticancer vaccines shown remarkable immunogenicity, resulting in the production of both IgM and IgG type antibodies. We have shown that the synthesis of polysialic acid type *C*-disaccharide can be accomplished via the same procedure. Although

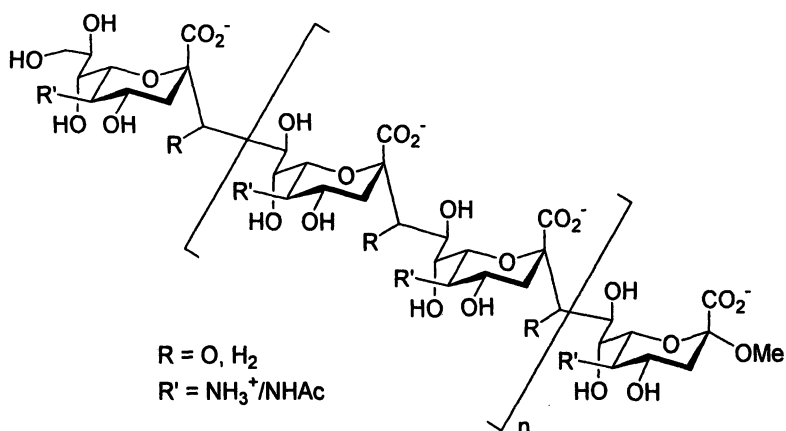


Figure 6. C-linked polysialic acid target.

very promising, the yield of the C-glycosylation need to be improved to apply this reaction to the polymerization reactions that would allow us to prepare C-linked polysialic acid derivatives (Figure 6).

References

1. Schauer, R. *Adv. Carbohydr. Chem. Biochem.* **1982**, *40*, 131-234.
2. Schauer, R., *Sialic Acids, Chemistry, Metabolism and Function*. Springer Verlag; Ed.; Wien: New York, NY, **1982**.
3. Jeanloz, R.; Codington, J. The biological role of sialic acid at the surface of the cell. In: *Biological Roles of Sialic Acids*; Rosenberg, A.; Schengrund, C.; Eds.; Plenum Press: New York, NY, **1976**, 201-238.
4. Varki, A. *Glycobiology* **1992**, *2*, 25-40.
5. Rosenberg, A.; Schengrund, C. *Biological roles of sialic acids*, Plenum Press, New-York and London, **1976**.
6. Weiss, L. *Nature*, **1961**, *191*, 1108-1109.
7. Weiss, L. *Exp. Cell Res.*, **1963**, *30*, 509-520.
8. Finne, J. *J. Biol. Chem.*, **1982**, *257*, 11966-11970.
9. Rosen, S.R. *Immunology* **1993**, *5*, 237-247.
10. McEver, R.P. *Current Opinion in Immunology* **1994**, *6*, 75-84.
11. Nelson, R.M.; Venot, A.; Bevilacqua, M.P.; Linhardt, R.J.; Stamenkovic, I. *Ann. Rev. Cell Dev. Biol.* **1995**, *11*, 601-631.
12. Hirst, G.K. *J. Exp. Med.*, **1943**, *78*, 99-109.

13. Wessels, M.R.; Poszgay, V.; Kasper, D.L.; Jennings, H.J. *J. Biol. Chem.*, **1987**, *262*, 8262-8267.
14. Bhattacharjee, A.K.; Jennings, H.J.; Kenny, C.P. *Biochemistry*, **1978**, *17*, 645-651.
15. Bomsel, M.; Alfsen, A. *Nat. Rev. Mol. Cell Bio.*, **2003**, *4*, 57-68.
16. Trimble, R.B.; Tarentino, A.L. *J. Biol. Chem.* **1991**, *266*, 1646-1651.
17. Arentino, A.L.; Gomez, C.M.; Plummer, T.H. Jr. *Biochem.* **1985**, *24*, 4665-4671.
18. Elder, J.H.; Alexander, S. *Proc. Nat. Acad. Sci. (USA)*. **1982**, *79*, 4540-45404.
19. Ghosh, P.; Bachhawat, B.K.; Surolia, A. *Arch. Biochem. Biophys.* **1981**, *206*, 454-457.
20. Baenziger, J.U.; Maynard, Y. *J. Biol. Chem.* **1980**, *255*, 4607-4613.
21. Hasegawa, A.; Terada, T.; Ogawa, H.; Kiso, M. *J. Carbohydr. Chem.* **1992**, *11*, 319-331.
22. Ress, D.K.; Linhardt, R.J. *Curr. Org. Synthesis*, **2004**, *1*, 31-46.
23. Suzuki Y.; Sato K.; Kiso M.; Hasegawa A. *Glycoconjugate journal*, **1990**, *7*, 349-354.
24. Witczak, Z. *J. Curr. Med. Chem.*, **1999**, *6*, 165-178.
25. Sakamoto J.-I.; Koyama T.; Miyamoto D.; Yingsakmongkon S.; Hidari Kazuya I. P. J.; Jampangern W.; Suzuki T.; Suzuki Y.; Esumi Y.; Hatano K.; Terunuma D.; Matsuoka K. *Bioorg. Med. Chem. Lett.*, **2007**, *17*, 717-721.
26. Schauer, R. *Trends Glycosci. Glyc.* **1997**, *9*, 315-330.
27. Traving, C.; Schauer, R. *Cell. Mol. Life Sci.* **1998**, *54*, 1330-1349.
28. Air, G. M.; Laver, W. *Proteins: Struct. Funct. Genet.* **1989**, *6*, 341-356.
29. Carlos, M.P.; Anderson, D.E.; Gardner, M.B.; Torres, J.V. *AIDS Res. Hum. Retroviruses* **2000**, *16*, 153-161.
30. Hakomori, S.I. *Cancer Res.* **1989**, *49*, 257-331.
31. Schwarz, J.B.; Kuduk, S.D.; Chen, X.; Sames, D.; Glunz, P.W.; Danishefsky, S.J. *J. Am. Chem. Soc.* **1999**, *121*, 2662-2673.
32. Liebe, B.; Kunz, H. *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 618-621.
33. Kuberan, B.; Linhardt, R.J. *Curr. Org. Chem.* **2000**, *4*, 653-677.
34. Ragupathi, G.; Koganty, R.R.; Qiu, D.; Lloyd, K.O.; Livingston, P.O. *Glycoconjugate J.* **1998**, *15*, 217-221.
35. Ravishankar, R.; Surolia, A.; Vijayan, M.; Lim, S.; Kishi, Y. *J. Am. Chem. Soc.* **1998**, *120*, 11297-11303.
36. Poveda, A.; Asensio, J.L.; Polat, T.; Bazin, H.; Linhardt, R.J.; Jiménez-Barbero J. *Eur. J. Org. Chem.* **2000**, *9*, 1805-1813.
37. Kuberan, B.; Sikkandar, S.A.; Tomiyama H.; Linhardt, R.J. *Angew. Chem. Int. Ed. Engl.* **2003**, *42*, 2073-2075.
38. Koenigs, W.; Knorr, E. *Chem. Ber.* **1901**, *34*, 957-981.

39. Helferich, B.; Zirner, J. *Chem. Ber.* **1962**, *95*, 2604-2611.
40. Hasegawa, A.; Okhi, H.; Nagahama, T.; Ishida, H.; Kiso, M. *Carbohydr. Res.* **1991**, *212*, 277-281.
41. Hasegawa, A.; Nagahama, T.; Okhi, H.; Hotta, K.; Ishida, H.; Kiso, M. *J. Carbohydr. Chem.* **1991**, *10*, 493-498.
42. Cao, S.; Meunier, S.; Andersson, F.; Letellier, M.; Roy, R. *Tetrahedron Asymm.* **1994**, *5*, 2303-2312.
43. Ress, D.K.; Baytas, S.N.; Wang, Q.; Munoz, E.M.; Tokuzoki, K.; Tomiyama, H.; Linhardt, R.J. *J. Org. Chem.* **2005**, *70*, 8197-8200.
44. Yuan, X.; Ress, D.K.; Linhardt, R.J. *J. Org. Chem.* **2007**, *72*, 3085-3088.
45. Marra, A.; Sinaÿ, P. *Carbohydr. Res.* **1989**, *187*, 35-42.
46. Shull, B.K.; Wu, Z.; Koreeda, M. *J. Carbohydr. Chem.* **1996**, *15*, 955-964.
47. Lemieux, R.U.; Ratcliffe, R.M. *Can. J. Chem.* **1979**, *57*, 1244-1251.
48. Schultz, M.; Kunz, H. *Tetrahedron-Asymm.* **1993**, *4*, 1205-1220.
49. Rosen, T.; Lico, I. M.; Chu, D. T. W. *J. Org. Chem.* **1988**, *53*, 1580-1582.
50. Horrobin, T.; Tran, C.H.; Crout, D. *J. Chem. Soc. Perkin Trans. 1* **1998**, *1*, 1069-1080.
51. Kuberan, B.; Wang, Q.; Koketsu, M.; Linhardt, R.J. *Synth. Commun.* **2002**, *32*, 1421-1426.
52. Tidwell, T.T. *Synthesis* **1990**, 857-870.
53. Du, Y.; Linhardt, R. J. *Carbohydr. Res.* **1998**, *308*, 161-164.
54. Cipolla, L.; Ferla, B.L.; Lay, L.; Peri, F.; Nicotra, F. *Tetrahedron: Asymmetry* **2000**, *11*, 295-303.
55. Yadav, J.S.; Reddy, B.V.S.; Basak, A.K.; Narsaiah, A.V. *Tetrahedron* **2004**, *60*, 2131-2135.
56. Sorensen, A.L.; Reis, C.A.; Tarp, M.A.; Mandel, U.; Ramachandran, K.; Sankaranarayanan, V.; Schwientek, T.; Graham, R.; Taylor-Papadimitriou, J.; Hollingsworth, M.A.; Burchell, J.; Clausen, H. *Glycobiology* **2006**, *16*, 96-107.
57. Ragupathi, G.; Koide, F.; Sathyan, N.; Kagan, E.; Spassova, M.; Bornmann, W.; Gregor, P.; Reis, C.A.; Clausen, H.; Danishefsky, S.J.; Livingston, P.O. *Cancer Immunology Immunotherapy* **2003**, *52*, 608-616.
58. Angata, T.; Varki, A. *Chem. Rev.* **2002**, *102*, 439-469.
59. Reglero, A.; Rodrigues-Aparicio, L. B.; Luengo, J. M. *Int. J. Biochem.* **1993**, *25*, 1517-1527.
60. Roth, J.; Zuber, C.; Wagner, P.; Taatjes, D. J.; Weisgerber, C.; Heitz, P. U.; Goridis, C.; Bitter-Suermann, D. *Proc. Natl. Acad. Sci. U.S.A.* **1988**, *85*, 2999-3003.
61. Scheidegger, E. P.; Lackie, P. M.; Papay, J.; Roth, J. *Lab. Invest.* **1994**, *70*, 95-106.
62. Du, Y.; Linhardt, R. J.; Vlahov, I. R. *Tetrahedron* **1998**, *54*, 9913-9959.
63. Demchenko, A. V.; Boons, G. J. *Chem. Eur. J.* **1999**, *5*, 1278-1283.

64. Kimbonguila, A. M.; Boucida, S.; Guibé, F.; Loffet, A. *Tetrahedron* **1997**, *53*, 12525-12538.
65. Guo, Z.-C.; Han, L.; Hu, B.; Cao, C.-S. *Youji Huaxue (Chinese)* **2004**, *24*, 946-949.
66. Gizur, T.; Harsanyi, K. *Syn. Commun.* **1990**, *20*, 2365-2371.
67. Dess, D. B.; Martin, J. C. *J. Org. Chem.* **1983**, *48*, 4155-4156.
68. Vlahov, I. R.; Vlahova, P. I.; Linhardt, R. J. *J. Am. Chem. Soc.* **1997**, *119*, 1480-1481.

Chapter 11

An Automated Method for Determining Glycosylation and Site Diversity in Glycoproteins

Hyun Joo An¹, John Tillinghast², and Carlito B. Lebrilla^{1,*}

¹Department of Chemistry and ²Division of Biostatistics, University of California at Davis, One Shields Avenue, Davis, CA 95616

The determination of glycosylation sites and oligosaccharide heterogeneity is key toward understanding the biological roles of glycoproteins. There are no previous methods that can reliably determine the site of glycosylation and the glycan heterogeneity at specific sites. We have developed a procedure for the determination of glycosylation sites and oligosaccharide heterogeneity in glycoproteins. The method is based on a combination of nonspecific proteolysis, deglycosylation, and high mass accuracy mass spectrometry analysis. The glycoprotein was proteolytically degraded using a nonspecific protease into glycopeptide fragments. The exact peptide mass was calculated by subtracting the glycan mass from the observed glycopeptide mass. The amino acid sequence of a matched peptide mass was identified from the protein database. A computer program, GlycoX, was developed in MATLAB to aid in the determination of the glycosylation sites and oligosaccharide heterogeneity in glycoproteins.

Introduction

Glycosylation is one of the most common forms of post-translational modification in eukaryotic proteins and is involved in many cell communication and signaling event (1). It has become apparent that a fundamental understanding of the structure and function of many glycoproteins is in no small part dependent upon knowledge of the location and composition of oligosaccharides linked to the protein. There are two major types of glycosylation: N-linked and O-linked. In N-glycosylation, an N-acetylglucosamine (GlcNAc) residue is attached to the amide group of an asparagine residue. Among eukaryotes, all N-linked glycans share a common trimannosyl core structure consisting of mannose and GlcNAc residues ($\text{Man}_3\text{GlcNAc}_2$). N-linked glycosylation is known to occur at asparagine residues in the consensus protein sequence NX(S/T), where X may be any amino acid residue, proline excepted. The presence of the consensus sequence is required for N-linked glycosylation; however, N-glycosylation of each potential site is not obligatory. Hence, a glycoprotein may contain a number of potentially N-glycosylated sites, each of which may or may not be glycosylated. Additionally, O-glycosylation may occur at any serine or threonine residue with no single common core or consensus protein sequence. Finally, the population of glycans occurring at a given glycosylated site is often not homogeneous; that is, a particular site of N-linked or O-linked glycosylation may be occupied by structurally distinct glycans on different copies, or glycoforms, of the protein. This is referred to as microheterogeneity, or site heterogeneity.

Traditionally, site-specific glycosylation analysis has been extremely challenging due to the high complexity of glycoproteins. There have been several recent reports on glycosylation site analysis (2-11). Typical approaches to this task are based on some combination of specific enzymatic proteolysis (usually with trypsin), fractionation of glycopeptides (most often by liquid chromatography or affinity chromatography), and glycopeptide analysis by mass spectrometry (6-9). In some cases, deglycosylation of glycopeptides is concurrently performed with the information regarding the glycan discarded (6). Unfortunately, many glycoproteins are resistant to trypsin (12, 13). Furthermore, the glycopeptides that are formed by specific enzymatic proteolysis are often too large for MS analysis because glycosylation may produce missed tryptic cleavages (8). For these reasons, the information regarding glycosylation is very often incomplete.

An alternative method for the determination of N-glycosylation sites and oligosaccharide heterogeneity in glycoproteins was developed in the our laboratory (3). Experimental scheme is shown in Figure 1. In this approach, a glycoprotein is extensively digested by nonspecific proteolysis, using a highly active mixture of proteases known as pronase to produce glycopeptides with small peptide moiety (2-8 residues). Nonglycosylated regions of the glycoprotein are digested to amino acids and dipeptides, while glycosylated

regions are protected from protease activities by the oligosaccharides. The resulting mixture of glycopeptides is easily separated from the amino acids and salts using solid-phase extraction (SPE) with porous graphitic carbon (PGC). In parallel, the glycoprotein is treated with PNGase F, an enzyme that specifically releases *N*-linked glycans from glycoproteins. However, this procedure is necessary only for large glycoproteins with several glycosylation sites. Smaller glycoproteins do not typically require the independent determination of the oligosaccharide constituents. In any case, the released glycans are also purified by SPE with PGC. The purified glycopeptides and glycans are analyzed by matrix-assisted laser desorption/ionization (MALDI) and high-accuracy MS analysis, in this case Fourier transform ion cyclotron resonance-mass spectrometry (FTICR-MS), allowing the assignment of site-specific glycosylation and elucidation of glycan heterogeneity at individual sites.

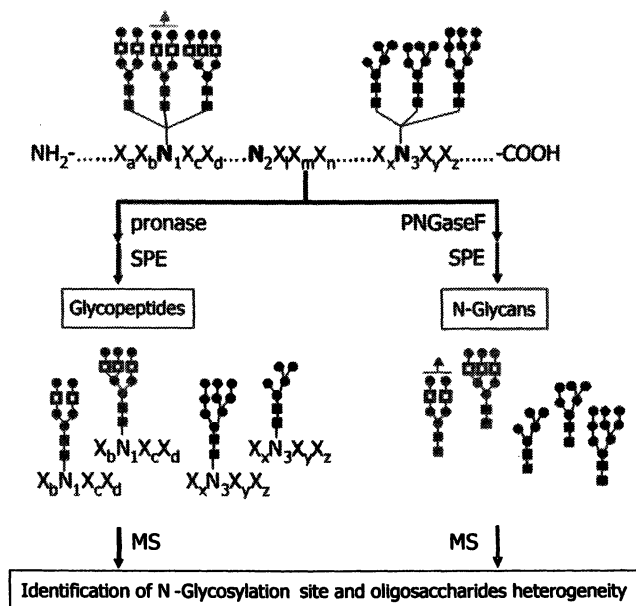


Figure 1. Experimental strategy of pronase digestion.

Pronase digestion

Glycoprotein (1~10 nmol) was dissolved in 0.1 M Tris buffer (pH 7.5) and treated with pronase E (~10 units). The solution was incubated at 37 °C for 24-36 hr. In parallel, a glycoprotein was digested with PNGase F at 37 °C for 12 hr

to release *N*-linked glycans. Glycopeptides produced by nonspecific proteolysis and oligosaccharides released by PNGase F were purified by solid phase extraction (SPE) using a porous graphitized carbon (PGC) cartridge. A PGC cartridge was washed with H₂O followed by 0.05% (v/v) trifluoroacetic acid (TFA) in 80% acetonitrile (AcN)/H₂O (v/v). The solution of digested glycoprotein or oligosaccharide was applied to the PGC cartridge. Subsequently, the cartridge was washed with deionized water at a flow rate of about 1 mL/min to remove salts and buffer. Glycopeptides and glycans were eluted with 10% AcN in H₂O, 20% AcN in H₂O, and 40% AcN in 0.05% TFA in H₂O. Each fraction was collected and concentrated *in vacuo* prior to MALDI analysis.

Mass spectra were recorded on an external source MALDI-FTICR instrument (HiResMALDI, IonSpec Corporation, Irvine, CA) equipped with a 7.0 T superconducting magnet and a pulsed Nd:YAG laser 355 nm. A solution of 2, 5-dihydroxybenzoic acid (DHB) was used as the matrix for all oligosaccharide analyses (0.05 mg/μL in 50% AcN). For glycopeptide analyses, a mixture of DHB and 2, 6-dihydroxyacetophenone was used as the matrix (0.030 mg/μL and 0.015 mg/μL, respectively, in 50% AcN). For negative mode analyses, 1 μL of the glycopeptide or oligosaccharide solution was applied to the MALDI probe, followed by 1 μL of the appropriate matrix solution. The sample was dried under a stream of air and subjected to mass spectrometric analysis. For positive mode analyses, the same sample preparation was applied, except that 1 μL 0.01 M NaCl in 50% AcN was added to the matrix-analyte mixture to enrich the Na⁺ content and produce primarily sodiated species.

Determination of Glycosylation Sites

To illustrate the experimental strategy, a model glycoprotein, ribonuclease B, with a single glycosylation site (⁶⁰N) and known oligosaccharide structures is described (14). Glycopeptides prepared by pronase digestion were analyzed by MALDI-MS in positive mode (Figure 2). The only peaks observed above *m/z* 1000 corresponded to predicted glycopeptides. Large, nonglycosylated peptide fragments were not present since they were digested to their amino acid constituents. Ionic species with *m/z* greater than 1000 were glycopeptides since the oligosaccharide moieties corresponded to at least 800 mass units (GlcNAc₂Man₃). In parallel to this analysis, oligosaccharides were released from ribonuclease B with PNGase F digestion, desalted by PGC, and analyzed by MALDI-MS. Based on the masses, the *N*-linked glycans of ribonuclease B were determined to consist of high-mannose-type GlcNAc₂Man_{*n*} (*n*=5-9) (data not shown). There were two glycopeptide series present in the spectrum, based on peaks that differed by 162 Da (one mannose residue). Doping the sample with NaCl provided additional information. The dopant enhanced the relative intensities of the five peaks in series 2, indicating that the ionic species were

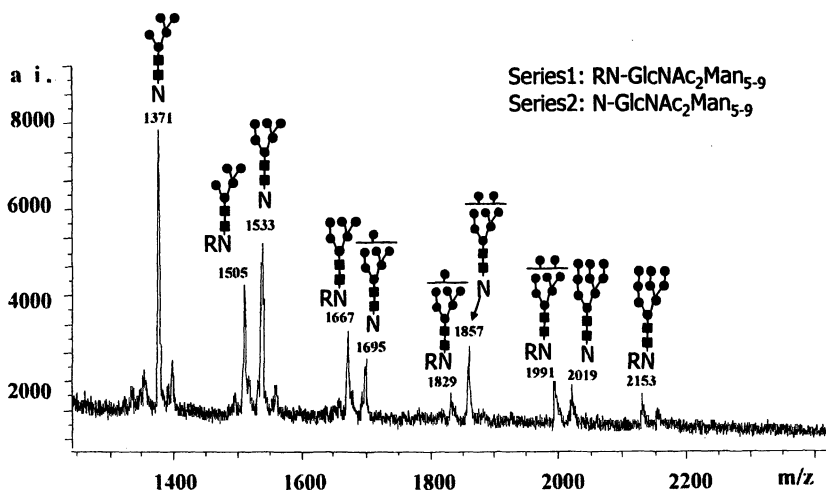


Figure 2. Glycopeptides from pronase digestion of ribonuclease B. Two series are observed corresponding to the dipeptide RN and single amino acid N.

sodium-coordinated (Figure 2). The other series was found to be protonated species whose intensities diminished upon the addition of the NaCl dopant.

The peptide moieties were identified by subtracting the masses of the observed glycans from the masses of the glycopeptides. For example, the first peak of series 1 (m/z 1505.456, Figure 2) corresponded to the glycopeptide with the glycan moiety GlcNAc₂Man₅ (m/z 1257.420). The peptide mass was obtained by subtracting the observed glycan mass 1216.42 Da ($[M+Na]^+$ minus Na^+ and H_2O) from the observed glycopeptide mass 1504.45 Da ($[M+H]^+$ minus H^+). The peptide mass (288.12 D) corresponded to the dipeptide RN (theoretical mass 288.15 D), with the glycosylation site at position ⁶⁰N. The theoretical mass routinely varied by less than 20 ppm from the experimental mass. The glycopeptides of series 1 therefore consisted of the dipeptide RN with high mannose oligosaccharides ranging in size from GlcNAc₂Man₅ to GlcNAc₂Man₉. It should be noted that while ribonuclease B produced mainly dipeptides, other glycoproteins produced peptides ranging in lengths from three to 10 even at long reaction periods allowing specific site determination. The lowest mass of series 2 (m/z 1371.39) corresponded to the glycopeptide with the oligosaccharide component GlcNAc₂Man₅ (m/z 1257.42). The peptide mass of 132.07 D corresponded to asparagine (theoretical mass 132.05 D). Glycopeptides of series 2 therefore contained a single N residue. Oligosaccharide heterogeneity found on ⁶⁰N of ribonuclease B was shown in Figure 3. In addition, the relative abundances of glycan species with the peptide RN could be determined from the

relative ion abundances. As these glycans have similar structures and sizes, their ionization efficiencies are similar and permit calculation of the relative oligosaccharides abundances in the glycoprotein (15).

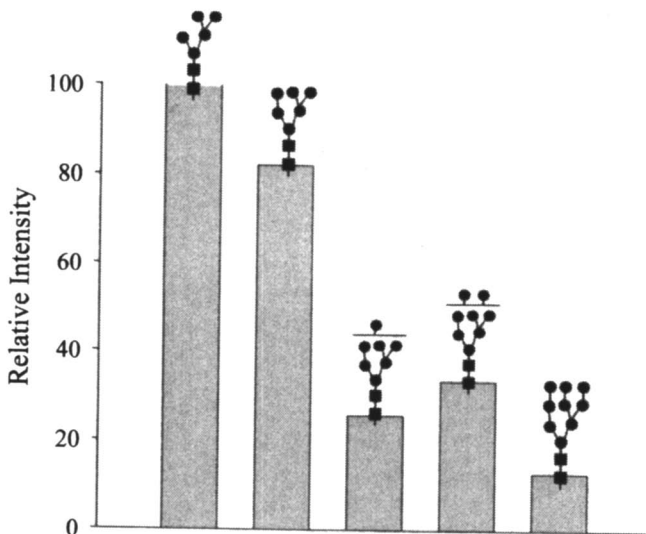


Figure 3. Oligosaccharides found on ^{60}N of ribonuclease B and their relative abundances.

The strategy described allows the specific sites of glycosylation to be established and yields the inherent site heterogeneity information from simple to complicated glycoproteins. Model glycoproteins such as ribonuclease B and chicken ovalbumin (with two potential sites, one of which is occupied) yielded the correct results (16). For examples of more complicated glycoproteins with unknown glycosylation sites, cortical granule lectins (CGL1 and CGL2) from *Xenopus laevis* (XL) eggs and CGL from *X. tropicalis* (XT) eggs were examined. The distribution of N-linked glycans on the glycosylation sites of XL CGL1 is shown in Figure 4. The complete profile of glycosylation sites of more complex glycoprotein, glucose oxidase, containing seven unknown potential glycosylation sites was also achieved using this procedure.

Glycoproteomics Tool: GlycoX

A new computer program, GlycoX, was developed to aid in the interpretation of the resulting data (17). The software was developed in

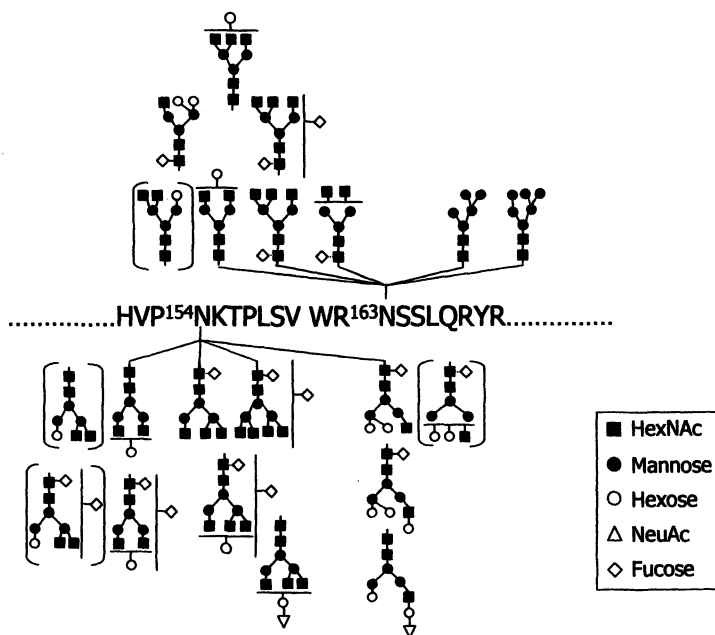


Figure 4. Distribution of the glycans on *N*-linked glycosylation sites of XL CGL1.

MATLAB and requires the entry of mass spectra as ASCII files. The program makes use of accurate masses belonging to the glycopeptides and glycans with the known protein sequence for assigning *N*-linked and *O*-linked glycosylation with the additional determination of microheterogeneity. The program automatically interprets the mass spectra of glycopeptides prepared by nonspecific or reduced-specificity proteolysis (e.g., with pronase E or proteinase K). A schematic flowchart of GlycoX is shown in Figure 5.

The mass intensity (*M/I*) table (saved in ASCII format) and the corresponding glycoprotein sequences from the SWISS-PROT/TrEMBL database (in FASTA format, saved as a text file) were entered into GlycoX. To determine the glycosylation sites and the accompanying glycans, the isotope filtered masses were then evaluated to determine whether they corresponded to combinations of glycan and peptide masses. The peptide composition is determined from the calculated mass by comparing all possible sequences from the monopeptide to the user-specified maximum peptide length. For glycan masses, the program generates possible combinations that are consistent with the given mass. Alternatively, masses from the glycan profile spectra can be used.

For complicated glycosylation, use of the measured glycan profile is preferred. To validate the program, several glycoproteins were characterized. These glycoproteins range from the simple, with one site of glycosylation, to the more complex, with at least seven sites of glycosylation.

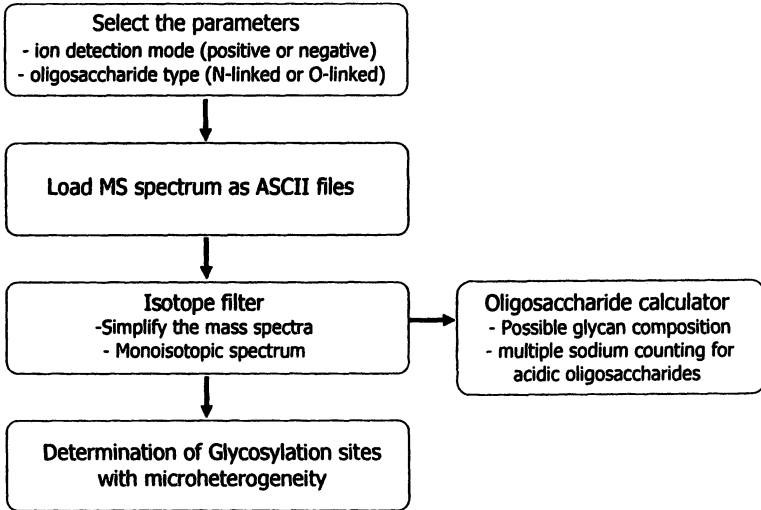


Figure 5. Schematic flowchart of GlycoX.

GlycoX includes an oligosaccharide calculator for assigning oligosaccharide compositions for N-linked, O-linked, and chemically modified oligosaccharides. An isotope filter is also included to determine the monoisotopic quasimolecular ions in mixtures of glycopeptides or glycans readily.

Conclusion

The new procedure described in this book chapter will substantially improve the ease and efficiency for the determination and quantitation of glycosylation sites with oligosaccharide heterogeneity. Additionally, the pronase digestion can be utilized to overcome some limitation of trypsin digestion, which is currently used. We have performed experiments using immobilized pronase that show them to be more effective in producing glycopeptides. Sufficient glycopeptides could be obtained after 1.5 hours to produce the desired results. The program GlycoX aids in the interpretation of MS data from a non-specific protease treatment. We employed primarily FT-MS for the analysis, however, similar

analyses can be performed on other high mass accuracy instruments that are now becoming widely available. The approach of nonspecific protease digestion with the automated data analysis using the GlycoX computer program is potentially adaptable to a high-throughput format for large-scale glycoproteomic investigation.

Acknowledgements

We gratefully acknowledge the financial support by the National Institutes of Health (R01 GM049077). We also thank Drs. Jerry Hedrick and Thomas Peavy for providing the sample of the *Xenopus cortical granule lectin*.

References

1. Varki, A., *Glycobiology* **1993**, *3*, 97-130.
2. Juhasz, P.; Martin, S. A., *Int J Mass Spectrom* **1997**, *169*, 217-230.
3. An, H. J.; Peavy, T. R.; Hedrick, J. L.; Lebrilla, C. B., *Anal Chem* **2003**, *75*, 5628-5637.
4. Imre, T.; Schlosser, G.; Pocsfalvi, G.; Siciliano, R.; Molnar-Szollosi, E.; Kremmer, T.; Malorni, A.; Vekey, K., *J Mass Spectrom* **2005**, *40*, 1472-1483.
5. Mormann, M.; Paulsen, H.; Peter-Katalinic, J., *Eur J Mass Spectrom* **2005**, *11*, 497-511.
6. Hagglund, P.; Bunkenborg, J.; Elortza, F.; Jensen, O. N.; Roepstorff, P., *J Proteome Res* **2004**, *3*, 556-566.
7. Medzihradsky, K. F.; Maltby, D. A.; Hall, S. C.; Settineri, C. A.; Burlingame, A. L., *J Am Soc Mass Spectr* **1994**, *5*, 350-358.
8. Krokhin, O.; Ens, W.; Standing, K. G.; Wilkins, J.; Perreault, H., *Rapid Commun Mass Sp* **2004**, *18*, 2020-2030.
9. Zhang, H.; Li, X. J.; Martin, D. B.; Aebersold, R., *Nat Biotechnol* **2003**, *21*, 660-666.
10. Bykova, N. V.; Rampitsch, C.; Krokhin, O.; Standing, K. G.; Ens, W., *Anal Chem* **2006**, *78*, 1093-1103.
11. Temporini, C.; Perani, E.; Calleri, E.; Dolcini, L.; Lubda, D.; Caccialanza, G.; Massolini, G., *Anal Chem* **2007**, *79*, 355-363.
12. Dezutterdambuyant, C.; Schmitt, D. A.; Dusserre, N.; Hanau, D.; Kolbe, H. V. J.; Kieny, M. P.; Gazzolo, L.; Mace, K.; Pasquali, J. L.; Olivier, R.; Schmitt, D., *Res Virology* **1991**, *142*, 129-138.
13. Bezouska, K.; Sklenar, J.; Novak, P.; Halada, P.; Havlicek, V.; Kraus, M.; Ticha, M.; Jonakova, V., *Protein Sci* **1999**, *8*, 1551-1556.

14. Fu, D. T.; Chen, L.; Oneill, R. A., *Carbohydr Res* **1994**, *261*, 173-186.
15. Cancilla, M. T.; Wang, A. W.; Voss, L. R.; Lebrilla, C. B., *Anal Chem* **1999**, *71*, 3206-3218.
16. Suzuki, T.; Kitajima, K.; Emori, Y.; Inoue, Y.; Inoue, S., *P Natl Acad Sci USA* **1997**, *94*, 6244-6249.
17. An, H. J.; Tillinghast, J. S.; Woodruff, D. L.; Rocke, D. M.; Lebrilla, C. B., *J Proteome Res* **2006**, *5*, 2800-2808.

Chapter 12

Chemical Approaches to Glycobiology

Nicholas J. Agard

Department of Pharmaceutical Chemistry, University of California at San Francisco, San Francisco, CA 94158–2552

Post-translational modifications play fundamental roles in regulating protein structure, localization, and function. Presented ubiquitously on eukaryotic proteins, these modifications provide a layer of structural and functional complexity beyond what is available from the genome. Despite their essential roles, post-translational modifications remain challenging to study in large part because researchers cannot apply traditional genetic manipulations to these non-templated structures. This lack of direct genetic control has driven researchers to pursue chemical approaches to modify and characterize these modifications. Here I will discuss one chemical method, the bioorthogonal chemical reporter strategy, which has been used extensively to study the most common post-translational modification, glycosylation. I will focus first on ways to metabolically introduce monosaccharides bearing unnatural functionality into cellular glycoconjugates. This will be followed by a survey of chemical methods used to modify these unnatural glycans. Lastly, I will highlight recent applications of these chemical ligations in the profiling, proteomic identification, and fluorescence imaging of glycosylation.

Introduction

Glycosylation is a ubiquitous post-translational modification affecting nearly 50% of all proteins and greater than 90% of those that traffic through the secretory pathway (1, 2). Glycans mediate both intra- and intercellular interactions and regulate diverse biological processes, including development (3), homeostasis (4), and immunity (5). Cell surface carbohydrates are a primary marker of a cell's physiological state. Changes in glycosylation are known to correlate with inflammatory (6), pathogenic (7, 8), and carcinogenic signals via the dynamic action of membrane recycling and glycan biosynthetic enzymes (i.e., glycosidases and glycosyltransferases) (9, 10). Thus, it is possible for a cell to remodel its glycoconjugates in the absence of *de novo* translation.

Despite the paramount importance of glycans in cellular regulation, details of their function remain poorly understood in part because carbohydrate structures are not directly genetically encoded. In eukaryotes, carbohydrates (with the exception of the β -O-GlcNAc modification) are attached to protein scaffolds as they traffic through the endoplasmic reticulum and the Golgi apparatus (11). Glycosyltransferases and glycosidases act sequentially on these scaffolds, but the addition of each monosaccharide relies on stochastic binding events. The glycans that emerge from this ordered, non-uniform process are heterogeneous, yet they contain common underlying structures. This non-templated biosynthesis has complicated elucidation of carbohydrate function in two important ways. First, heterogeneity in glycan expression has frustrated attempts to establish direct structure-function relationships involving individual carbohydrate epitopes. Second, glycans are not amenable to the direct genetic manipulations that facilitate protein purification and labeling. As a result, it remains difficult to visualize glycoconjugates or inventory their expression in living organisms.

The detection and purification of glycoconjugates has been addressed through the development of a two-step strategy, termed the bioorthogonal chemical reporter strategy (12, 13). In this approach, a monosaccharide bearing an unnatural functional group (the chemical reporter) is metabolically incorporated into glycoconjugates (Figure 1) (14). The chemical reporter is then covalently tagged by exogenous reagents appended to detection or affinity probes. Successful labeling depends on both the capacity of biosynthetic enzymes to accept the unnatural substrate and the functional group tolerance and specificity of the subsequent reaction. Enzymes involved in monosaccharide activation and glycan biosynthesis are highly specific and will tolerate only very slight modifications to the native structure. This specificity prohibits direct incorporation of fluorescent molecules or epitope tags and limits the size of the chemical reporter to simple functional groups. In addition, the covalent chemistry used must rapidly and exclusively label the unnatural glycans under

conditions compatible with the target glycoconjugates. In the most restricted systems, living cells and animals, this means the reactions must take place at 37 °C in aqueous solvent and in the presence of competing functionalities, including amines, thiols, and molecular oxygen. Chemical reporters that successfully traverse biosynthetic pathways and react selectively under physiological conditions despite competing functionality are termed “bioorthogonal”.

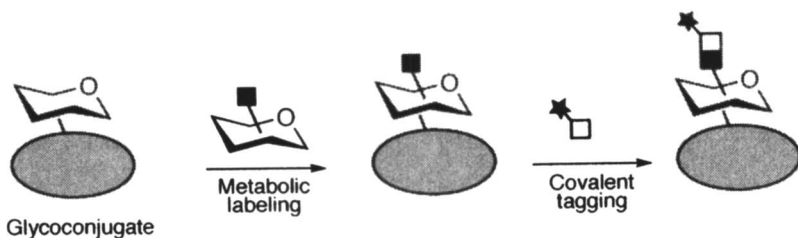


Figure 1. The bioorthogonal chemical reporter strategy. First, unnatural monosaccharides are metabolically incorporated into glycoconjugates. Next, the unnatural functionality (black square) is covalently tagged using appropriately derivatized reagents (white square).

Here I will review the bioorthogonal chemical reporter strategy. First, I will outline glycan biosynthetic pathways and their tolerance for unnatural monosaccharides. Next, I will discuss the corresponding chemical methods for their derivatization, including hydrazone/oxime formation, thiol alkylation, Staudinger ligation, and [3+2] azide-alkyne cycloadditions (both copper-catalyzed and strain-promoted). Last, I will outline the current applications of the bioorthogonal chemical reporter strategy.

Metabolic Labeling of Unnatural Monosaccharides

Glycoproteins are biosynthesized by the transfer of monosaccharides from activated nucleotide-sugar donors to target proteins (15, 16). Nearly all of the activated sugars can be synthesized *de novo* from glucose and glucosamine (17). However, salvage pathways can also recycle monosaccharides from degraded glycoproteins and glycolipids (18). In some cell lines, salvaged monosaccharides compose up to 80% of the total glycans on the cell surface (19). These salvage pathways can be taken advantage of to incorporate unnatural functionalities into glycans. Indeed, most unnatural monosaccharides are activated via homologous salvage pathways.

Salvage pathways process free sugars in a stepwise fashion to form activated nucleotide donors. For three sugars, galactose (Gal), fucose (Fuc) and *N*-acetylgalactosamine (GalNAc), the salvage pathway begins by anomeric phosphorylation of the free sugar (18). Condensation of the phosphosugar with a nucleotide triphosphate gives the activated nucleotide sugar (Figure 2). Salvage of three additional monosaccharides, *N*-acetylglucosamine (GlcNAc), glucose (Glc), and mannose (Man), begins with kinase-mediated phosphorylation of the 6-hydroxyl to give the 6-phosphosugar (20). This compound is in turn isomerized and condensed with the nucleotide triphosphate to give the activated donor sugar.

In contrast to other monosaccharides, activated sialic acid donors are biosynthesized from *N*-acetylmannosamine (ManNAc) or directly from sialic acids (Sia), including *N*-acetylneuraminic acid (NeuAc), via a more complex pathway (21). ManNAc is phosphorylated at the 6-hydroxyl group and condensed with phosphoenolpyruvate to give *N*-acetylneuraminic acid-9-phosphate (NeuAc-9-P). Phosphate ester hydrolysis is followed by direct condensation with CTP to give CMP-NeuAc (Figure 3). Sialic acids can intercept this pathway directly via enzymatic reaction with CTP.

Diverse unnatural monosaccharides are tolerated by these metabolic pathways (22-34). However, only molecules modified with bioorthogonal chemical reporters are capable of subsequent bioorthogonal ligation. Four functional groups have been used for this purpose: azides, alkynes, ketones, and thiols (25, 26, 32, 34, 35). While ketones and thiols occur naturally on the interior of the cell and are therefore not bioorthogonal in the strictest sense, these functionalities are not highly abundant on the cell surface. Thus, selective conjugation is possible in this location. Numerous bioorthogonally functionalized monosaccharides have been incorporated into cellular glycoconjugates, including derivatives of ManNAc, Sia, GalNAc, GlcNAc, and Fuc (Figure 4) (23, 25, 26, 28-30, 32-34, 36).

The tolerance of each pathway to unnatural functionalized monosaccharides depends on its unique suite of enzymes. Introduction of unnatural monosaccharides into glycans relies on the promiscuity of the salvage pathway enzymes and glycosyltransferases. The efficiency of this incorporation has been determined both by glycan analysis, which determines what percentage of the total monosaccharide content is composed by the unnatural analog, and by *in vitro* kinetic analysis of the associated enzymes. Cellular sialic acid content is analyzed by fluorescence detection on a reversed-phase HPLC after liberation of the labile sialosides via mild acid hydrolysis and subsequent condensation with 1,2-diamino-4,5-methylenedioxybenzene (DMB). Analysis of cell lines incubated with azide-derivatized ManNAc (*N*-azidoacetylmannosamine (ManNAz)) showed variable incorporation of the metabolized unnatural monosaccharide, ranging from 4-41% of the total sialosides (37). This high variability results from a number of cell-line-dependent factors, including the

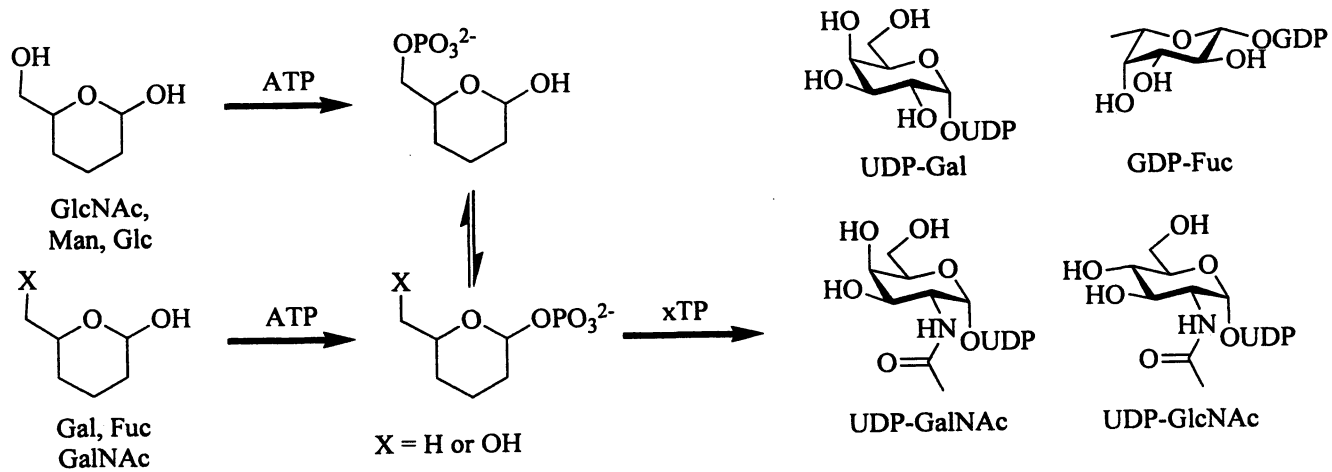


Figure 2. Monosaccharides are metabolized via salvage pathways to give activated nucleotide sugars. Man and Glc (not pictured) are activated via the same pathway as GlcNAc.

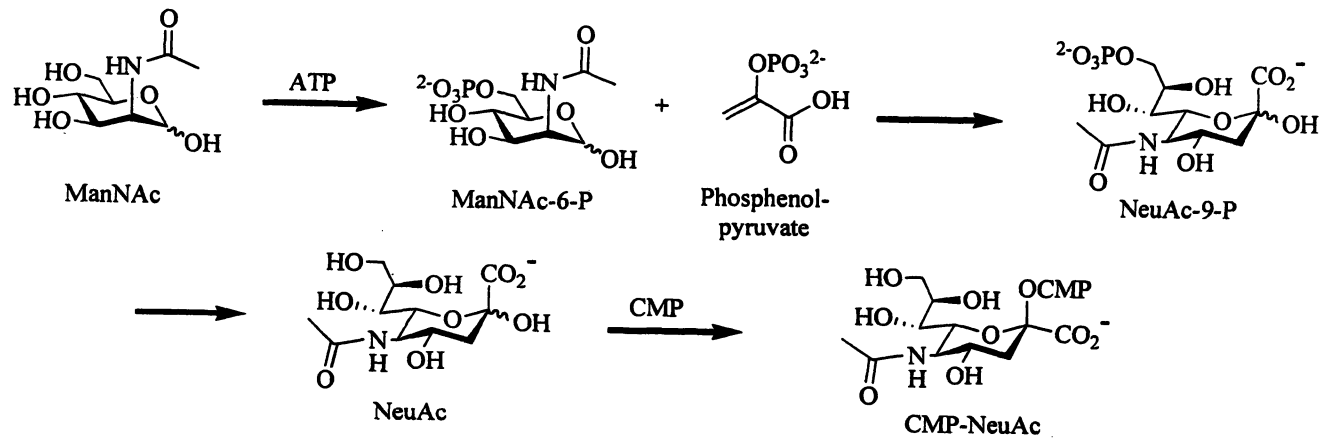


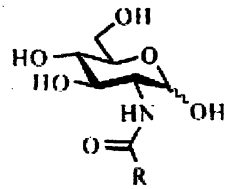
Figure 3. The biosynthesis of CMP-sialic acid. Two free sugars, ManNAc and NeuAc, intercept the pathway and are subject to replacement with unnatural analogs.

native concentrations of competing metabolites, the rate of cell surface turnover, and the ability of metabolic enzymes to accommodate the unnatural monosaccharides. In living organisms, substrate incorporation is thought to be further limited by tissue accessibility and metabolic degradation of the substrates. Glycan analysis of murine heart tissue showed that ~2% of the sialic acid was replaced by the azide-containing sialosides (SiaNAz) (38). Western blot analysis found azide-containing glycoproteins in the heart, kidney, liver, and spleen, though not in the brain or thymus. Mice treated with *N*-azidoacetylglactosamine (GalNAz), an analog of GalNAc, showed similar organ distribution, though no quantitative analysis of GalNAz incorporation has been performed (39).

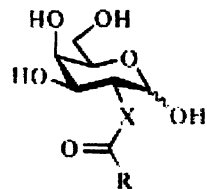
The enzymatic tolerance for an azide-derived glucosamine analog (*N*-azidoacetylglucosamine (GlcNAz)) has been assayed directly. Each enzyme in the salvage pathway showed ~3.5-fold lower efficiency (k_{cat}/K_M) for GlcNAz compared to GlcNAc. This difference appears to be almost entirely due to the less efficient binding (K_M) of the unnatural analog (40). Similarly, Khidekel and coworkers showed a ketone-derived UDP-GalNAc analog (UDP-2-acetyl-2-deoxygalactose) was a poor substrate for β -1,4-galactosyltransferase, though a single mutation in the glycosyltransferases resulted in highly efficient transfer (41). Furthermore, Hang and coworkers showed that UDP-GalNAz is a substrate for multiple polypeptide GalNAc transferases (ppGalNAcTs) with k_{cat} values ranging from ~15-50% of the k_{cat} values for the natural substrate (42). The K_M values of ppGalNAcT-1 for UDP-GalNAz and UDP-GalNAc were equivalent within experimental error. These data suggest that intercepting the salvage pathway at downstream steps should increase the efficiency of metabolic incorporation by limiting the number of enzymes that must tolerate the modification. In biochemical assays, activated nucleotide sugars are likely to be efficient substrates for glycosyl transferases. In living systems, however, these highly charged molecules are unlikely to traverse the cell membrane and gain access to the secretory compartments. To the best of our knowledge, there are currently no published applications using synthetic functionalized nucleotide sugars in living cells.

Cellular incorporation of unnatural monosaccharides is limited not only by enzyme specificity but also substrate availability. Monosaccharides are highly hydrophilic with poor membrane permeability. Thus, high extracellular concentrations of free monosaccharide (1-10 mM) are necessary to achieve efficient labeling of glycoproteins and cell surfaces (25). Peracetylated monosaccharides increases their membrane permeability and thus their intracellular concentration. Inside the cell, these sugars are deacetylated by non-specific cytosolic esterases and incorporated into cellular glycans via the salvage pathway (43). Increasing cell permeability allows much lower concentrations of

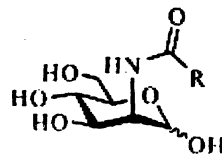
Unnatural Sugars



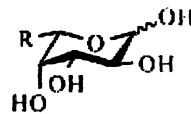
GlcNAc



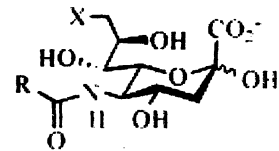
GalNAc



ManNAc



Fuc



Sia

References detailing the metabolic incorporation of bioorthogonally functionalized monosaccharides:

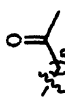
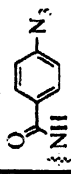
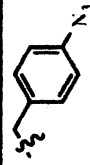

R =	Monosaccharide					
	GlcNAc	GalNAc	ManNAc	Fuc	Sia	
CH_2N_3	30, 51	28, 39 X = NH	26, 27, 38	31, 32, 33	23 X = OH	
	23 n = 2		23, 25 n = 2,3		23 n = 2-4 X = OH	
CH_3		29 X = CH_2			23, X = 	
CH_2SH			34			
			23		23 X = OH	
			33 n = 2	32, 33 n = 0		

Figure 4. Literature precedents for the incorporation of functionalized monosaccharides into cellular glycans.

unnatural peracetylated sugars (10-100 μM) to efficiently label glycoconjugates (26).

Covalent Tagging of Chemical Reporters Using Bioorthogonal Chemistries

The second step in the bioorthogonal chemical reporter strategy is the covalent attachment of epitope tags to the labeled sugar. Acceptable conditions for bioorthogonal reactions vary with the substrate (e.g., isolated proteins, cell lysates, fixed cells, living cells, or living organisms). However, each reaction must be highly specific and rapid under conditions that solvate and do not damage the target. Typically, this includes aqueous solvent, neutral pH, and reaction at 37 °C (14). Numerous reactive functionalities are present in living systems, including nucleophilic thiols and amines, electrophilic ketones and disulfides, and both molecular and enzymatic oxidants (e.g., O_2 and cytochrome P450s). Given the high biological concentrations of these functionalities, non-specific labeling would result in substantial background. In addition, chemical or enzymatic degradation of either the chemical reporter or its reaction partner can dramatically reduce labeling efficiency.

A key to efficient bioorthogonal ligation is a rapid reaction rate. Biomolecules are frequently present at lower concentrations (nM or μM) than those used in typical organic reactions (mM). Concentrations of exogenous reagents are limited by solubility and toxicity to the system. While reaction yields on isolated materials can frequently be increased with longer reaction times, the reaction duration in living systems is limited by the stability of the biomolecule and/or the dynamics of the biological phenomenon under investigation. Thus, effective glycoconjugate labeling must take place in the span of minutes or hours instead of days.

To date, four bioorthogonal reactions have been used to label glycans on cells and in lysates: hydrazone/oxime formation with ketones, thiol alkylation with maleimides, Staudinger ligation of azides with triaryl phosphines, and copper-catalyzed or strain-promoted [3+2] cycloadditions of alkynes and azides (Figure 5) (25, 26, 32, 34, 35). While each reaction has been used extensively, most recent applications have employed azido- or alkynyl-sugars due to their superior metabolic incorporation and efficient ligations.

Historically, ketosugars, including derivatives of ManNAc, Sia, and GlcNAc, were the first unnatural monosaccharide endowed with chemical reporters to be metabolically incorporated into cells (23, 26). Hydrazides and aminoxy compounds react specifically and essentially irreversibly with ketones in aqueous solution. However, this reaction is optimal in slightly acidic conditions ($\sim\text{pH}$ 5.5) that are detrimental to cell viability. Furthermore, ketones,

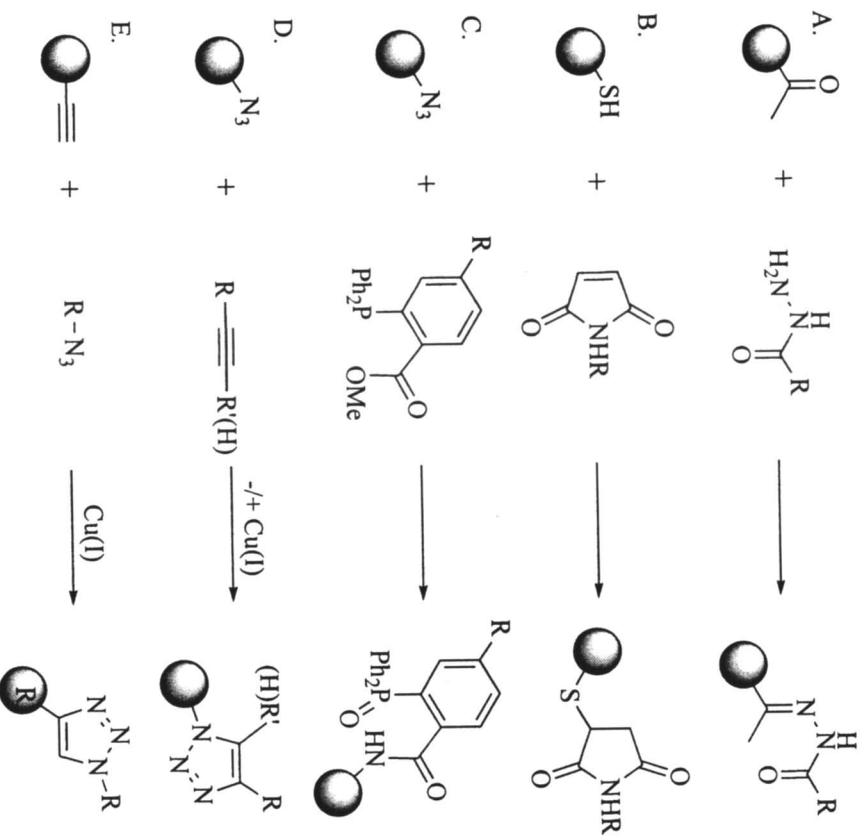


Figure 5. Bioorthogonal reactions on sugars. **A.** Ketones react with hydrazides to give hydrazones. **B.** Thiols undergo Michael Reaction with maleimides. **C.** Azides undergo Staudinger ligand with phosphines or **D.** strain-promoted or copper catalyzed [3+2] cycloaddition with alkynes. **E.** Alkynes undergo copper-catalyzed [3+2] cycloaddition with azides.

while virtually absent from the cell surface, are present in a number of intracellular metabolites which could result in background labeling. Despite these liabilities, hydrazone/oxime bioconjugations continue to be used in experiments with limited competing functionalities, including proteomic labeling of lysates and derivatization of isolated proteins. Commercial availability of diverse hydrazide conjugates coupled with the synthetic accessibility of ketones will likely ensure this ligation's continued widespread use.

Thiol-maleimide conjugation is a well-known protein conjugation methodology that has only recently been applied to unnatural sugars. Thioglycolylmannosamine (ManNTGc) can be readily incorporated into glycoconjugates expressed on the cell surface (34). However, in the oxidizing environment outside of the cell, the sulfur atoms are largely complexed as disulfides and unreactive to electrophilic reagents. Exposure of ManNTGc treated cells to reducing agents (i.e., triscarboxyethylphosphine (TCEP)), reduces the disulfide bonds and activates the thiols for further reaction. While the reaction with maleimides is rapid under physiological conditions, cysteine thiols are common components of proteins. On the surface of ManNTGc-treated cells, the majority of TCEP-exposed thiols are due to the unnatural sugar and can be specifically detected by maleimide conjugation. Intracellularly, millimolar concentrations of free thiols would swamp out the ManNTGc signal. Thus, this technique is not applicable to labeling intracellular glycans or lysates, nor for situations (such as proteomic labeling) where high specificity is essential.

Three bioorthogonal reactions have been developed for selective tagging of azide-containing glycoconjugates. The first of these, the Staudinger ligation, is a modification of the classic Staudinger reduction. In the ligation, an intramolecular electrophilic ester captures the nucleophilic aza-ylide intermediate prior to hydrolysis (Figure 6) (44). Following hydrolysis, this process results in an amide linkage between the phosphine reagent and the azide-bearing glycan with concomitant oxidation of the phosphine. The Staudinger ligation has been used for tagging azide-labeled cell lysates, live cell surfaces, and even tissues of living animals (26, 38, 39, 45). Despite its extensive use, the Staudinger ligation is somewhat limited by its sluggish reaction rate and non-specific oxidation of the requisite phosphines. Degradation of the reagents *in situ* is particularly troublesome in experiments when quantitative labeling is essential (e.g., proteomics). Physical organic studies attempting to increase the rate of the Staudinger ligation found that nucleophilic attack on the azide is the rate-determining step of the reaction. Introduction of electron-donating groups onto the phenyl rings increased the phosphine nucleophilicity and subsequently the rate of the reaction (44). However, the increased nucleophilicity was accompanied by an increased rate of non-specific oxidation, decreasing the utility of these modified reagents.

A second reaction used for bioorthogonal labeling of glycans is a copper-catalyzed modification of the classic Huisgen-type alkyne-azide [3+2]

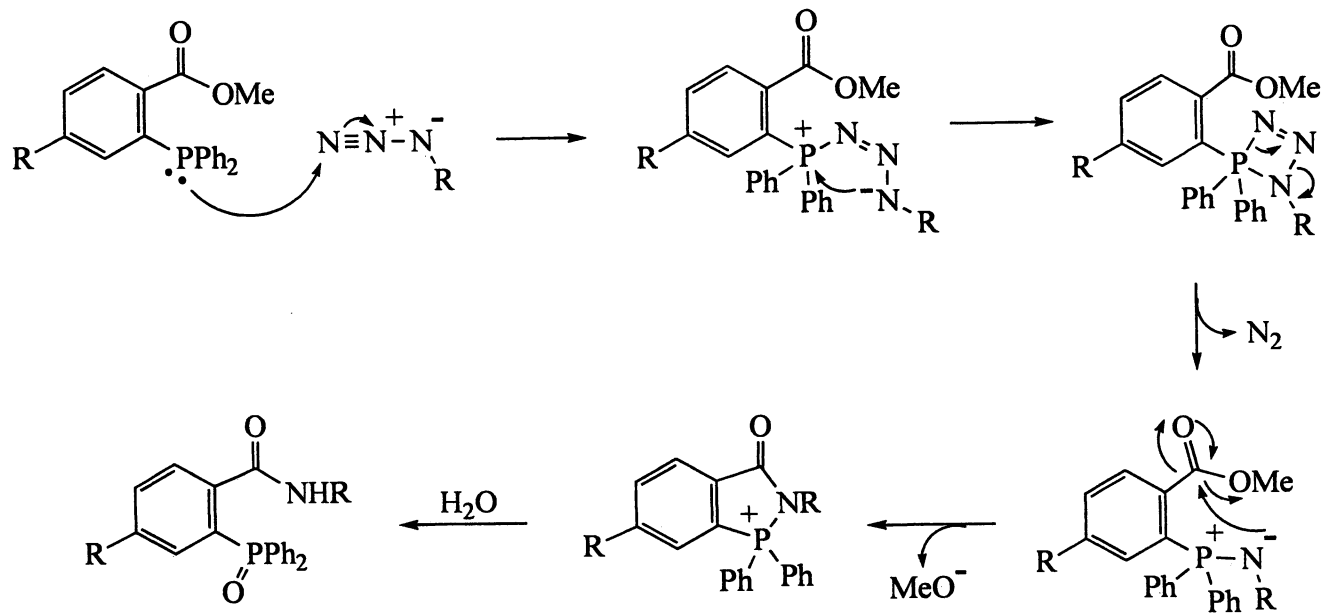


Figure 6. Mechanism of the Staudinger ligation.

cycloaddition (46), better known as click chemistry. This remarkable reaction, developed independently by Sharpless and Meldal, rapidly tags labeled sugars in the presence of endogenous biological functionality (47, 48). Catalytic copper (I) reacts with a terminal alkyne to form a copper acetylide under mild conditions (49). This metal complex reacts rapidly with the azide to form the 1,4-substituted triazole product and regenerate the catalytic copper. This reaction is markedly more efficient than the Staudinger ligation, enabling near stoichiometric labeling of low-abundance substrates. Click chemistry has the added advantage that both components, azides and alkynes, are sufficiently small and robust to traverse biosynthetic pathways. Consequently, azides and alkynes can each serve as the chemical reporter. An alkyne derivative was found to be both more efficiently incorporated and less toxic than the corresponding azido sugar in the fucose salvage pathway (33). The sole drawback to this chemistry is that the copper catalyst is cytotoxic, preventing its use in experiments where cell viability is essential (50).

To address toxicity concerns, a copper-free variant of the [3+2] cycloaddition was developed which accelerates reactions with azides through the use of ring strain (35, 50). Constraining the alkyne to an eight-membered ring results in a significant bond angle deformation of 17° and nearly 18 kcal/mol of strain energy (35). A portion of this strain energy is released in the transition state of the [3+2] cycloaddition, accelerating the rate of the reaction compared to the unstrained version. The first reported cyclooctyne reagent was somewhat unstable and hydrophobic, reacting with organic azides with second-order rate constants approximately equal to those of the Staudinger ligation. Subsequent derivatives have addressed stability, solubility and reactivity concerns (Figure 7) (50, 51). Of particular note, introduction of a *gem*-difluoro substituent adjacent to the alkyne increases the second-order rate constant of the reaction by ~30 fold (51). This electron-withdrawing group decreases the energy of the molecule's LUMO, resulting in more efficient orbital overlap with the HOMO of the azide and an increased rate of reaction. In labeling experiments, this difluorinated cyclooctyne, termed DIFO, tagged azide-bearing proteins as efficiently as click chemistry. In addition, this molecule shows efficient labeling of living cells and animals without apparent toxicity.

The optimal choice of the three azide-specific bioorthogonal chemistries is dependent on the context of the experiment. These reactions have been compared via NMR analyses of reactions with isolated small molecules, Western blot analyses of protein labeling in lysates, and flow cytometry analyses of cultured and primary cells (50, 51). Direct kinetic comparison of click chemistry to the other bioorthogonal chemistries is complicated by reaction order (two for the Staudinger ligation and strain-promoted cycloadditions, higher for click chemistry). Functionally, however, click chemistry is the fastest of the reactions in cell lysates, and the optimal reaction for *in vitro* labeling of biomolecules or fixed cells. In addition, click reagents are commercially available, facilitating

their use by the non-specialist. However, copper's cytotoxicity precludes the use of this reaction on live cells or in living organisms. DIFO reactions approximate the speed of click chemistry in Western blot analysis and are non-toxic and highly effective in flow cytometry and live cell imaging applications (51). Inside of living mice, however, DIFO shows reduced potency potentially due to poor solubility or enzymatic degradation. In addition, DIFO is the product of a lengthy multi-step synthesis, rendering it less available than the alternative technologies. The Staudinger ligation, despite its relatively low second-order rate constant, remains the most efficient reagent inside living mice. Potential advances, including the development of tight-binding ligands that reduce copper's cytotoxicity or further optimization of the cyclooctynes, could render a

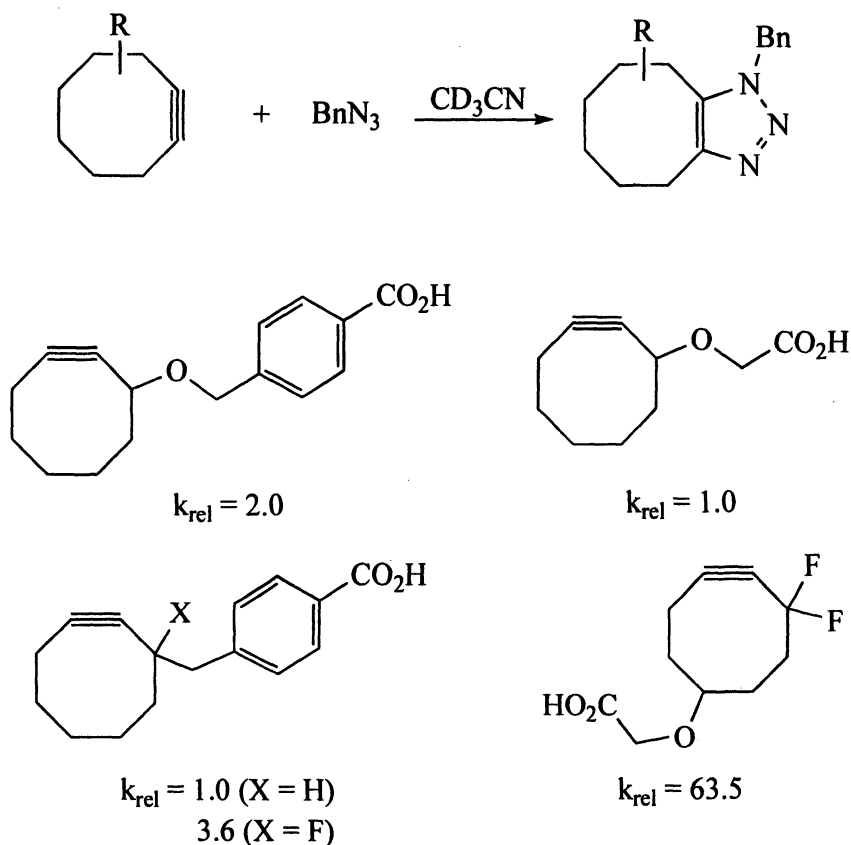


Figure 7. The collection of cyclooctyne reagents and relative second-order rate constants for reaction with benzyl azide.

single ligation the best for bioconjugation. Currently, however, each of these reagents remains relevant to the bioorthogonal labeling of azides.

Applications of Metabolic Engineering

Broadly speaking, the bioorthogonal chemical reporter strategy has been used for two applications in glycobiology: glycan profiling and glycan imaging. Tagging bioorthogonally labeled glycans with epitope tags enables rapid analysis of metabolic flux and/or isolation of these glycoconjugates from complex mixtures of biomolecules. Isolated biomolecules can be subjected to further analysis to identify the collection of labeled species. Alternatively, covalent attachment of detection tags allows rapid and direct visualization of changes in glycosylation state in cellular or organismal analysis.

Metabolic turnover of mucin-type *O*-linked glycosylation has been studied using the bioorthogonal chemical reporter strategy. Cell lines incubated with peracetylated GalNAz (Ac₄GalNAz) and tagged via the Staudinger ligation showed distinct glycoproteins and differential levels of azides displayed on the cell surface (28). This result suggests that incorporation of the unnatural sugar is a cell-specific phenomenon, which can vary widely from cell type to cell type. Mice injected with Ac₄GalNAz also incorporate GalNAz (39). Analysis of murine splenocytes found that B-cells incorporate the unnatural sugar more efficiently than T-cells. Surprisingly, lectin analysis of total *O*-linked glycosylation showed T-cells to have greater quantities of *O*-linked glycosylation than B-cells, suggesting that the unnatural sugars monitor dynamic glycosylation instead of steady-state levels. This hypothesis was confirmed by incubating primary splenocytes with ¹⁴C-labeled Ac₄GalNAc. In these native cells the labeled sugar once again labeled B-cells more robustly.

Bioorthogonal tagging has been used to isolate *O*-GlcNAcylated proteins for use in proteomic assays. Intracellular *O*-GlcNAc is a reciprocal modification to phosphorylation and a key regulator of cellular homeostasis (4). Nuclear pore protein p62 is known to be extensively *O*-GlcNAcylated (52). Immunoprecipitated p62 from Ac₄GlcNAz treated cells was efficiently labeled via the Staudinger ligation, in contrast to p62 from Ac₄GlcNAc-treated cells (40). Comprehensive analysis of *O*-GlcNAcylated proteins was performed by tagging Ac₄GlcNAz-treated cell lysate with a biotinylated phosphine via the Staudinger ligation (53). The tagged proteins were purified by affinity chromatography and analyzed by mass spectrometry. This procedure identified 199 *O*-GlcNAcylated proteins, of which 12% had been previously identified. Western blot analysis on 10% of the identified proteins confirmed that they were indeed *O*-GlcNAc-modified.

In addition to its uses in glycan profiling and proteomic purification, the bioorthogonal chemical reporter strategy can be used for direct imaging of glycans. Sawa and coworkers were able to image fucosylation in fixed cells by labeling them with an azide-bearing fucose derivative and tagging the unnatural sugar with a fluorescent dye via click chemistry (32). Fluorescence imaging of the tagged cells showed labeling on the cell surface and punctate staining inside the cell. Colocalization with wheat germ agglutinin (WGA) showed the intracellular staining to be located in the Golgi apparatus. Subsequent work by Hsu and coworkers extended this fixed cell labeling to alkyne-bearing fucose and mannosamine derivatives (33). Again, labeling was present on the cell surface and in the Golgi. However, the use of the alkyne-bearing sugars enabled tagging with an activatable azido-coumarin which fluoresced upon ligation. This modification noticeably increased the signal-to-noise for probe detection and thus increased sensitivity for the comparatively rare fucose modifications.

While click chemistry has enabled static imaging of fixed cells, dynamic imaging of glycans would allow direct visualization of changes as they progress in response to physiological stimuli. Real-time imaging has been accomplished using Alexa Fluor conjugates of DIFO (51). SiaNAz-labeled glycans on the cell surface can be tagged with fluorescent DIFO conjugates in as little as one minute. Subsequent fluorescence analysis showed rapid retrograde transport into the Golgi in approximately 5 minutes, and later slow (~ 12 h) transfer of the labeled glycoconjugates into lysosomes. Sequential labeling every 24 h, with serially red-shifted dyes, showed that this labeling pattern is continuously repeated until all of the fluorescent dyes colocalize into the lysosome. This technology may be applicable to imaging development or infection by identifying newly synthesized glycans and metabolic "hot spots" from the steady-state levels of glycosylation that are detected via antibody or lectin staining.

The ultimate application of each of these technologies is understanding and detecting the role of glycosylation in human health and disease. In basic research, identifying changes in glycan and glycoprotein expression may help to elucidate the mechanism of physiological changes. The bioorthogonal chemical reporter strategy facilitates separating the glycoconjugates of interest from complex biological mixtures. Subsequent analysis of these glycoconjugates can identify changes in low-abundance substrates that otherwise would be masked by the preponderance of irrelevant biomolecules. In addition, this approach is able to access dynamic changes in glycosylation even if steady-state levels remain relatively constant. In this way, it may be possible to identify glycoconjugates which are rapidly biosynthesized and degraded, or whose glycans are regularly recycled.

In addition to biochemical approaches to identifying dynamic glycosylation and glycoconjugates in biological mixtures, activated alkynes and phosphines can be used for direct imaging of glycosylation. In principle, these reagents can

be used for histological analysis, providing alternative staining methodologies and increasing the scope of targets from classic antibody- or lectin-based methods. *In vivo* imaging of glycosylation may facilitate detecting disease or tracking the progression of both normal and diseased physiological states. Following changes in glycosylation may help to elucidate the mechanism of these processes.

Both glycan profiling and glycan imaging could be translated to clinical settings. Cancerous cells are known to secrete highly glycosylated proteins (54). Regular screening of patients for serum glycoproteins found to be associated with cancer might enable early detection and more successful treatment. Similarly, many cancer lines have elevated levels of glycosylation on their cell surface, which may be detected directly via conjugation of fluorescent, positron emission tomography (PET), or magnetic resonance imaging (MRI) probes to bioorthogonal reagents.

Conclusions

The bioorthogonal chemical reporter strategy continues to enable the study of biological systems that were previously inaccessible using conventional experimental techniques. This straightforward approach (metabolically label the biomolecules of interest, covalently tag the unnatural sugar, and detect/purify the biomolecules via the attached biophysical probe) can be applied to a large swath of research questions. Currently inaccessible classes of glycans might be targeted by new monosaccharide analogs or cell-permeable derivatives of downstream metabolites (i.e., protected phospho-sugars or nucleotide sugars). Novel bioorthogonal reaction partners would create a plethora of potential new monosaccharides which might be incorporated via salvage pathways. Additionally, improved chemistries with increased sensitivity and biocompatibility could expand the applications of these reagents.

Acknowledgements

I would like to thank Jeremy Baskin, Isaac Miller, Jennifer Prescher, David Rabuka, and Michael Schelle for their critical reading of this manuscript and Carolyn Bertozzi for invaluable advice and support.

References

1. Apweiler, R.; Hermjakob, H.; Sharon, N. *Biochim Biophys Acta* **1999**, *1473*, 4-8.

2. Jung, E.; Veuthey, A. L.; Gasteiger, E.; Bairoch, A. *Proteomics* **2001**, *1*, 262-268.
3. Grobe, K.; Ledin, J.; Ringvall, M.; Holmborn, K.; Forsberg, E.; Esko, J. D.; Kjellen, L. *Biochim Biophys Acta* **2002**, *1573*, 209-215.
4. Slawson, C.; Housley, M. P.; Hart, G. W. *J Cell Biochem* **2006**, *97*, 71-83.
5. Hanisch, F. G.; Ninkovic, T. *Curr Protein Pept Sci* **2006**, *7*, 307-315.
6. Parish, C. R. *Nat Rev Immunol* **2006**, *6*, 633-643.
7. Herrmann, M.; von der Lieth, C. W.; Stehling, P.; Reutter, W.; Pawlita, M. *J Virol* **1997**, *71*, 5922-5931.
8. Suzuki, Y. *Biol Pharm Bull* **2005**, *28*, 399-408.
9. Pelham, H. R. *Curr Opin Cell Biol* **1991**, *3*, 585-591.
10. Whelan, S. A.; Hart, G. W. *Circ Res* **2003**, *93*, 1047-1058.
11. Hanisch, F. G. *Biol Chem* **2001**, *382*, 143-149.
12. Dube, D. H.; Bertozzi, C. R. *Curr Opin Chem Biol* **2003**, *7*, 616-625.
13. Prescher, J. A.; Bertozzi, C. R. *Cell* **2006**, *126*, 851-854.
14. Prescher, J. A.; Bertozzi, C. R. *Nat Chem Biol* **2005**, *1*, 13-21.
15. McDowell, G.; Gahl, W. A. *Proc Soc Exp Biol Med* **1997**, *215*, 145-157.
16. Trujillo, J. L.; Gan, J. C. *Biochim Biophys Acta* **1973**, *304*, 32-41.
17. Sullivan, F. X.; Kumar, R.; Kriz, R.; Stahl, M.; Xu, G. Y.; Rouse, J.; Chang, X. J.; Boodhoo, A.; Potvin, B.; Cumming, D. A. *J Biol Chem* **1998**, *273*, 8193-8202.
18. Pastuszak, I.; Drake, R.; Elbein, A. D. *J Biol Chem* **1996**, *271*, 20776-20782.
19. Rome, L. H.; Hill, D. F. *Biochem J* **1986**, *235*, 707-713.
20. Niehues, R.; Hasilik, M.; Alton, G.; Korner, C.; Schiebe-Sukumar, M.; Koch, H. G.; Zimmer, K. P.; Wu, R.; Harms, E.; Reiter, K.; von Figura, K.; Freeze, H. H.; Harms, H. K.; Marquardt, T. *J Clin Invest* **1998**, *101*, 1414-1420.
21. Hinderlich, S.; Stasche, R.; Zeitler, R.; Reutter, W. *J Biol Chem* **1997**, *272*, 24313-24318.
22. Murrey, H. E.; Gama, C. I.; Kalovidouris, S. A.; Luo, W. I.; Driggers, E. M.; Porton, B.; Hsieh-Wilson, L. C. *Proc Natl Acad Sci U S A* **2006**, *103*, 21-26.
23. Luchansky, S. J.; Goon, S.; Bertozzi, C. R. *ChemBiochem* **2004**, *5*, 371-374.
24. Mahal, L. K.; Charter, N. W.; Angata, K.; Fukuda, M.; Koshland, D. E., Jr.; Bertozzi, C. R. *Science* **2001**, *294*, 380-381.
25. Mahal, L. K.; Yarema, K. J.; Bertozzi, C. R. *Science* **1997**, *276*, 1125-1128.
26. Saxon, E.; Bertozzi, C. R. *Science* **2000**, *287*, 2007-2010.
27. Saxon, E.; Luchansky, S. J.; Hang, H. C.; Yu, C.; Lee, S. C.; Bertozzi, C. R. *J Am Chem Soc* **2002**, *124*, 14893-14902.
28. Hang, H. C.; Yu, C.; Kato, D. L.; Bertozzi, C. R. *Proc Natl Acad Sci U S A* **2003**, *100*, 14846-14851.

29. Hang, H. C.; Bertozzi, C. R. *J Am Chem Soc* **2001**, *123*, 1242-1243.
30. Vocadlo, D. J.; Hang, H. C.; Kim, E. J.; Hanover, J. A.; Bertozzi, C. R. *Proc Natl Acad Sci U S A* **2003**, *100*, 9116-9121.
31. Rabuka, D.; Hubbard, S. C.; Laughlin, S. T.; Argade, S. P.; Bertozzi, C. R. *J Am Chem Soc* **2006**, *128*, 12078-12079.
32. Sawa, M.; Hsu, T. L.; Itoh, T.; Sugiyama, M.; Hanson, S. R.; Vogt, P. K.; Wong, C. H. *Proc Natl Acad Sci U S A* **2006**, *103*, 12371-12376.
33. Hsu, T. L.; Hanson, S. R.; Kishikawa, K.; Wang, S. K.; Sawa, M.; Wong, C. H. *Proc Natl Acad Sci U S A* **2007**, *104*(8), 2614-2619.
34. Sampathkumar, S. G.; Li, A. V.; Jones, M. B.; Sun, Z.; Yarema, K. J. *Nat Chem Biol* **2006**, *2*, 149-152.
35. Agard, N. J.; Prescher, J. A.; Bertozzi, C. R. *J Am Chem Soc* **2004**, *126*, 15046-15047.
36. Han, S.; Collins, B. E.; Bengtson, P.; Paulson, J. C. *Nat Chem Biol* **2005**, *1*, 93-97.
37. Luchansky, S. J.; Argade, S.; Hayes, B. K.; Bertozzi, C. R. *Biochemistry* **2004**, *43*, 12358-12366.
38. Prescher, J. A.; Dube, D. H.; Bertozzi, C. R. *Nature* **2004**, *430*, 873-877.
39. Dube, D. H.; Prescher, J. A.; Quang, C. N.; Bertozzi, C. R. *Proc Natl Acad Sci U S A* **2006**, *103*, 4819-4824.
40. Vocadlo, D. J.; Bertozzi, C. R. *Angew Chem Int Ed Engl* **2004**, *43*, 5338-5342.
41. Khidekel, N.; Ficarro, S. B.; Peters, E. C.; Hsieh-Wilson, L. C. *Proc Natl Acad Sci U S A* **2004**, *101*, 13132-13137.
42. Hang, H. C.; Yu, C.; Pratt, M. R.; Bertozzi, C. R. *J Am Chem Soc* **2004**, *126*, 6-7.
43. Sarkar, A. K.; Fritz, T. A.; Taylor, W. H.; Esko, J. D. *Proc Natl Acad Sci U S A* **1995**, *92*, 3323-3327.
44. Lin, F. L.; Hoyt, H. M.; van Halbeek, H.; Bergman, R. G.; Bertozzi, C. R. *J Am Chem Soc* **2005**, *127*, 2686-2695.
45. Hang, H. C.; Bertozzi, C. R. *Bioorg Med Chem* **2005**, *13*, 5021-5034.
46. Huisgen, R. *Angew Chem Int Ed Engl* **1963**, *2*, 565-598.
47. Rostovtsev, V. V.; Green, L. G.; Fokin, V. V.; Sharpless, K. B. *Angew Chem Int Ed Engl* **2002**, *41*, 2596-2599.
48. Tornøe, C. W.; Christensen, C.; Meldal, M. *J Org Chem* **2002**, *67*, 3057-3064.
49. Rodionov, V. O.; Fokin, V. V.; Finn, M. G. *Angew Chem Int Ed Engl* **2005**, *44*, 2210-2215.
50. Agard, N. J.; Baskin, J. M.; Prescher, J. A.; Lo, A.; Bertozzi, C. R. *ACS Chem Biol* **2006**, *1*, 644-648.
51. Baskin, J. M.; Prescher, J. A.; Laughlin, S. T.; Agard, N. J.; Chang, P. V.; Miller, I. A.; Lo, A.; Codelli, J. A.; Bertozzi, C. R. **2007**, *287*(5460), 2007-2010.

52. Lubas, W. A.; Smith, M.; Starr, C. M.; Hanover, J. A. *Biochemistry* **1995**, *34*, 1686-1694.
53. Nandi, A.; Sprung, R.; Barma, D. K.; Zhao, Y.; Kim, S. C.; Falck, J. R.; Zhao, Y. *Anal Chem* **2006**, *78*, 452-458.
54. Brockhausen, I. *EMBO Rep* **2006**, *7*, 599-604.

Chapter 13

Automated Solution-Phase Oligosaccharide Synthesis and Carbohydrate Microarrays: Development of Fluorous-Based Tools for Glycomics

Nicola L. Pohl

Department of Chemistry and the Plant Sciences Institute, Iowa State University, Ames, IA 50011

Carbohydrates are increasingly recognized as key players in a range of disease processes, but their study and the development of glycotherapeutics has been limited in part by the lack of the same ready access to well-defined structures as has been achieved by automated synthesis of nucleic acids and peptides. After a review of the challenges of developing automated protocols for oligosaccharide synthesis, a new approach based on solution-phase rather than solid-phase iterative protocols is discussed. This new approach to automated synthesis based on light-fluorous tags also allows the direct formation of carbohydrate microarrays for biological screening of carbohydrate-protein interactions.

Carbohydrates specifically mediate numerous biological processes that include cell adhesion and differentiation, pathogen invasion, tumor cell metastasis, and inflammatory responses and are therefore of great interest for the generation of therapeutics (1-3). Unfortunately, the molecular basis for most of

these carbohydrate interactions is not understood, in part because homogeneous well-defined oligosaccharides are extremely challenging to procure (3). Although pure nucleic acids and peptides can easily be purchased because of robust automated synthesis protocols, an analogous commercial process for oligosaccharide synthesis is still unavailable. The toil of carbohydrate synthesis also limits the range of structures that can be incorporated into microarrays for screening of sugar-protein interactions. Data about these interactions is definitely needed to demystify the role of carbohydrates in the life cycles of organisms as well as to aid the development of sugar-based therapeutics such as vaccines that intervene in these carbohydrate functions.

The Challenges of Automated Oligosaccharide Synthesis

In his attempts to make peptides in the 1960s, Bruce Merrifield had the key insight of attaching an amino acid to a solid support for subsequent coupling to additional amino acid monomer units with masked nucleophilic groups (Figure 1) (4). Although the biphasic reaction conditions require the use of significant excesses of reagents for reasonable coupling times, these excess reagents can be rinsed from the solid resin in a process that is easily automated. The application of this principle has revolutionized the synthesis of peptides and nucleic acids and made these classes of biopolymers readily commercially available for now extensive structure-function studies. In his Nobel Prize acceptance speech in 1984, Prof. Merrifield mentioned that his insights should also be applicable to the synthesis of oligosaccharides (5). Indeed, many researchers have attempted to apply the principles of solid-phase chemistry to the problem of iterative carbohydrate synthesis and in 2001 an automated peptide synthesizer was modified to produce a few oligosaccharides (6). However, after more than 40 years, researchers interested in the functions of this class of biomolecules are still not able to go online and purchase a custom-made oligosaccharide in the manner that they purchase nucleic acids and peptides. Clearly, the synthesis of oligosaccharides is an especially arduous challenge that may well require development of a completely different principle for its automation.

Solid-phase Approaches to Building Carbohydrate Structures

Several factors have likely contributed to the relative difficulty in making oligosaccharides when compared to producing peptides and oligonucleotides. Unlike formation of an amide or phosphate linkage, creation of a glycosidic bond results in a new stereocenter that is often difficult to control. Nonenzymatic carbohydrate synthesis relies on protecting groups to permanently block some

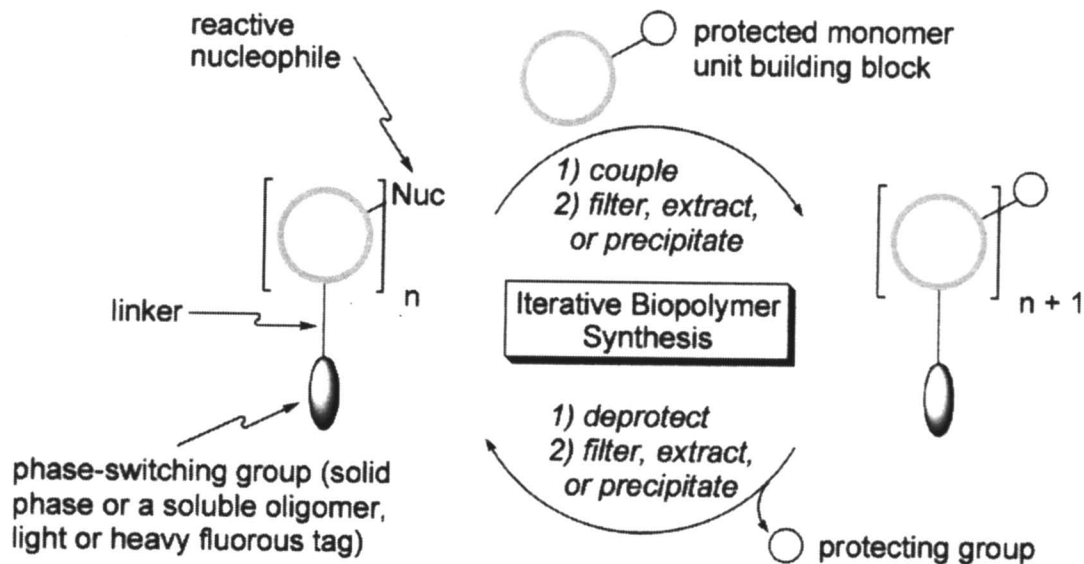


Figure 1. The basic strategy for phase-switching approaches to iterative biopolymer synthesis for simplified purification of intermediates.

reactive functional groups and to temporarily mask future reaction sites. The production of these building blocks require four or even fourteen reaction steps—many more than are required for blocking amino acid building blocks for example. Therefore, the use of large excesses of building blocks to achieve reasonable coupling yields on a solid resin necessitates starting with large excesses of the sugar to make enough building blocks (6). For example to make a trisaccharide, presuming unrealistic 100% reaction yields, with 10 equivalents of building block per coupling cycle, three equivalents end up on the solid-phase and 27 equivalents are relegated to the waste bin. Numerous research groups have managed to carry out syntheses of carbohydrates on a range of solid supports with various coupling chemistries, but research continues in finding ways around the problems associated with this approach (7).

Solution-phase Alternatives to Building Carbohydrate Structures

The problems of solid-phase chemistry are inherent to biphasic reactions and, since protected oligosaccharide chains do not have the same problems with aggregation that plague peptide synthesis and make spatial segregation of these molecules on a solid-phase so appealing for synthesis, a solution-phase approach seemed to be in order. The biggest challenge would be how to automate such an approach to iterative biopolymer synthesis. To avoid poor reaction kinetics inherent to biphasic systems using solid-phase supports, soluble tags with unique physical properties can be attached to the growing biopolymer chain to aid in purification of intermediates by tag precipitation, extraction into a liquid phase, or affinity chromatography/solid-phase extraction (8). Because tag precipitation is not quantitative, tags for extraction methods are more attractive options. Soluble hydrocarbon and fluorocarbon tags have already been employed in the synthesis of some oligosaccharides for over a decade. Attachment of a hydrocarbon (lipid) tag requires a large tag (at least two 12 to 18-carbon-containing chains) for facile separation from untagged compounds on C₁₈-modified supports (9). Unfortunately, these large tags make solubility in reaction solvents and characterization of intermediates by standard solution-phase proton NMR difficult. On the other hand, fluorocarbons will phase separate from not only aqueous solutions but also lipids and conventional organic solvents. Several fluorocarbon chains can be incorporated into a protecting group to allow extraction of the compound containing the “heavy” fluorocarbon tag into a liquid fluorocarbon layer or a single fluorocarbon chain, a “light” fluorocarbon tag, can capture the tagged molecule by fluorocarbon-derivatized silica gel in a solid-phase extraction process (10). The reaction mixture is loaded on fluorocarbon silica gel, untagged compounds are eluted, and then a change of solvent allows elution of the pure tagged compound. Several carbohydrate protecting groups as well as an anomeric activating group for glycosylation reactions have been designed with

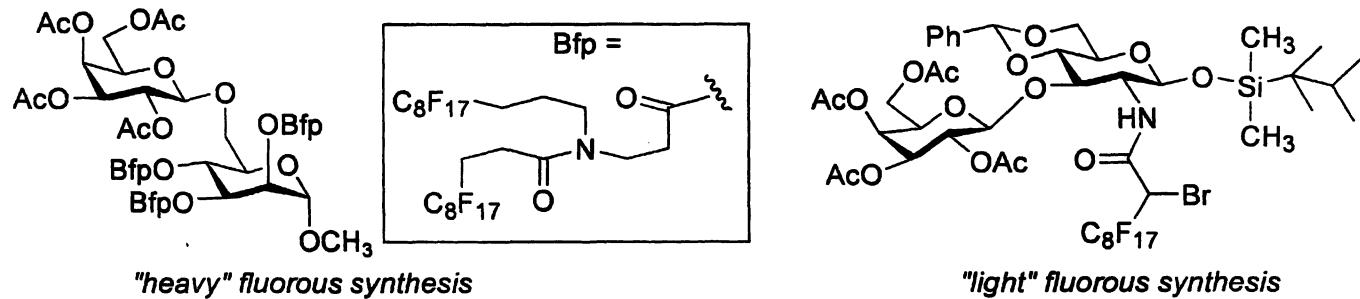


Figure 2. Examples of protecting groups with multiple fluorous tags for "heavy" fluorous synthesis by liquid-liquid extraction (11i) and with single fluorous tags for "light" fluorous synthesis of carbohydrates with fluorous solid-phase extraction of intermediates (11a).

single fluoros tags to simplify purification schemes (Figure 2), but their use is not wide-scale (11-12).

An Automated Solution-phase Strategy for Oligosaccharide Synthesis

Unfortunately, relatively little is known about these "fluorophobic" interactions. Could this binding interaction be robust enough so that even a robot could run a separation process reliably? And, if so, what sort of robotics platform would be needed to carry out not only the solution-phase chemistry, but also these purification steps? Rinsing solid-phase resins is a far easier process.

To probe the strength of noncovalent fluoros-fluoros interactions, the synthesis of oligomannose structures found on the surface of HIV and *Leishmania*-associated parasites was undertaken (13). An easily removable allyl-containing C_8F_{17} tag (14) was first attached to the anomeric carbon of a suitably protected mannose building block (Figure 3). By growing the oligosaccharide chain from the reducing end outward, as nature does, different capping structures can be easily added just as natural sugars are diversified at their ends. From the starting mannose block, the 2-position of mannose was deblocked and glycosylated Schmidt's trichloroacetimidate activation protocol (15) with another building block in successive cycles to form a tetrasaccharide as well as a branched mannose structure. Each time, the fluoros-tagged sugar bound to the fluoros silica gel selectively for facile elution of reaction byproducts and reagents and then was readily eluted for subsequent reaction steps. In addition, purification of fluoros-tagged intermediates was also always possible using standard silica gel chromatography. Thereby, the growing chain could be purified even if a reaction step was not incredibly clean without loss of the purification handle as would be the case with a solid-phase resin.

This fluoros-phase approach has also proved amenable to automation (16). Dozens of automation platforms exist to carry out iterative solid-phase chemistry; none are commercially available to do liquid handling, fluoros-solid-phase extractions, product transfers, and evaporation steps repetitively without intervention. Because heat dissipation and dangerously reactive intermediates are not problems in glycosylation or standard deprotection chemistry, at least microreactors are unnecessary, and standard reaction vessels can be used. A representative scheme for the synthesis of polyrhmannose is depicted in Figure 4. Reaction yields were 80% per reaction step in the automation cycle with use of only 1.5 to two equivalents of building block at each coupling cycle. Therefore, a 10% overall yield can be expected after 2 coupling and 2 deblocking/deprotection reactions. For comparison with solid-phase approaches, even

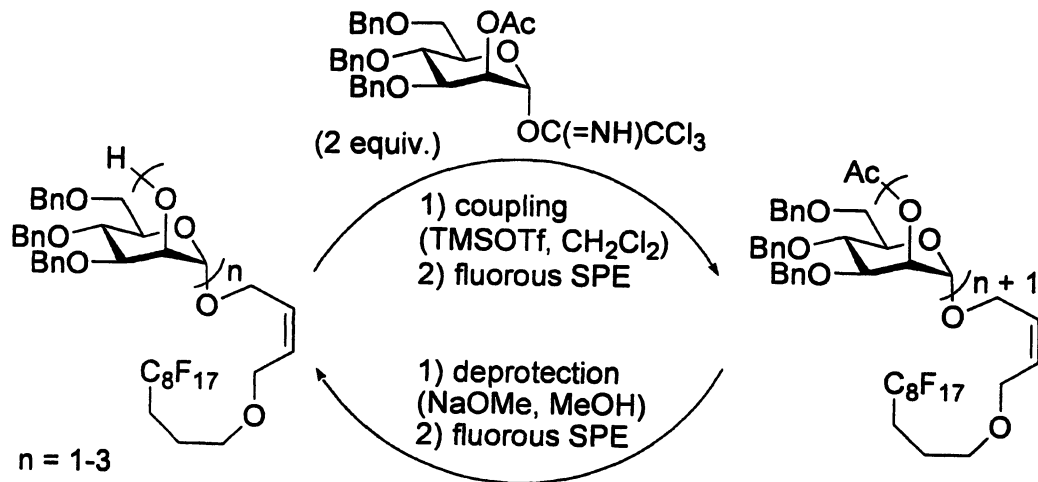


Figure 3. Synthesis of polymannose using a single C_8F_{17} -tag for fluorous solid-phase extraction of intermediates.

presuming 100% yields per reaction, 2 coupling and 2 deblocking/deprotection reactions will only provide a 1% yield based on building block use when 10 equivalents are required in each coupling cycle. Also, unlike solid-phase approaches, the automation cycle can be stopped at any point, and the intermediate can be purified by standard silica gel chromatography methods before mistakes build up and then characterized by proton and carbon NMR. Production of the polyrihamnose fragment from the denoted building block also required only one hour of manual labor with 24 hours of machine time compared to ~24 hours of manual labor required without automation. Polymannose was also synthesized in comparable yields using the same automation protocol.

Direct Formation of Fluorous-based Carbohydrate Microarrays

Fluorous-phase methods can simplify the iterative synthesis of carbohydrates with the added benefits of standard solution-phase reaction monitoring, reduced building block requirements at each coupling cycle, and the possibility of using other purification techniques as needed with intermediates. Noncovalent fluorous-fluorous interactions are even robust enough to allow automation of purification steps between reactions and thereby allow automation of iterative oligosaccharide synthesis. However, even with automated protocols, synthetic carbohydrates ideally would be used frugally to allow numerous protein/antibody screens in the course of glycotherapeutics development. Microarrays such as DNA chips allow such minimal sample usage and therefore have spurred innovative methods for protein-detection assays on glass slides (17). Some of these technologies have been applied to the fabrication of carbohydrate chips (18), but almost all of these microarray methods rely on covalent attachment of compounds to the slide and therefore require distinct functional handles. Given a new method for the automated synthesis of oligosaccharides that relies on noncovalent rather than covalent attachment for phase-switching behavior, the question becomes whether such noncovalent fluorous-fluorous interactions are also strong enough to support a fluorous-based noncovalent strategy for attachment of carbohydrates to an array format (Figure 5).

To test the suitability of using noncovalent fluorous-fluorous interactions for surface patterning, several sugars were synthesized with a fluorous-tagged allyl linker (19). The allyl linkers were then transformed to more flexible and less reactive alkyl linkers by palladium-mediated hydrogenation. After synthesis of the requisite fluorous-tagged carbohydrates (Figure 6), a suitable fluorinated surface was required. An optically and fluorescently clear surface was obtained by reaction of a glass microscope slide with a fluoroalkylsilane for subsequent spotting using a standard arraying robot.

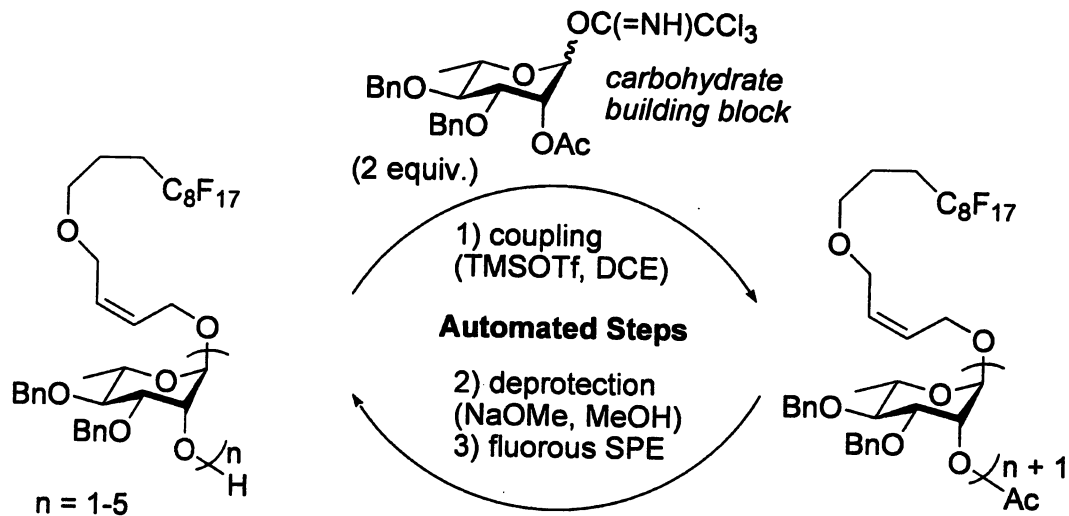


Figure 4. Automated solution-phase synthesis of polyrhamnose.

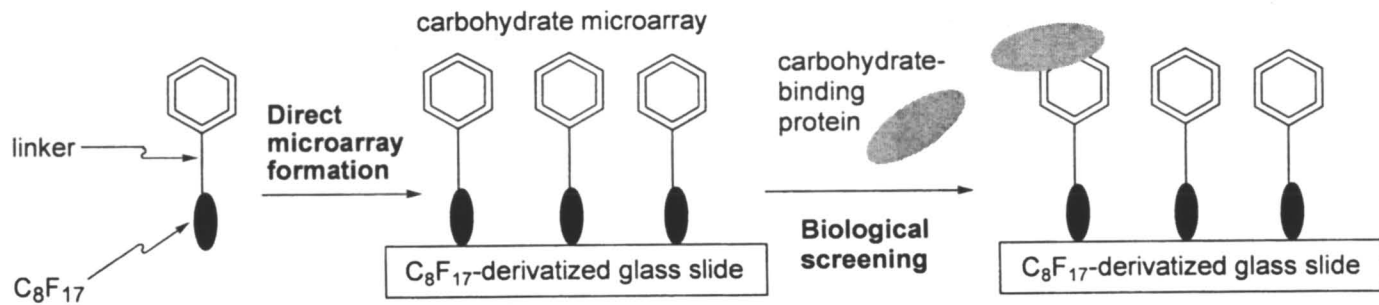


Figure 5. A strategy for use of a single fluororous tag to allow direct formation of microarrays for biological screening.

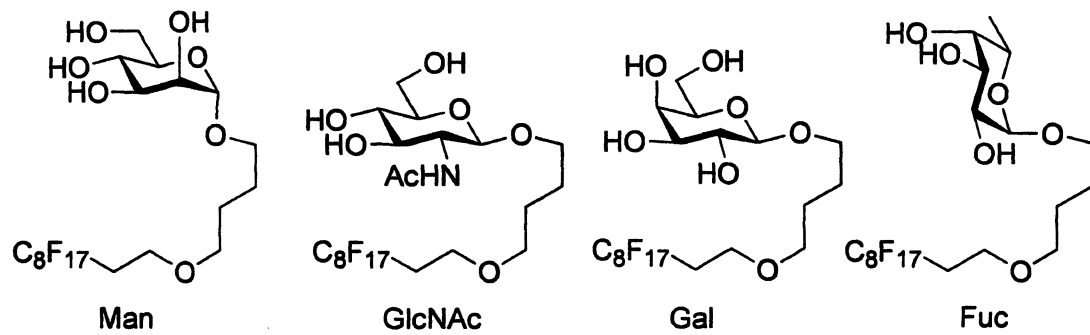


Figure 6. Fluorous-tagged carbohydrates synthesized for microarray formation.

The spotted slide was then incubated for 20 minutes with a solution of fluorescein isothiocyanate-labeled concanavalin A (FITC-ConA), rinsed repeatedly with assay buffer and distilled water, and then scanned with a standard fluorescent slide scanner (Figure 7). The scan clearly showed binding of FITC-ConA only to the mannose-containing spots. The anomeric position also could be discriminated as the beta-linked GlcNAc was not bound by the lectin. An assessment of the ability of the array to withstand detergents often included in biological screens was appraised with the labeled plant lectin from *Erythrina cristagalli* (FITC-ECA) and the hydrocarbon detergent Tween-20. Because hydrocarbons phase separate from fluorocarbons, the noncovalent fluororous-based array should be more stable to regular detergents than a noncovalent array method based on hydrocarbon-hydrocarbon interactions. Indeed, the fluororous-based microarray tolerated the 20 minute incubation time and repeated rinsing with this detergent-containing buffer. Of note, however, is that the noncovalent fluororous-fluororous interactions are not stable enough to use under the continuous flow conditions required by surface plasmon resonance or quartz crystal microbalance measurements (20).

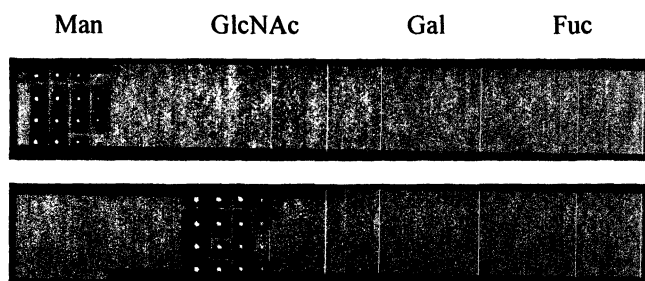


Figure 7. Fluorescence images of arrayed carbohydrates assayed with FITC-labeled lectins. Columns of 4 spots each of 2, 1, 0.5 and 0.1 mM carbohydrates were incubated for 20 min with FITC-ConA (top) or FITC-ECA with 1% TWEEN-20 detergent (bottom) with BSA. (Reproduced with permission from reference 19. Copyright 2005 American Chemical Society.)

Carbohydrate arrays that include disaccharides as well as a charged amino sugar have also been produced with this strategy (14). In addition, a fluororous tag capped with an aminooxy functionality allows the incorporation of reducing sugars directly into a fluororous-based microarray (Figure 8) (21). Applications of these fluororous-based strategies for microarray formation to larger saccharides as well as other compounds such as glycopeptides and nucleic acids can easily be imagined.

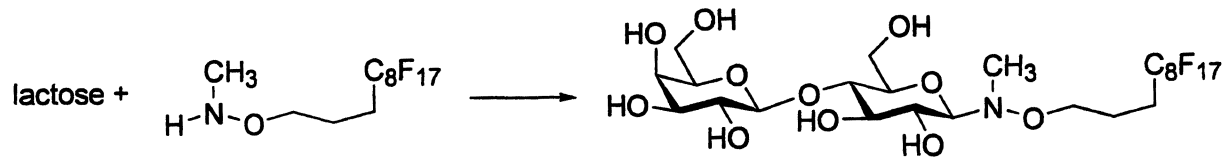


Figure 8. An aminoxy fluorous tag for incorporation of reducing sugars into a noncovalent fluorous-based microarray platform.

Future Directions

Noncovalent fluororous-fluororous interactions do appear to be robust enough to support robotically programmed compound separation of growing saccharide chains based on a fluororous tag. The ability to directly form compound microarrays with a fluororous-tail makes a fluorocarbon-based phase-switching approach to compound synthesis even more appealing. Reactions can be monitored using traditional solution-phase techniques and, unlike solid-phase approaches, large excesses of reagents are not required for high yields. An automated solution-phase-based synthesis of oligosaccharides also allows purification of intermediates prior to the final cleavage of the phase-tag using other methods if necessary to prevent the excessive build up of undesired reaction products. Extensions to a range of oligosaccharides ultimately will test the robustness and utility of these new fluororous-based surface patterning and automated solution-phase synthesis approaches to drive new discoveries in glycobiology.

References

1. Dube, D. H.; Bertozzi, C. R. *Nat. Rev. Drug Discov.* **2005**, *4*, 477-488.
2. Rudd, P. M.; Wormald, M. R.; Dwek, R. A. *Trends Biotechnol.* **2004**, *22*, 524-530.
3. Seeberger, P. H.; Werz, D. B. *Nature*, **2007**, *446*, 1046-1051.
4. Merrifield, R. B. *J. Am. Chem. Soc.* **1963**, *85*, 2149-2154.
5. Merrifield, B. *Biosci. Rep.* **1985**, *5*, 353-376.
6. Plante, O. J.; Palmacci, E. E.; Seeberger, P. H. *Science*, **2001**, *291*, 1523-1527.
7. For recent examples see: (a) Ako, T.; Daikoku, S.; Ohtsuka, I.; Kato, R.; Kanie, O. *Chem. Asian J.* **2006**, *1*, 798-813. (b) Wu, J.; Guo, Z. *J. Org. Chem.* **2006**, *71*, 7067-7070. (c) Jonke, S.; Liu, K. G.; Schmidt, R. R. *Chemistry*, **2006**, *12*, 1274-1290. (d) Dondoni, A.; Marra, A.; Massi, A. *Angew. Chem. Int. Ed.* **2005**, *44*, 1672-1676. (e) Ojeda, R.; Terenti, O.; de Paz, J. L.; Martin-Lomas, M. *Glycoconj. J.* **2004**, *21*, 179-1795. For reviews of approaches to solid-phase oligosaccharide synthesis see: (f) Grathwohl, M.; Drinnan, N.; Broadhurst, M.; West, M. L.; Meuterms, W. *Methods Enzymol.* **2003**, *369*, 248-267. (g) Seeberger, P. H.; Haase, W. C. *Chem. Rev.* **2000**, *100*, 4349-4394.
8. Ito, Y.; Manabe, S. *Chem. Eur. J.* **2002**, *8*, 3077-3084 and references therein.
9. (a) Bauer, J.; Rademann, J. *J. Am. Chem. Soc.* **2005**, *127*, 7296-7297. (b) Nilsson, U. J.; Fournier, E. J.; Hindsgaul, O. *Bioorg. Med. Chem.* **1998**, *6*, 1563-1575.

10. (a) Zhang, W. Fluorous Chemistry for Synthesis and Purification of Biomolecules: Peptides, Oligosaccharides, Glycopeptides, and Oligonucleotides. In *Current Fluoroorganic Chemistry: New Synthetic Directions, Technologies, Materials, and Biological Applications*; Soloshonok, V. A.; Mikami, K.; Yamazaki, T.; Welch, J. T.; Honek, J. F., Eds. ACS Symposium Series 949; American Chemical Society: Washington, DC, 2007, p. 207-220. (b) Curran, D. P. Separations with Fluorous Silica Gel and Related Materials. In *The Handbook of Fluorous Chemistry*; Gladysz, J.; Horváth, I.; Curran, D. P.; Wiley-VCH: Weinheim, 2004; pp 101-127. (c) Horváth, I. T. *Acc. Chem. Res.* 1998, 31, 641-650.
11. (a) Manzoni, L.; Castelli, R. *Org. Lett.* 2006, 8, 955-957. (b) Miura, T.; Satho, A.; Goto, K.; Murakami, Y.; Imai, N.; Inazu, T. *Tetrahedron: Asymmetry* 2005, 16, 3-6. (c) Miura, T.; Goto, K.; Waragai, H.; Matsumoto, H.; Hirose, Y.; Ohmae, M.; Ishida, H.-k.; Satoh, A.; Inazu, T. *J. Org. Chem.* 2004, 69, 5348-5353. (d) Manzoni, L.; Castelli, R. *Org. Lett.* 2004, 6, 4195-4198. (e) Jing, Y.; Huang, X.; *Tetrahedron Lett.* 2004, 45, 4615-4618. (f) Manzoni, L. *Chem. Commun.* 2003, 2930-2931. (g) Miura, T.; Goto, K.; Hosaka, D.; Inazu, T. *Angew. Chem. Int. Ed.* 2003, 42, 2047-2051. (h) Palmacci, E. E.; Hewitt, M. C.; Seeberger, P. H. *Angew. Chem. Int. Ed.* 2001, 40, 4433-4437. (i) Miura, T.; Hirose, Y.; Ohmae, M.; Inazu, T. *Org. Lett.* 2001, 3, 3947-3950. (j) Röver, S.; Wipf, P. *Tetrahedron Lett.* 1999, 40, 5667-5670. (k) Curran, D. P.; Ferrito, R.; Hua, Y. *Tetrahedron Lett.* 1998, 39, 4937-4940.
12. For a recent review of fluorous-phase approaches to oligosaccharide synthesis see: Pohl, N. L. Carbohydrate Microarrays and Fluorous-Phase Synthesis: Interfacing Fluorous-Phase Tags with the Direct Formation of Glycoarrays. In *Current Fluoroorganic Chemistry: New Synthetic Directions, Technologies, Materials, and Biological Applications*; Soloshonok, V. A.; Mikami, K.; Yamazaki, T.; Welch, J. T.; Honek, J. F., Eds. ACS Symposium Series 949; American Chemical Society: Washington, DC, 2007, pp. 261-270.
13. Jaipuri, F. A.; Pohl, N. L. *Abstracts of Papers.* 230th National Meeting of the American Chemical Society, Washington, DC, Aug 28-Sept 1, 2005; American Chemical Society: Washington, DC, 2005, ORGN-185.
14. Mamidyala, S. K.; Ko, K.-S.; Jaipuri, F. A.; Park, G.; Pohl, N. L. *J. Fluorine Chem.* 2006, 127, 571-579.
15. Schmidt, R. R.; Jung, K.-H. *Trichloroacetimidates in Carbohydrates in Chemistry and Biology, Part I: Chemistry of Saccharides*, Vol 1 (B. Ernst, G.W. Hart, P. Sinay, Eds.) Wiley-VCH, 2000, Weinheim, pp. 5-59.
16. Park, G.; Pohl, N. L. *Abstracts of Papers.* 233th National Meeting of the American Chemical Society, Chicago, IL, March 25-29, 2007; American Chemical Society: Washington, DC, 2007, ORGN-629.

17. (a) Uttamchandani, M.; Wang, J.; Yao, S. Q. *Mol. BioSyst.* **2006**, *2*, 58-68.
(b) Tomizaki, K.; Usui, K.; Mihara, H. *ChemBioChem* **2005**, *6*, 782-799.
18. For recent reviews on carbohydrate microarrays see: (a) Horlacher, T.; Seeberger, P. H. *OMICS* **2006**, *10*, 490-498. (b) Culf, A. S.; Cuperlovic-Culf, M.; Oullette, R. J. *OMICS* **2006**, *10*, 289-310. (c) Larsen, K.; Thygesen, M. B.; Guillaumie, F.; Willats, W. G.; Jensen, K. J. *Carbohydr. Res.* **2006**, *341*, 1209-1234. (d) Paulson, J. C.; Blixt, O.; Collins, B. E. *Nat. Chem. Biol.* **2006**, *2*, 238-248.
19. Ko, K.-S.; Jaipuri, F. A.; Pohl, N. L. *J. Am. Chem. Soc.* **2005**, *127*, 13162-13163.
20. Camci-Unal, G.; Pohl, N. L. *Abstracts of Papers*. 233th National Meeting of the American Chemical Society, Chicago, IL, March 25-29, 2007; American Chemical Society: Washington, DC, 2007, CARB-118.
21. Chen, G.-S.; Pohl, N. L. *Abstracts of Papers*. 233th National Meeting of the American Chemical Society, Chicago, IL, March 25-29, 2007; American Chemical Society: Washington, DC, 2007, CARB-138.

Chapter 14

Polypeptide-Based Glycopolymers for the Study of Multivalent Binding Events

**Ronak Maheshwari, Shuang Liu, Brian D. Polizzotti, Ying Wang,
and Kristi L. Kiick***

**Department of Materials Science and Engineering, University of Delaware,
and the Delaware Biotechnology Institute, Newark, DE 19716**

Important recognition events in biology are mediated by multivalent interactions between relevant oligosaccharides and multiple saccharide receptors present on lectins, viruses, toxins, and cell surfaces. Accordingly, considerable effort has been devoted toward the development of multivalent polymeric ligands for carbohydrate-binding proteins. This work describes the synthesis of new polypeptide-based, galactose-functionalized glycopolymers, via a combination of protein engineering and chemical methods. These methodologies permit control over the number and the spacing of saccharides on a scaffold, as well as the conformation of the glycopolymer backbone. Glycopolypeptides were characterized via ^1H NMR, MALDI TOF mass spectrometry, SDS-PAGE analysis, circular dichroic spectroscopy, and spectrophotometric assays. They were tested as inhibitors of the binding of the cholera toxin B subunit via direct enzyme-linked assays. These experiments confirm that near quantitative modification of the polypeptides is possible and that the conformations of the polypeptides are not significantly altered upon glycosylation. The binding assays indicate the relevance of appropriate saccharide spacing on controlling the binding event, and that other polymer chain attributes, such as chain extension, also play an important role in binding. These experiments offer insights in the design of linear glycopolymers for the manipulation of multivalent binding events.

Introduction

Many biological processes depend on molecular recognition and binding events that occur between proteins and carbohydrates, either on cell surfaces or in isolation. In particular, protein-carbohydrate interactions are responsible for various processes such as cell signaling, differentiation and growth, and adhesion of various viral and bacterial toxins (1, 2). These processes occur via multivalent oligosaccharide-mediated binding events, as multivalency results in improvements in binding often referred to as the "cluster glycoside effect" (3). Considerable effort has been devoted to the understanding and development of high affinity multivalent ligands, both for inhibition of carbohydrate and protein interactions as well as for the selective activation of desired cell-signaling processes (4-12).

The general study of multivalent binding to a variety of targets has been of wide interest in the development of therapeutic molecules, and many research groups have synthesized oligo- and multivalent inhibitors of various forms including linear polymers, dendritic macromolecules, or small molecule based ligands (5, 6, 12-18). Although these investigations have elucidated general architectural guidelines that are believed to govern multivalency in protein-carbohydrate interactions, a more detailed understanding of multivalent binding by polymeric materials has been hindered by the lack of architectural control possible in chemically derived glycopolymers. Such detailed investigation requires the development of well defined glycopolymers in which architectural features are better defined and controlled. We have therefore employed a combination of protein engineering and chemical approaches to develop well-defined random coil and α -helical glycopolypeptides that offer control in the placement and valence of saccharide groups on the polypeptide chain, and have assessed differences in their binding, on the basis of their architectures, to the B pentamer subunit of the cholera toxin (CT B₅).

Cholera toxin is produced by the *Vibrio cholerae* bacterium and has an AB₅ architecture; other examples of the AB₅ toxin family include shiga-like toxins and heat labile enterotoxins. The A subunit of AB₅ bacterial toxins is responsible for the deleterious (for the host cell) enzymatic activity while the five B subunits, which have identical saccharide binding sites, mediate adhesion to the host cell. The B subunits of cholera toxin bind specifically to the ganglioside GM1 (Gal β 1-3GalNAc β 1-4(Neu5Ac α 2-3)Gal- β 1-Glc-ceramide) present on the surface of intestinal epithelial cells. After entry into the cell via endocytosis, the A subunit is cleaved and initiates the catalytic events that result in host dehydration and diarrhea (19). From the crystal structure of the binding complex between CT B₅ and GM1, it is known that two terminal sugars of the GM1 pentasaccharide, galactose and sialic acid, are mainly responsible for the interaction with the binding sites of CT B₅, which are approximately 35 Å apart (19). The ability of the CT B₅ to bind to simple mono- and oligo-saccharides with reasonably high

affinity, coupled with its well defined structure, makes this toxin a desirable model system in the study of structure-based design of glycomolecule inhibitors, although most inhibitors of the AB₅ toxins investigated to date have been oligovalent small molecules and bivalent linear inhibitors rather than glycopolymers (10, 11, 20, 21).

In the work reported here, two families of random coil glycopolypeptides have been produced – [(AG)₃PEG]_x (x=10 and 16) and {[(AG)₂PSG]₂ [(AG)₂PEG] [(AG)₂PSG]₂}₆ – as well as two families of α-helical polypeptides with the sequences [AAAEAAAAQAAAQAEAAQAAQ]_y (y=3 and 6), and [AAAQAAQAQAAAE AAAQAAQAAQ]₆ (22-24). Their synthesis has permitted comparison of the role of backbone conformation and saccharide spacing on multivalent binding events. The carboxylic acid functionality of glutamic acid, incorporated at desired positions in the polypeptides, allows the coupling of various amine functionalized saccharides. The ability of the resulting glycopolypeptides to act as inhibitors of the CT B₅ subunit has been compared to determine the impact of specific architectural variables on the binding event. Galactose-functionalized poly(glutamic acid) (PGA) based glycopolymers have also been produced as controls to permit assessments of the effects of backbone flexibility. Direct enzyme-linked assays (DELA) have been conducted to study the inhibition of CT B₅ by the various glycopolypeptides, and the enhancement of IC₅₀ values over the monovalent saccharide galactose has been compared.

Experimental Section

Materials and Methods

Plasmids (pET3b) encoding the [(AG)₃PEG]_x (x=10 and 16) polypeptides were donated by Sakata and Tirrell (25). Other cloning and expression plasmids employed in these studies were purchased from Novagen (San Diego, CA) and EMD Biosciences (San Diego, CA). The general molecular biology methods, protein engineering protocols, and characterization methods employed were previously described (23, 24, 26). All polypeptides were expressed with deca-histidine tags at the N-terminus to facilitate purification via metal chelate affinity chromatography under denaturing conditions. The purity, molecular weight, and composition of the random coil and α-helical polypeptides were confirmed via SDS-PAGE (sodium dodecyl sulfate-polyacrylamide gel electrophoresis), reverse-phase HPLC (RP-HPLC), MALDI-TOF mass spectrometry, ¹H NMR spectroscopy, and amino acid analysis.

Synthesis of Glycopolypeptides

The polypeptides were glycosylated via reaction of the pendant glutamic acids (and C-termini) with either β -D-galactosylamine or *N*-(ϵ -aminocaproyl)- β -D-galactosylamine. The detailed procedure for modification of the polypeptides and purification of the glycopolypeptides has been previously reported (26).

Characterization of Glycopolypeptides

SDS-PAGE Analysis

The successful glycosylation of the polypeptides was monitored via SDS-PAGE with visualization via coomassie blue and periodate staining. In periodate staining, the carbohydrates were oxidized to aldehydes via treatment with periodic acid, followed by treatment with Gel-Code glycoprotein staining solution (Pierce, Rockford, IL).

Degree of Glycosylation

The degree of glycosylation of the glycopolypeptides was estimated via ^1H NMR, MALDI-TOF, and periodate assay (Glycoprotein Carbohydrate Estimation Kit (Pierce, Rockford, IL)). ^1H NMR spectra were acquired on a Bruker DRX-400 NMR spectrometer under standard quantitative conditions at 25 °C, and the standard protocols for the periodate assay were as described by the manufacturer. Comparison of sample solutions to a calibration curve of proteins of known carbohydrate content, combined with quantitative amino acid analysis of the samples, permitted estimation of the degree of substitution of the glycopolypeptides.

Circular Dichroic Spectroscopy (CD)

The random coil and helical conformations of polypeptides and glycopolypeptides were characterized via circular dichroic spectra recorded on a JASCO 810 spectrophotometer (Jasco, Inc., Easton, MD) using a 1-mm path length quartz cuvette in the single-cell mount setup. Background scans of buffer (10 mM phosphate, 150mM NaCl (PBS) pH 7.3) were recorded and automatically subtracted from the sample scans. Samples were made with the

appropriate buffer at a concentration range of approximately 0.1-0.3 mg/ml in PBS buffer pH 7.3. Data points for the wavelength-dependent CD spectra were recorded with 1 nm bandwidth. Mean residue ellipticity was calculated according to standard equations, using the cell pathlength and polypeptide concentrations determined via amino acid analysis.

Gel Permeation Chromatography

The relative hydrodynamic volume of polypeptides and glycopolypeptides was assessed via gel permeation chromatography (GPC). Samples were dissolved at 1mg/mL in phosphate-buffered saline pH 7.3 and filtered through a 0.22 μm filter prior to injection. The samples were separated using a Waters Ultrahydrogel Linear column (7.8 x 300 mm) followed by a Ultrahydrogel 250 (7.8 x 300 mm) column. Detection was achieved via the use of a Waters 2996 photodiode array detector. Additional experiments were conducted for polypeptides dissolved in 0.1M NaNO₃ or in buffer (PBS or NaNO₃) containing 20% acetonitrile, to confirm that neither electrostatic interactions nor hydrophobic interactions were prominent in the observed trends in retention. Identical trends in elution time with polymer identity were observed under all conditions, confirming the lack of interference from slight enthalpic effects during the GPC separation; results for the samples in PBS are given below.

Direct Enzyme Linked Assay (DELA)

DELA assays were carried out in a 96-well format using a GD1b-modified, toxin-binding surface, as previously reported (26). Samples consisted of 6 ng/mL cholera toxin B subunit conjugated to horseradish peroxidase (CT B₅-HRP) incubated at r.t. for 2 hrs in the presence of glycopolypeptide ligands at varying concentrations. IC₅₀ values were calculated from at least five different concentrations of competitive ligand via nonlinear regression, with the statistical package Microcal Origin, as described previously (23, 26, 27).

Results

Synthesis of Glycopolypeptides

Well defined random coil and helical polypeptides were synthesized in good yield via standard protein engineering protocols (23, 24) and modified with β -D-

galactosylamine and *N*-(ϵ -aminocaproyl)- β -D-galactosylamine, via HBTU-activated amide bond formation between the amine groups on the desired saccharides and the glutamic acid side chains of the polypeptides. Glycosylation of polypeptides was confirmed via SDS-PAGE analysis and the degree of glycosylation was determined via at least two of the following methods: MALDI-TOF mass spectroscopy, spectrophotometric analysis, or ^1H NMR spectroscopy. Table 1 shows a list of polypeptides and glycopolypeptides, with their architectural features and degree of glycosylation as determined experimentally. NMR spectra of these glycopolypeptides are similar to that of our previously reported glycopolypeptides (23, 24, 26); the degree of glycosylation was estimated through integration of galactose ring protons with respect to the β protons of alanine. Although some degrees of glycosylation are less than 100%, on average even the incompletely substituted glycopolypeptides (ca. 85%; 72%) carry a sufficient number of saccharides (e.g., 13/17 5/7) to yield a well defined average distance between the saccharides.

Secondary Structure of Glycopolypeptides

Circular dichroism experiments were conducted on all polypeptides and glycopolypeptides to confirm that the secondary structure of the polypeptides did not change as a result of glycosylation. Representative CD spectra for random coil and α -helical polypeptides and glycopolypeptides, modified with either Cap or Gal, are shown in Figure 1.

The spectra of the random coil glycopolypeptides show a minimum at 198 nm, characteristic of a random coil conformation (28), while the CD spectra of the helical glycopolypeptides show the characteristic minima at 208 nm and 222 nm (29). Furthermore, the spectra for the α -helical polypeptide and Gal-functionalized glycopolypeptide are similar, with a mean residue ellipticity of nearly $-40,000 \text{ deg cm}^2 \text{ dm}^{-1}$ and a high fractional helicity of approximately 71% under the experimental conditions.

The conformation of the helical polypeptides changes from helical to non-helical with an increase in temperature, and the glycopolypeptides and polypeptides show similar thermal transition behavior (not shown) (23). CD spectra for all other random coil and α -helical glycopolypeptides were recorded and found to be consistent with our previously reported data (23, 24), with some variations in absolute intensities, regardless of the identity of the saccharides used in the modification (Gal or Cap). These observations confirmed that the glycosylation of polypeptides with different saccharides does not significantly alter the secondary structure of the polypeptides.

Table 1. Polypeptide sequence, molecular weight, approximate distance between glutamic acid residues and degree of substitution for synthesized glycopolypeptides.

<i>Polypeptides and glyco-polypeptides</i>	<i>Polypeptide sequences</i>	<i>Degree of Substitution, (DS) (mol %)</i>	<i>Molecular Mass (kDa)</i>
17-RC-10 ^a	[AGAGAGPEG] ₁₀		7.1
Gal-17-RC-10 ^b		90 ± 1	8.5
Cap-17-RC-10		100 ± 1	9.7
17-RC-16	[AGAGAGPEG] ₁₆		11.1
Gal-17-RC-16		100 ± 1	13.6
Cap-17-RC-16		71 ± 2	14.3
35-RC-6	{[(AG) ₂ PSG] ₄ [(AG) ₂ PEG]} ₆		18.1
Cap-35-RC-6		72 ± 1	19.5
17-H-6	[AAAEQAAAAQAAAEAA AAQAAQ] ₆		14.8
Cap-17-H-6		90 ± 1	17.9
35-H-6	[AAAEQAAQAAAEAA QAAQAAQ] ₆		14.2
Gal 35-H-6		100 ± 1	15.2
Cap-35-H-6		97 ± 7	16.0
Cap-Poly(Glu) ₃₀	Poly(L-Glutamic acid) DP=30	25 ± 1	5.9
Cap-Poly(Glu) ₈₆	Poly(L-Glutamic acid) DP=86	45 ± 4	17.9
Cap-Poly(Glu) ₁₁₃	Poly(L-Glutamic acid) DP=113	11 ± 2	21.8

^a In the notation x-RC-y or x-H-y, x represents the approximate distance between the adjacent glutamic acids or saccharides, H or RC represents helical or random coil conformations of polypeptide in solution, and y represents the multimerization number.

^b Gal = β-D-galactosylamine; Cap = N-(ε-aminocaproyl)-β-D-galactosylamine.

Inhibition Assays

The capability of these glycopolypeptides to act as efficient inhibitors of the CT B₅ was tested in a competitive DELA format developed by Minke et al. (30). Inhibition curves from the DELA experiments for galactose and for select Cap-modified polypeptides are shown in Figure 2. The Gal-modified polypeptides showed poorer inhibition than the Cap-modified polypeptides; indeed, compounds Gal-17-RC-X and Gal-17-H-X did not show any significant inhibition up to a saccharide concentration of 1 mM. IC₅₀ values were calculated via nonlinear regression fits to a standard dose-response curve, employing at least five different concentrations (saccharide concentration) of competitive ligand (23). The reported IC₅₀ value of a ligand is given as the average of the calculated IC₅₀ values from at least two independent assays.

IC₅₀ values calculated for all the synthesized glycopolypeptide inhibitors, reported as saccharide concentrations, are shown in Table 2. Monovalent galactose exhibited an IC₅₀ of approximately 35 mM in these assays, which is consistent with previously reported values (30, 31). The random coil glycopolypeptides – Cap-17-RC-10, Cap-17-RC-16, Cap-35-RC-6, Cap-PGA30, Cap-PGA86 and Cap-PGA113 – show IC₅₀ values of 1150, 1400, 250, 150, 80 and 55 μM, respectively, as shown in Table 2, corresponding to 29, 27, 140, 230, 440, 550-fold improvements over galactose. In comparison, the helical glycopolypeptides show similar ranges of improvement in inhibition. Specifically, Gal 35-H-6 shows a 200-fold improvement in inhibition upon multivalent presentation of saccharide ligands on the helical polypeptide scaffold (23), and the Cap-35-H-6 shows approximately 340-fold improvement. The modified 17-H-X scaffolds showed lower inhibition values (not shown), consistent with observations for the random-coil scaffolds. Differences in the Hill slopes may suggest differences in the valency and/or association state during binding (32); additional characterization via light scattering will assist in the interpretation of these data. Unmodified polypeptides at concentrations of up to 1mM generally showed no inhibition, confirming that binding is mediated by the pendant saccharides and not the polypeptide chain, although polypeptides at concentrations greater than 1 mM exhibited a small degree of non-specific binding (scattered values, 10-15%). These results clearly indicate that the recombinantly derived glycopolypeptides showed increased inhibitory activity relative to the monovalent galactose with variations related to saccharide placement and not polypeptide composition.

Gel Permeation Chromatography

GPC experiments were conducted to evaluate the effects of charges and backbone flexibility on the inhibition of CT by comparing the relative

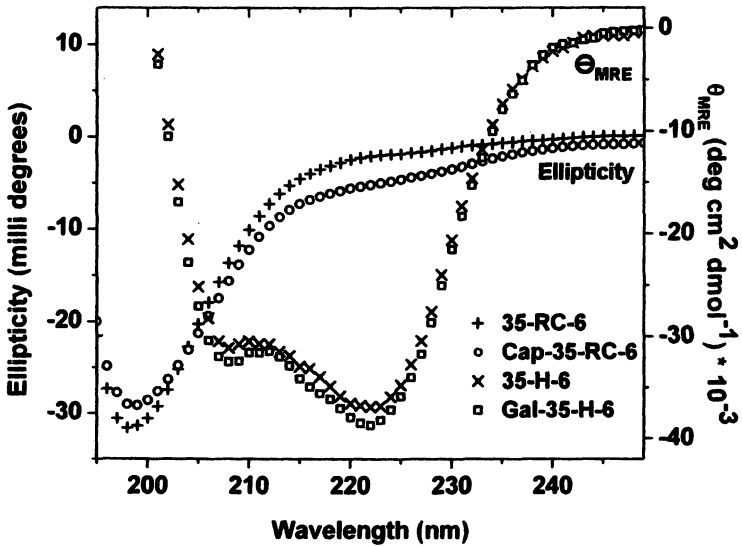


Figure 1. CD spectra of random coil and α -helical polypeptides before and after modification with different saccharides at 4°C in PBS buffer pH 7.3

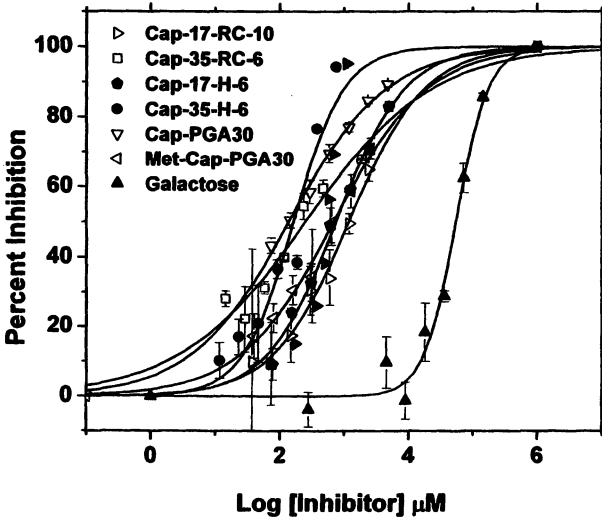


Figure 2. Inhibition of CT B₅ by compounds Cap-17-RC-10, Cap-35-RC-6, Cap-17-H-6, Cap-35-H-6, Cap-PGA30, Met-Cap-PGA30 and Galactose.

hydrodynamic volume (V_h) of polypeptides, glycopolypeptides and Cap-PGAs, as indicated via their retention times in a GPC-based experiment under aqueous conditions (Figure 3). The retention times obtained from these experiments and the corresponding variation in IC_{50} values are shown in Table 2. In these GPC experiments, recombinantly derived random coil and helical glycopolypeptides elute at greater retention times when compared to the PGAs, suggesting the lower hydrodynamic volume of the recombinant polypeptides. Partial modification of PGAs with saccharides had no effect on hydrodynamic volume, due to the retention of charge of the unmodified glutamic acids.

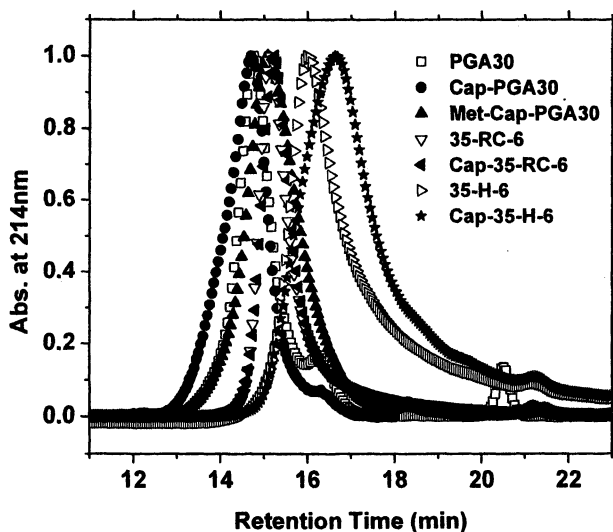


Figure 3. GPC results for PGA30, Cap-PGA30, Met-Cap-PGA30, 35-RC-6, Cap-35-RC-6, 35-H-6 and Cap-35-H-6.

The comparison of the retention times and inhibition improvements, as listed in Table 2, suggests a correlation between measured IC_{50} values and the hydrodynamic volume of the glycopolypeptides. To confirm this relationship between hydrodynamic volume and inhibition efficiency, the charge density of the glycosylated PGA30 was reduced, in order to reduce V_h , via formation of methyl esters of the unglycosylated glutamic acid residues through treatment with trimethylsilyl diazomethane via previously reported methods (33). Methylation of Cap-PGA30 results in an increase in its retention time, confirming the expected reduction in hydrodynamic volume. The inhibition of CT B₅ by the methylated Cap-PGA30 was also reduced, from 230-fold to 50-fold, as shown in Table 2.

Table 2. DELA results for glycopolypeptide and enhancement factor values over monovalent Galactose and retention times from GPC of polypeptides and glycopolypeptides.

<i>Samples</i>	<i>Retention time (min)</i>	<i>IC₅₀ (μM)</i>	<i>Enhancement factor over Galactose</i>
Galactose	-	35000 - 70000 ^a	1
Cap-17-RC-10	-	1150 ± 10	29 ± 19
Cap-17-RC-16	15.2	1400 ± 20	27 ± 10
Cap-35-RC-6	15.2	250 ± 50	140 ± 20
Cap-17-H-6	15.8	725 ± 50	74 ± 3
Cap-35-H-6	16.6	160 ± 26	340 ± 30
Cap-PGA30	14.7	150 ± 20	230 ± 20
Cap-PGA86	14.6	80 ± 5	440 ± 40
Cap-PGA113	-	55 ± 5	550 ± 120
Met-Cap-PGA30 (DM ^b = 80%)	15.1	700 ± 180	50 ± 10

^aValues vary slightly depending on the specific DELA trial.

^bDM = Degree of Methylation

Discussion

Design of Glycopolypeptides

There have been many investigations reported on the use of linear and branched glycopolymers for the inhibition of protein receptors of known structure, which generally show improvements in IC₅₀ values in the range of 10⁰ to 10²-fold relative to monovalent saccharides (16, 34-38). Most of these designs have focused on improvements in binding based on simple multivalency and the identity of saccharides attached to the polymer backbone, but have lacked precise control of the distance between pendant carbohydrate moieties. In these investigations, we produced a series of glycopolypeptides with pendant saccharide spacing intended to match (or not match) the distance between the adjacent binding sites in the CT B₅ subunit (35-37 Å separation). The saccharides present on the polymeric backbones should be accessible to the binding sites of the CT B₅ as the path between the adjacent binding sites is not severely obstructed by amino acid residue side chains (19).

By employing the combination of protein engineering and chemical strategies we have developed random coil glycopolypeptides (Cap-17-RC-10,

Cap-17-RC-16 and Cap-35-RC-6) and corresponding helical glycopolypeptides (Cap-17-H-6 and Cap-35-H-6), which provide a controlled variation in the nominal distance between saccharide groups via the controlled presentation of a chemically reactive amino acid on an otherwise chemically unreactive polypeptide. The Cap-functionalized polypeptides show improved inhibition over all Gal-functionalized polypeptides, confirming the benefit of the hydrophobic linker arm from the Cap-based ligands, as has been observed previously (26, 31, 39). Alanine-rich sequences were selected for the helical glycopolypeptides owing to the very high helical propensity of alanine; the inclusion of glutamine increases aqueous solubility (22, 40). Alanine-containing random-coil polypeptides, similarly equipped with glutamic acid residues, permit assessments of the importance of backbone conformation on the binding event. The combined use of these polypeptides and PGA-based glycopolypeptides has also permitted assessments of the relative binding activity on the basis of valence, charge density, and backbone flexibility.

Effect of Glycosylation

Circular dichroism characterization of the random coil and α -helical glycopolypeptides confirms that glycosylation does not change the conformation of polypeptides, although on average the alanine-rich α -helical glycopolypeptides have lower helicities than their corresponding polypeptides. With increasing temperature, the conformation of helical polypeptides and glycopolypeptides changes to random coil; the ΔH value of this thermal transition is similar for the α -helical polypeptides and glycopolypeptides (not shown), suggesting a similar enthalpy of unfolding for the cooperative units (41). Glycosylation therefore does not appear to disrupt the backbone conformation of the polypeptide, which is consistent with the fact that the glutamic acid residues do not play a key role in stabilizing the secondary structure for either the α -helical or random coil polypeptides. Thus design of this kind of chemically "neutral" scaffold, based on secondary structure preferences of natural amino acids, is promising for the presentation of desired functional groups.

Effects of Saccharide Spacing

The effects of saccharide spacing were compared by changing the nominal distance between the glutamic acids on these random coil and helical polypeptides, which were calculated via root mean square (24) and energy minimization calculation methods (22, 40) respectively. The inhibition exhibited by polypeptides with functional group spacing that is either matched or

unmatched with the spacing between the galactose-binding sites of the CT B₅ subunit (19) were compared. As shown in the DELA data (Table 2 and Figure 2), the recombinant glycopolypeptides Cap-35-RC-6 and Cap-35-H-6, with approximate saccharide spacing of 35 Å, exhibit greater inhibition relative to the glycopolypeptides with saccharides nominally presented at a 17 Å spacing, thus confirming the accepted importance of matching the presentation of the multivalent ligands with those of the target receptor. Although these improvements in binding represent a relatively small difference in free binding energies, the results suggest the potential for varying the avidity of the binding event via controlled and small changes in polymer architecture at a level of structural detail not previously possible. The smaller improvements in inhibition by the glycopolypeptides Cap-17-H-6 and Cap-17-RC-X (X = 10 and 16), which display saccharides at a higher density and number than Cap-35-H-6 and Cap-35-RC-6, likely result from a combination of statistical effects and multivalent binding which are commonly observed in glycopolymeric inhibitors and effector (37, 38, 42). In contrast, the interactions of ligands in Cap-35-RC-6 and Cap-35-H-6 with the receptor are likely due to a primarily multivalent mechanism rather than to statistical effects, owing to the matched saccharide spacing (to CT B₅) and the low valency of saccharides. These results therefore suggest the potential for binding improvements for linear glycopolymers that are mediated by primarily a multivalent mechanism rather than a statistical one.

Effects of Backbone Conformation

Effects of backbone conformation were studied by comparing random coil and α -helical glycopolypeptides with same number of saccharide ligands. The random coil glycopolypeptide Cap-35-RC-6 shows a 140-fold inhibition enhancement over monovalent galactose (24). The helical glycopolypeptide Cap-35-H-6, with the same valency, similar spacing and lower molecular weight, shows slightly higher inhibition (340-fold enhancement). These results suggest that differences in conformational entropy do not have a large impact on inhibition in these systems, as observed in other glycopolymers (38), although the sub-100% fractional helicity of the helical glycopolypeptides may reduce the inherent differences in conformational flexibility.

Effects of Backbone Composition

The inhibition by PGA-based glycopolypeptides was compared with that by recombinantly synthesized glycopolypeptides in order to assess the effects of backbone composition on inhibitory potency. In spite of their similar molecular

mass range, 'average' distance between saccharides, and more heterogeneous structures, the Cap-PGAs exhibit greater inhibition than Cap-35-H-6 and Cap-35-RC-6. The lack in inhibition of CT B₅ by unmodified PGAs at polypeptide concentrations up to 1mM suggests that electrostatic interactions are not responsible for the improvements in inhibition. GPC experiments were therefore conducted to examine the difference in hydrodynamic volume and their effects on toxin inhibition. As shown in Figure 3 for polypeptides eluted with PBS buffer, all polymers elute over a somewhat narrow range of elution times (14-16 minutes), and all PGA polypeptides elute earlier than the recombinant random coil and helical polypeptides under these conditions. Cap-17-RC-16 and Cap-35-RC-6 have similar retention times despite significant difference in their molecular weight, which may be due to the similar percentage of glycine in both sequences, which in the absence of any electrostatic repulsion causes the glycopolypeptides to adopt a compact conformation. The retention times of the PGA polypeptides were roughly dependent on the degree of polymerization of the PGA, and partial glycosylation of the PGA polypeptides had no effect on their retention time. The apparently higher hydrodynamic volume of the PGA polypeptides relative to that of Cap-35-RC-6 and Cap-35-H-6 suggests their more extended conformation in solution; the impact of this chain extension on hydrodynamic volume is highlighted by the earlier elution of even the Cap-PGA30 despite its much lower molecular weight. Taken together with the inhibition results (Table 2), these GPC data illustrate that even these small observed differences in V_h have measurable impact on inhibition properties.

Improvements in inhibition (lower IC_{50} values) due to increases in hydrodynamic volume (lower retention time) are not suggested to result from steric stabilization effects (1, 37, 43), as both DELA and fluorescence titration assays of related glycopolypeptides indicate similar inhibitory potency (not shown)(26). Changes in saccharide accessibility as a function of backbone flexibility is therefore the likely reason for the difference in these inhibition values. To validate the role of V_h , PGA polymers were partially methylated to remove negative charges and their inhibition was tested. Table 2 shows a comparison of the elution profile and inhibition results for the heterogeneous Met-Cap-PGA30 polypeptide relative to that of the recombinant glycopolypeptide; the data illustrate that the methylated PGA-based glycopolypeptide shows a reduced hydrodynamic volume and poorer inhibition after methylation. Previous reports have shown the correlation of polymer molecular mass with inhibition (37, 44, 45); our results indicate that comparisons of molecular size rather than molecular mass are necessary for predicting inhibitory potency of these classes of glycopolypeptides. Thus, appropriate placement of saccharide ligands on the polymer backbone may alone not be sufficient for large inhibition improvements unless the polymer chains can adopt a more extended conformation to increase the accessibility of the ligands to the binding sites.

Despite the relatively low molecular weight of these glycopolypeptides, the enhancements in saccharide binding to the CT B₅ subunit by these glycopolypeptides are of the range observed for large linear synthetic glycopolymers and dendrimers of high valency (10^1 - 10^3 fold improvements)(16, 34-38) and are significantly improved relative to saccharide-modified globular proteins and linear glycopolymers of similar valencies (10^0 - 10^1 fold improvements) (38). Although the inhibition exhibited by these glycopolypeptides is smaller than that observed for high-affinity pentavalent small molecule inhibitors reported by Fan and coworkers (46), our results provide additional guiding principles for designing glycopolymers of greater binding avidity via appropriate manipulation of saccharide placement and chain extension, and suggest opportunities in the design of macromolecules for the manipulation of multivalent phenomena on longer length scales than those accessible by small molecules.

Conclusions

A combination of protein engineering and chemical methods has been successful for the production of a variety of well-defined random coil and α -helical glycopolypeptides. Variations in the number and spacing of the pendant carbohydrate moieties were easily achieved via these methods. Competitive enzyme-linked immunosorbent assays of the recombinantly derived glycopolypeptides suggest that multivalent, well defined glycopolypeptides with saccharides spaced approximately 35 Å apart exhibit greater inhibitory ability toward CT B₅. Inhibition by these glycopolypeptides is comparable to large linear synthetic glycopolymers of high valency. The combination of gel permeation chromatography experiments and immunochemical assays indicates that polypeptides of greater hydrodynamic volume, independent of molecular weight comparisons, exhibit improved inhibition, and that homogeneous glycopolypeptides offer inhibition improvements even when the saccharides have decreased accessibility. The difference in inhibition potency by these well defined glycopolypeptides may allow assessments of the origin of binding improvements based on polymer architectural features, which will serve as the basis for designing new glycopolypeptides of greater binding avidity.

Acknowledgements

This work was funded in part by the Camille and Henry Dreyfus Foundation and the National Science Foundation (DMR 0210223). The project described was also supported in part by grant numbers 1-P20-RR017716-01 and 1-P20-RR015588 (instrument facilities) from the National Center for Research

Resources (NCRR), a component of the National Institutes of Health (NIH). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NCRR or NIH. Wei Zhang and Yong Duan (University of California Davis) are thanked for conducting molecular dynamics simulations, and Jill Sakata and David Tirrell (California Institute of Technology) are thanked for the kind donation of the plasmids encoding polypeptides 1 and 2.

References

1. Mammen, M.; Chio, S.-K.; Whitesides, G. M. *Angew. Chem. Int. Ed.* **1998**, *37*, (20), 2755-2794.
2. Varki, A. *Glycobiology* **1993**, *3*, (2), 97-130.
3. Lee, Y. C. L., R.T. *Acc. Chem. Res.* **1995**, *28*, 321-327.
4. Gordon, E. J.; Sanders, W. J.; Kiessling, L. L. *Nature* **1998**, *392*, 30-31.
5. Nagahori, N.; Nishimura, S.-I. *Biomacromolecules* **2001**, *2*, (1), 22-24.
6. Mammen, M.; Dahmann, G.; Whitesides, G. M. *J. Med. Chem.* **1995**, *38*, 4179-4190.
7. Kamitakahara, H.; Suzuki, T.; Nishigori, N.; Suzuki, Y.; Kanie, O.; Wong, C.-H. *Angew. Chem. Int. Ed.* **1998**, *37*, (11), 1524-1528.
8. Baird, E. J.; Holowka, D.; Coates, G. W.; Baird, B. *Biochemistry* **2003**, *42*, 12739-12748.
9. Lis, H.; Sharon, N. *Chem. Rev.* **1998**, *98*, (2), 637-674.
10. Kitov, P. I.; Shimizu, H.; Homans, S. W.; Bundle, D. R. *J. Am. Chem. Soc.* **2003**, *125*, (11), 3284-3294.
11. Fan, E. K.; Zhang, Z. S.; Minke, W. E.; Hou, Z.; Verlinde, C.; Hol, W. G. J. *J. Am. Chem. Soc.* **2000**, *122*, (11), 2663-2664.
12. Mangold, S. L.; Morgan, J. R.; Strohmeyer, G. C.; M., G. A.; Cloninger, M. J. *Org. Biomol. Chem.* **2005**, *3*(12), 2354-2358.
13. Arosio, D.; Vrasidas, I.; Valentini, P.; Liskamp, R. M. J.; Pieters, R. J.; Bernardi, A. *Org. Biomol. Chem.* **2004**, *2*, (14), 2113-2124.
14. Arosio, D.; Fontanella, M.; Baldini, L.; Mauri, L.; Bernardi, A.; Casnati, A.; Sansone, F.; Ungaro, R. *J. Am. Chem. Soc.* **2005**, *127*, (11), 3660-3661.
15. 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*, (9), 1205-1215.
16. Mangold, S. L.; Cloninger, M. J. *Org. Biomol. Chem.* **2006**, *4*, (12), 2458-2465.
17. Gestwicki, J. E.; Cairo, C. W.; Mann, D. A.; Owen, R. M.; Kiessling, L. L. *Anal. Biochem.* **2002**, *305*, (2), 149-155.
18. Dam, T. K.; Oscarson, S.; Roy, R.; Das, S. K.; Page, D.; Macaluso, F.; Brewer, C. F. *J. Bio. Chem.* **2005**, *280*, 8640-8646.

19. Merritt, E. A.; Sarfaty, S.; Vandenakker, F.; Lhoir, C.; Martial, J. A.; Hol, W. G. J. *Protein Sci.* **1994**, *3*, (2), 166-175.
20. 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*, (6770), 669-672.
21. Fan, E.; O'Neal, C. J.; Mitchell, D. D.; Robien, M. A.; Zhang, Z.; Pickens, J. C.; Tan, X.-J.; Korotkov, K.; Roach, C.; Krumm, B.; Verlinde, C. L. M. J.; Merritt, E. A.; Hol, W. G. J. *Int. J. Med. Microbiol.* **2004**, *294*, (4), 217.
22. Farmer, R. S.; Argust, L. M.; Sharp, J. D.; Kiick, K. L. *Macromolecules* **2006**, *39*, 162-170.
23. Wang, Y.; Kiick, K. L. *J. Am. Chem. Soc.* **2005**, *127*, 16392-16393.
24. Polizzotti, B. D.; Maheshwari, R.; Vinkenburg, J.; Kiick, K. L. *in revision* **2007**.
25. McGrath, K. P.; Fournier, M. J.; Mason, T. L.; Tirrell, D. A. *J. Am. Chem. Soc.* **1992**, *114*, (2), 727-33.
26. Polizzotti, B. D.; Kiick, K. L. *Biomacromolecules* **2006**, *7*, (2), 483-490.
27. Bowen, W. P.; Jerman, J. C. *Trends Pharmac. Sci.* **1995**, *16*, (12), 413-417.
28. Saxena, V. P.; Wetlaufer, D. B. *Proc. Natl. Acad. Sci. U. S. A.* **1971**, *66*, 969.
29. W. Curtis Johnson, J. *Proteins: Structure, Function, and Genetics* **1990**, *7*, (3), 205-214.
30. Minke, W. E.; Roach, C.; Hol, W. G. J.; Verlinde, C. L. M. J. *Biochemistry* **1999**, *38*, (18), 5684-5692.
31. Vrasidas, I.; De Mol, N. J.; Liskamp, R. M. J.; Pieters, R. J. *Eur. J. Org. Chem.* **2001**, (24), 4685-4692.
32. Shoichet, B. K. *J. Med. Chem.* **2006**, *49*, (25), 7274-7277.
33. Strong, L. E.; Kiessling, L. L. *J. Am. Chem. Soc.* **1999**, *121*, 6193-6196.
34. Schlick, K. H.; Udelhoven, R. A.; Strohmeyer, G. C.; Cloninger, M. J. *Mol. Pharmaceutics* **2005**, *2*, (4), 295-301.
35. Appeldoorn, C. C. M.; Joosten, J. A. F.; el Maate, F. A.; Dobrindt, U.; Hacker, J.; Liskamp, R. M. J.; Khan, A. S.; Pieters, R. J. *Tetrahedron-Asymmetry* **2005**, *16*, (2), 361-372.
36. Yamaguchi, S.; Nishida, Y.; Sasaki, K.; Kambara, M.; Kim, C. L.; Ishiguro, N.; Nagatsuka, T.; Uzawa, H.; Horiuchi, M. *Biochem. Biophys. Res. Commun.* **2006**, *349*, (2), 485-491.
37. Gestwicki, J. E.; Cairo, C. W.; Strong, L. E.; Oetjen, K. A.; Kiessling, L. L. *J. Am. Chem. Soc.* **2002**, *124*, (50), 14922-14933.
38. Lundquist, J. J.; Toone, E. J. *Chem. Rev.* **2002**, *102*, (2), 555-578.
39. Minke, W. E.; Hong, F.; Verlinde, C. L. M. J.; Hol, W. G. J.; Fan, E. J. *Biol. Chem.* **1999**, *274*, (47), 33469-33473.
40. Farmer, R. S.; Kiick, K. L. *Biomacromolecules* **2005**, *6*, 1531-1539.
41. MacCallum, J. L.; Moghaddam, M. S.; Chan, H. S.; Tieleman, D. P. *PNAS* **2007**, *104*, (15), 6206-6210.

42. Woller, E. K.; Walter, E. D.; Morgan, J. R.; Singel, D. J.; Cloninger, M. J. *J. Am. Chem. Soc.* **2003**, *125*, 8820-8826.
43. Choi, S. K.; Mammen, M.; Whitesides, G. M. *Chem. Biol.* **1996**, *3*, 97-104.
44. Spaltenstein, A.; Whitesides, G. M. *J. Am. Chem. Soc.* **1991**, *113*, 686-687.
45. Totani, K.; Kubota, T.; Kuroda, T.; Murata, T.; Hidari, K. I. P. J.; Suzuki, T.; Suzuki, Y.; Kobayashi, K.; Ashida, H.; Yamamoto, K.; Usui, T. *Glycobiology* **2003**, *13*, 315-326.
46. Zhang, Z.; Pickens, J. C.; Hol, W. G. J.; Fan, E. *Org. Lett.* **2004**, *6*, (9), 1377-1380.

Author Index

- Agard, Nicholas J., 205
An, Hyun Joo, 195
Chen, Chi-Chang, 175
Chen, Wenlan, 135
Chen, Xi, 79
Chokhawala, Harshal A., 79
De Libero, Gennaro, 135
Gervay-Hague, Jacquelyn, 121
Guthrie, Katherine A., 157
Hesek, Dusan, 40
Holmberg, Leona A., 157
Huang, Fei, 175
Huang, Lijun, 23
Huang, Shengshu, 79
Huang, Xuefei, 23
Kiick, Kristi L., 232
Kulkarni, Suvarn S., 121
Lebrilla, Carlito B., 195
Lee, Mijoon, 40
Linhardt, Robert J., 175
Liu, Shuang, 232
Liu, Yang, 135
Lu, Xiaowei, 23
Maheshwari, Ronak, 232
Mobashery, Shahriar, 40
O'Doherty, George, 3
Pohl, Nicola L., 221
Polizzotti, Brian D., 232
Sandmaier, Brenda M., 157
Song, Jing, 135
Tillinghast, John, 195
Tokuzaki, Kazuo, 175
Tomiyama, Hiroshi, 175
Wang, Lai-Xi, 99
Wang, Peng George, 135
Wang, Ying, 232
Weïwer, Michel, 175
Witczak, Zbigniew J., 62
Xia, Chengfeng, 135
Yao, Qingjia, 135
Yu, Hai, 79
Yu, Xiaomei, 3
Yuan, Xuejun, 175
Zhang, Wenpeng, 135
Zhang, Yalong, 135
Zhou, Dapeng, 135

Subject Index

- A**
- Achmatowicz approach, hexopyranoses, 6–7, 10
 - Agelasphins
 - glycolipids, 154, 168, 170
 - See also* Glycolipid antigens
 - Alcohols
 - asymmetric synthesis of chiral furfuryl, 11
 - glycosylation of, 14, 15, 16
 - Anticancer vaccine candidates
 - carbohydrate antigen sialyl Tn (STn), 218
 - C-linked polysialic acid (PSA) target, 235*f*
 - derivatives of neuraminic acid (Neu5Ac, KDN, and KDO), 217
 - glycoprotein glycan catabolism, 220*f*
 - glycosylation of sialic acid type carbohydrates, 218, 221
 - immunological evaluation of O- and C-glycosides of Tn and STn, 226, 234*t*
 - keyhole limpet hemocyanin (KLH) conjugates of STn O- and C-glycosides, 226, 230*f*
 - linkage of neuraminic acid to glycoconjugates, 217–218
 - O-glycosides vs. C-glycosides stability, 219*f*
 - retrosynthetic scheme for PSA C-disaccharide, 228, 231
 - samarium mediated C-glycosylations, 218, 222
 - synthesis of C-glycoside acceptor, 223, 227
 - synthesis of C-glycoside analog of STn, 221, 223, 224, 225
 - synthesis of C-Neu5Ac disaccharide, 228, 232, 233
 - synthesis of double C-glycoside analog of STn, 223, 226, 227
 - synthesis of polysialic acid (PSA) C-disaccharide, 228, 234
 - tumor antigen STn O-glycoside and its C-glycoside analog, 218, 221*f*
 - Anti-tumor activity, cardiac glycosides, 20, 21*f*
 - Assays, high-throughput substrate specificity, 112, 116
 - Asymmetric synthesis, chiral furfuryl alcohols, 11
 - Autologous peripheral blood stem cell transplantation (ASCT)
 - breast and ovarian cancer treatment, 198–199
 - immune reconstitution after, 211–212
 - See also* Theratope (sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine
 - Automated oligosaccharide synthesis
 - basic strategy for phase-switching approaches, 274*f*
 - challenges, 273–279
 - solid-phase approaches, 273, 275
 - solution-phase chemistry to carbohydrate structures, 275, 277
 - solution-phase strategy, 277, 279
 - synthesis of polymannose, 278*f*
 - synthesis of polyrhannose, 280*f*
 - See also* Oligosaccharides
 - 3-Azidopropyl β -D-galactopyranose-(1 \rightarrow 4)- β -D-glucopyranoside, acceptor for α 2,6-linked sialosides, 102, 104*t*, 105*t*

B

- Backbone composition,
glycopolypeptides, 298–300
- Backbone conformation,
glycopolypeptides, 298
- Bacterial cell wall
basic chemical structure, 55
biosynthesis pathway, 55, 56*f*, 57*f*
chemoenzymatic preparation of
lipid II, 71, 74*f*
construction of backbone of lipid
IV tetrasaccharide, 71, 76*f*
construction of disaccharide core of
tetrasaccharide, 60–61, 62*f*
construction of NAG–NAM repeats
(*N*-acetylglucosamine–*N*-
acetylmuramic acid), 58, 59
hexadecasaccharide by
glycosylation, 65
immunostimulatory properties of
peptidoglycan and fragments,
61, 64
lactyl ether at 3-O position late in
NAM synthesis, 58, 60
lipid II as key building block, 55
NAG–NAM–NAG–NAM as
methyl lactate, 65, 68*f*
NAG–NAM–NAG–NAM
sequence leading to fragment of,
61, 63*f*
protecting groups, 60
solid-phase synthesis of branched
peptidoglycan derivative, 65–66,
70*f*
studies of synthetic cell wall
fragments, 61
syntheses of lipid II, 69, 73*f*
synthesis of heptaprenyl-lipid IV,
71, 75
synthesis of lipid I, 69, 72*f*
synthesis of NAG–MPP–NAG–
MPP (MPP = muramyl
pentapeptide), 58, 59
synthesis of octasaccharide, 65, 67*f*
trichloroacetimidate method, 60
- Bacterial glycolipids
activators cholesteryl 6-O-acyl- β -
D-galactopyranoside (BbGL1)
and 1,2-di-O-acyl-3-O- α -D-
galactopyranosyl-*sn*-glycerol
(BbGL2), 156
Natural Killer T (NKT) cells
activators, 155–156
synthesis of C-glycoside analog of
BbGL2, 160, 163
See also Glycolipid antigens
- Biological activities, hyaluronic acid
oligosaccharides (sHA), 30–31
- Biological processes, carbohydrates,
268–269
- Bioorthogonal chemical reporter
strategy
applications of metabolic
engineering, 266–268
bioorthogonal reactions on sugars,
261*f*
biosynthesis of CMP-sialic acid,
254, 256*f*
click chemistry, 264, 267
copper-catalyzed modification of
Huisgen-type alkyne-azide
[3+2] cycloaddition, 262, 264
copper-free variant of [3+2]
cycloaddition, 264
covalent tagging, 260, 262, 264–
266
glycoconjugates, 252–253
literature precedents, 259*f*
metabolic labeling of unnatural
monosaccharides, 253–254, 257,
260
metabolizing monosaccharides via
salvage pathways, 255*f*
salvage pathways, 254, 255*f*
Staudinger ligation, 262, 262*f*
thiol-maleimide conjugation, 262
tolerance of pathways, 254, 257,
260
unnatural sugars, 258

- Biosynthesis pathway, bacterial cell wall, 53, 54*f*, 55*f*
- Bovine ribonuclease B, glycosylation remodeling of, 139, 143, 144
- Breast cancer
- autologous peripheral blood stem cell transplantation (ASCT) treatment, 198–199
 - event free survival, 203, 204*f*
 - event free survival by estrogen receptor status and treatment, 203, 206*f*
 - factors associated with event free survival in vaccinated patients, 203, 209
 - failure counts and percentages for event free survival, 210*t*
 - immune function testing, 203, 208*t*
 - multivariable hazard ratios for overall survival, 211*t*
 - overall survival by estrogen receptor status and treatment, 203, 207*f*
 - overall survival by relapse risk and treatment, 203, 205*f*
 - patient characteristics, 200*t*
 - vaccination outcome with Theratope, 212
- See also* Ovarian cancer; Theratope (sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine
- C**
- Cancer. *See* Anticancer vaccine candidates; Breast cancer; Ovarian cancer; Theratope (sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine
- Carbohydrates
- Achmatowicz approach, 6–7, 10
 - aminoxy fluorouracil tag for incorporation of reducing sugars, 284*f*
 - design and developing new glycochemicals, 79–80
 - direct formation of fluorouracil-based carbohydrate microarrays, 279, 283
 - fluorouracil-tagged, for microarray formation, 282*f*
 - mediating biological processes, 272–273
 - strategy for use of single fluorouracil tag, 281*f*
- See also* Glycosyl thio-carboamino peptides
- Cardiac glycosides
- D- and L-, with anti-tumor activity, 21*f*
 - de novo approach to, 20, 22–24
 - digitoxin, 24, 26
 - digitoxin disaccharide, 23, 25
 - digitoxin monosaccharide, 23, 24
 - natural products to aglycons to new bioactive leads, 20, 22
- Cell surface presentation, modified sialic acids, 95
- Cell wall. *See* Bacterial cell wall
- Chemoenzymatic synthesis
- advantage, 140
 - extension to glycosylation remodeling of glycoproteins, 140–141
 - isoglobotrihexosylceramide (iGb3), 181, 182
- See also* Glycopeptides; Sialosides
- Cholera toxin
- study of structure-based design, 289–290
- See also* Polypeptide-based glycopolymers
- Click chemistry
- azides and alkynes, 264

- dynamic imaging of glycans, 267
- sialosides, 104, 111
- Cloning, *Escherichia coli* K-12 sialic acid aldolase, 100
- Cluster glycoside effect, multivalent binding, 287
- Conjugation
 - functionalized sialosides to biomolecules, 104, 108, 109
 - thiol-maleimide, 262
- Convergent chemoenzymatic synthesis. *See* Glycopeptides
- Covalent tagging, chemical reporters, 256, 258, 260–262

D

De novo synthesis

- Achmatowicz approach, 6–7, 10
- β -linked *N*-glycosides, 17, 18
- cardiac glycosides, 20, 22–24
- D- and L-cardiac glycosides with anti-tumor activity, 20, 21*f*
- digitoxin, 24, 26
- digitoxin disaccharide, 23, 25
- digitoxin monosaccharide, 23, 24
- enantio- and diastereoselective pyranone synthesis, 14, 16
- enantioselective synthesis of hexoses, 4
- furfuryl alcohols, 7, 11
- galactono-lactones, 5–6
- galacto-sugar γ -lactones from dienates, 5, 8, 9
- glycosylation of alcohols, 14, 15, 16
- 1,4-linked manno-trisaccharides, 20, 21
- 1,6-manno-trisaccharides, 18, 19, 20
- monosaccharide natural products, 17*f*, 18

- monosaccharides, 4–7
- natural and unusual
 - oligosaccharides, 14, 18
- natural products to aglycons to new bioactive leads, 20, 22
- oligosaccharides, 18–24
- 1,4-oligosaccharides, 20, 21
- 1,6-oligosugars, 18, 19, 20
- palladium-catalyzed α/β -selective glycosylation, 14, 15
- palladium π -allyl chemistry for pyranoses, 6, 8, 9
- papulacandins, 7, 12, 13
- term, 4

Digitoxin, synthesis, 24, 26

Digitoxin disaccharide, synthesis, 23, 25

Digitoxin monosaccharide, synthesis, 23, 24

Direct enzyme linked assay (DELA), glycopolypeptides, 290, 296*t*

Double glycosylation, endoglycosidase-catalyzed, 136, 138

E

Efficiency, chemoenzymatic synthesis of sialosides, 99

Endo- β -*N*-acetylglucosaminidases (ENGases)

substrate-assisted mechanism, 129*f*

synthesis of di- and tetra-saccharide sugar oxazolines, 129, 130

transglycosylation for glycopeptide synthesis, 124–125

transglycosylation with synthetic sugar oxazolines, 133

See also Glycopeptides

Escherichia coli K-12 sialic acid aldolase, cloning and characterization, 100

- F**
- Fluorous-based carbohydrate microarrays
 - aminoxy fluorous tag for incorporation into, 284*f*
 - direct formation, 279, 283
 - fluorescence images, 283*f*
 - fluorous-tagged, for microarray formation, 282*f*
 - mediating biological processes, 272–273
 - strategy for use of single fluorous tag, 281*f*
 - Furfuryl alcohols, asymmetric synthesis, 11
- G**
- Galactolipid KRN 7000
 - discovery, 154
 - synthesis of, analogs, 158, 161
 - synthesis using armed and disarmed concept, 172–173, 174
 - See also* Glycolipid antigens
 - Galactono-lactones, de novo synthesis, 5–6
 - Galactosyl ceramides (GalCer)
 - discovery of glycolipids, 154
 - first generation synthesis of, analogs, 158, 159
 - immunogenic analogs, 155*f*
 - modification on 3'- and 4'-positions, 172–175, 179
 - stimulation of Natural Killer T (NKT) cells by, analogs, 175, 179
 - streamlined synthesis, 158, 160, 161
 - synthesis of 3'-substituted analogs, 173, 176
 - synthesis of 4'-Nac-GalCer, 175, 177
 - tumor suppressing properties, 170
 - See also* Glycolipid antigens
 - Gel permeation chromatography (GPC), glycopolypeptides, 290, 293, 295
 - Glucose, building block for hyaluronic acid oligosaccharides, 31, 32
 - Glucoside, conversion into glucuronic acid for hyaluronic acid oligosaccharides, 38
 - Glucuronic acid
 - direct hyaluronic acid oligosaccharide (sHA) assembly, 38
 - glycosyl donor for hyaluronic acid oligosaccharides, 31, 32
 - Glycans
 - function, 252
 - incorporating functionalized monosaccharides into, 259*f*
 - profiling and imaging, 267–268
 - See also* Bioorthogonal chemical reporter strategy
 - Glycobiology tools. *See* Glycosyl thio-carboamino peptides
 - Glycochemicals, design and development from carbohydrates, 77–78
 - Glycoconjugates
 - applications of metabolic engineering, 266–268
 - bioorthogonal chemical reporter strategy, 252–253
 - bioorthogonal reactions on sugars, 261*f*
 - biosynthesis of CMP-sialic acid, 256*f*
 - click chemistry, 264, 267
 - copper-catalyzed modification of Huisgen-type alkyne-azide [3+2] cycloaddition, 262, 264
 - copper-free variant of [3+2] cycloaddition, 264
 - covalent tagging of chemical reporters using bioorthogonal

- chemistries, 260, 262, 264–266
- cyclooctyne reagents and relative second-order rate constants for reaction with benzyl azide, 265*f*
- detection and purification, 252–253
- direct imaging of glycosylation, 267–268
- literature precedents for incorporating functionalized monosaccharides into glycans, 259*f*
- metabolic labeling of unnatural monosaccharides, 253–254, 257, 260
- monosaccharides metabolized via salvage pathways, 255*f*
- optimal choice of azide-specific bioorthogonal chemistries, 264–266
- role of glycosylation in human health and disease, 267
- Staudinger ligation, 262, 263*f*, 266
- thiol-maleimide conjugation, 262
- tolerance of pathways, 254, 257, 260
- unnatural sugars, 258
- Glycolipid antigens
 - activators cholesteryl 6-O-acyl- β -D-galactopyranoside (BbGL1) and 1,2-di-O-acyl-3-O- α -D-galactopyranosyl-s-glycerol (BbGL2), 156
 - agelasphins, 154, 168, 170
 - α -anomeric glycolipids, 160, 162*f*
 - background of synthesis, 157
 - bacterial glycolipid ligands for NKT cell activation, 155–156
 - CD1d dependent mode of activation of Natural Killer T (NKT) cells, 154–155
 - discovery of immunopotent α -linked galactolipid KRN 7000, 154
 - discovery of lysosomal isoglobotrihexosylceramide (iGb3), 170*f*, 171
 - iGb3 by chemoenzymatic synthesis, 181, 182
 - immunogenic galactosyl ceramides (GalCer) analogs, 155*f*
 - in situ anomerization, 157, 159
 - mammalian cells, 171–172
 - measurement of binding with NKT cells, 175, 177*f*
 - metabolic stability of iGb3 analogs, 187–190
 - modification of α -GalCer on 3'- and 4'-positions, 172–175, 179
 - NKT cells activation by iGb3 analogs, 183, 185*f*
 - one-pot synthesis of α -anomeric glycolipids, 160, 162
 - preparation of ceramide acceptors, 158, 159
 - presentation to T lymphocytes, 168, 169*f*
 - α -stereoselective glycosylation, 158, 159
 - α -stereoselective glycosylation via 'in situ anomerization', 157, 159
 - stimulation of hybridomas NKT cells by α -GalCer analogs, 175, 178*f*
 - streamlined synthesis of α -GalCer and BbGL2 analogs, 158, 160, 161
 - structure activity relationships of iGb3, 182–183
 - structure of iGb3, 170*f*
 - structures of exogenous and endogenous, of NKT cells, 170*f*
 - synthesis of 3'-substituted α -GalCer analogs, 173, 176
 - synthesis of 4'-NAc-GalCer, 175, 177

- synthesis of α -linked glycolipids, 157–160
- synthesis of C-glycoside analog of BbGL2, 160, 163
- synthesis of iGb3 by chemo method, 179, 180
- synthesis of KRN 7000 analogs, 156, 161
- synthesis of KRN 7000 using armed and disarmed concept, 172–173, 174
- tumor suppressing properties of α -GalCer, 170
- Glycopeptides**
 - advantage of chemoenzymatic method, 140
 - endo-A-catalyzed synthesis of modified glycopeptide C34, 136
 - endo-A-catalyzed transglycosylation with different oxazolines, 135
 - endo- β -*N*-acetylglucosaminidase (ENGase)-catalyzed transglycosylation for, 124–125
 - endo-M-catalyzed transglycosylation with synthetic oxazolines, 139
 - ENGase-catalyzed double glycosylation, 138, 140
 - ENGase-catalyzed hydrolysis and transglycosylation, 126*f*
 - ENGase-catalyzed transglycosylation with synthetic sugar oxazolines, 132
 - extension of chemoenzymatic approach to glycosylation remodeling of glycoproteins, 140–141
 - glycosylation remodeling of bovine ribonuclease B, 141, 145, 146
 - modified sugar oxazoline as substrates, 132, 134, 138
 - structures of modified oligosaccharide oxazoline derivatives, 134*f*
 - structures of synthetic CD52 antigen with natural glycans, 127*f*
 - structures of synthetic sugar oxazoline derivatives, 137*f*
 - substrate-assisted mechanism for ENGase-catalyzed reactions, 129*f*
 - synthesis of cyclic HIV-1 V3, carrying two *N*-glycans, 142, 143
 - synthesis of di- and tetra-saccharide sugar oxazolines, 129, 130
 - synthesis of unnatural hexasaccharide oxazoline, 141, 144
 - transition-state analog substrates for ENGase-catalyzed transglycosylation, 125, 128, 132
- See also* Polypeptide-based glycopolymers
- Glycopolypeptides**
 - backbone composition, 300–302
 - backbone conformation, 300
 - characterization methods, 291–292
 - circular dichroism spectroscopy, 291–292, 296*f*
 - degree of glycosylation, 291, 294*t*
 - design of, 298–299
 - direct enzyme linked assay (DELA), 292, 298*t*
 - gel permeation chromatography, 292, 295, 297
 - inhibition assays, 295, 296*f*
 - SDS-PAGE analysis, 291
 - secondary structure, 293, 296*f*
 - synthesis, 291, 292–293
- Glycoproteins**
 - glycosylation remodeling of, 140–141

- See also* Glycopeptides;
Glycosylation
- Glycosaminoglycan (GAG)
superfamily, hyaluronic acid (HA),
30
- N*-Glycosides, β -linked, synthesis, 17,
18
- Glycosphingolipids
chemo and chemoenzymatic
synthesis of
isoglobotrihexosylceramide
(iGb3), 179, 182
metabolic stability of iGb3 analogs,
187–190
modification of α -
galactosylceramide (α -GalCer)
on 3'- and 4'-positions, 172–175,
179
NKT cells activation by iGb3
analogs, 183, 185*f*
structure activity relationship
studies of iGb3, 182–183
synthesis of 3'-substituted α -
GalCer analogs, 173, 176
synthesis of 4'-NAc-GalCer, 175,
177
synthesis of KRN 7000 using
armed and disarmed concept,
172–173, 174
synthesis of S-iGb3 through
reinversion protocol, 187, 188,
189
See also Glycolipid antigens
- Glycosylation
alcohols, 14, 15, 16
alternative method for
determination of *N*-linked sites
and oligosaccharide
heterogeneity, 242–243
computer program GlycoX, 241,
246–248
determination of, sites, 244–246
direct imaging of, 267–268
distribution of glycans on *N*-linked,
sites, 247*f*
ENCase-catalyzed double, 138,
140
experimental strategy of pronase
digestion, 243*f*
glycopeptides from pronase
digestion of ribonuclease B,
245*f*
glycopolypeptides, 299
glycoproteomics tool: GlycoX,
237, 246–248
N-linked, 242
oligosaccharides on ^{60}N of
ribonuclease B and relative
abundances, 246*f*
O-linked, 242
palladium-catalyzed α/β -selective,
14, 15
post-translational modification of
proteins, 123–124, 252
pronase digestion, 243–245
remodeling of glycoproteins, 140–
141
role in human health and disease,
267
schematic flowchart of GlycoX,
248*f*
sialic acid type carbohydrates, 218,
221, 222
site-specific, analysis, 242
stereoselective, for glycolipids,
157, 159
sulfoxide and trichloroacetimidate,
for hyaluronic acid
trisaccharides, 38, 42, 43
See also Glycopeptides
- Glycosyl thio-carboamino peptides
coupling of amino functionalized
thiodisaccharide with protected
amino acids, 87, 89
4,6-dideoxy-1,2-*O*-isopropylidene-
D-glycero-pent-4-enopyranose-
3-*ulose* (3), 80
divergent syntheses constructing S-
thio-carbopeptide libraries, 88,
89, 90

- enone (3) functionalized for
 formation of (1-3)-S-thio-3-nitro-3,4-dideoxysaccharides, 86-87
- formation of β -(1-5)-4-deoxy-5-C-thiodisaccharide, 85-86
- isolevoglucosenone, 80
- levoglucosenone, 80
- Michael addition of 1-thio- β -D-glucose to isolevoglucosenone for (1-2)-S-thiodisaccharides, 83-84
- Michael addition of 1-thioglucose to α -nitroalkene, 82-83
- new approach to β -(1-2)-2,3-dideoxy-2-C-acetamidomethyl-2-S-thiodisaccharides, 81-82
- new family of C-nitro-S-thio-addition products, 86-87
- new perspectives, 92
- one-step synthesis of (1,4)-S-thiodisaccharides, 81
- preparation methods of S-linked and C-S-thiodisaccharides, 80
- reactive enones class 4-deoxy-1,2-O-isopropylidene-L-glyceropent-4-enopyrano-3-ulose, 85
- Sn2 displacement of iodine in 3-iodo-levoglucosenone for formation of (1-3)-S-thiodisaccharides, 83-84
- three step approach involving Michael addition of methanol, nitromethane addition, and mesylation/elimination, 82-83
- GlycoX**
 computer program, 241, 246-248
 schematic flowchart, 248f
See also Glycosylation
- H**
- Heptaprenyl-lipid IV synthesis, 71, 75
- See also* Bacterial cell wall
- Hexasaccharide oxazoline, synthesis of unnatural, 139, 142
- Hexopyranoses, Achmatowicz approach, 6-7, 10
- High-throughput substrate specificity, sialidases, 112-113
- Huisgen-type alkyne-azide [3+2] cycloaddition
 copper-catalyzed, 262, 264
 copper-free variant, 264
See also Click chemistry
- Hyaluronic acid oligosaccharides (sHA)
 biological activities, 30-31
 comparison to hyaluronic acid (HA), 30
 conventional synthesis, 31-39
 conversion of glucoside into glucuronic acid for sHA building blocks, 38
 core sequences with odd number of monosaccharide units, 45, 47i
 deprotection order and oxidation strategy for one-pot, 46, 50
 disaccharides in Jacquinet synthesis, 38-39, 42, 43
 factors for designing synthetic route, 31
 gluconuric acid directly as glycosyl donor, 31, 32
 glucose as building block with post glycosylation, 31-33
 glucuronic acid, glucose and glucosamine derivatives for one-pot method, 39, 45
 glucuronic acid building blocks for sHA assembly, 38
 iterative one-pot synthesis, 39, 45-46, 49
 multiple one-pot glycosylations, 39, 44
 scale-up of one-pot synthesis, 45-46, 49

- sHA pentasaccharide and sHA hexasaccharide with glucosamine at reducing end, 33, 35, 36, 37
- sHA tetrasaccharide with glucuronic acid at reducing end, 31, 33, 34, 35
- sHA trisaccharides with combination of sulfoxide and trichloroacetimidate glycosylation, 38, 40, 41
- I**
- Immunostimulatory properties, peptidoglycan and fragments, 59, 62
- Inhibition assays, glycopolypeptides, 293, 294*f*
- Isoglobotrihexosylceramide (iGb3) activation of Natural Killer T (NKT) cells, 183, 185*f*
- chemoenzymatic synthesis, 181, 182
- chemo method, 179, 180
- deoxy galactosyl donors and terminal deoxy analogs, 183, 186
- discovery, 170*f*, 171
- metabolic stability, 187–190
- modification of ceramide part, 183, 184
- structure activity relationship studies, 182–183
- synthesis of S-iGb3 through reinversion protocol, 187, 188, 189
- See also* Glycolipid antigens
- Iterative one-pot synthesis, hyaluronic acid oligosaccharides, 39, 44, 45–46, 49
- J**
- Jacquinet synthesis, hyaluronic acid disaccharides, 38–39, 42, 43
- L**
- Lipid I synthesis, 69, 72*f*
See also Bacterial cell wall
- Lipid II bacterial cell wall, 55
chemoenzymatic preparation, 71, 74*f*
total synthesis of, 69, 73*f*
See also Bacterial cell wall
- Lipid IV construction of backbone of lipid IV tetrasaccharide, 71, 76*f*
synthesis of heptaprenyl-, 71, 75
See also Bacterial cell wall
- M**
- Metabolic engineering, applications, 266–268
- Metabolic labeling, unnatural monosaccharides, 253–254, 257, 260
- Metabolic stability, isoglobotrihexosylceramide (iGb3), 187–190
- Monosaccharide natural products, de novo synthesis, 17*f*, 18
- Monosaccharides Achmatowicz approach, 6–7, 10
de novo synthesis, 4–7
galactono-lactones, 5–6
literature precedents for

- incorporating functionalized, into glycans, 259*f*
- metabolic labeling of unnatural, 253–254, 257, 260
- palladium π -allyl chemistry to pyranoses, 6, 8, 9
- papulacandins, 7, 12, 13
- salvage pathways, 254, 255*f*
- Multivalent binding events
 - cluster glycoside effect, 289
 - general study, 289
 - See also* Polypeptide-based glycopolymers
- N**
- Natural Killer T (NKT) cells
 - activation by
 - isoglobotrihexosylceramide (iGb3) analogs, 183, 185*f*
 - bacterial glycolipid ligands for, activation, 155–156
 - mode of activation, 154–155
 - stimulation by α -galactosylceramide analogs, 175, 179
 - See also* Glycolipid antigens
- O**
- Oligosaccharide oxazolines
 - structures of, derivatives, 134*f*, 137*f*
 - transglycosylation substrates, 132, 134, 138
- Oligosaccharides
 - aminoxy fluorous tag for incorporating reducing sugars, 284*f*
 - automated solution-phase strategy, 277, 279
 - automated solution-phase synthesis of polyrhannose, 280*f*
 - basic strategy for phase-switching approaches, 274*f*
 - cardiac glycosides, 20, 22–24
 - challenges of automated synthesis, 273–279
 - de novo synthesis of natural and unusual, 14, 18
 - digitoxin, 24, 26
 - digitoxin disaccharide, 23, 25
 - digitoxin monosaccharide, 23, 24
 - direct formation of fluorous-based carbohydrate microarrays, 279, 283
 - fluorous-tagged, for microarray formation, 282*f*
 - future directions, 285
 - 1,4-oligosaccharides, 20, 21
 - 1,6-oligosugars, 18, 19, 20
 - solid-phase approaches to building carbohydrate structures, 273, 275
 - solution-phase alternatives to building carbohydrate structures, 275, 277
 - strategy for use of single fluorous tag, 281*f*
 - synthesis of polymannose, 278*f*
 - See also* Hyaluronic acid oligosaccharides (sHA)
- One-pot synthesis, hyaluronic acid oligosaccharides, 39, 44, 45–46, 49
- One-pot three-enzyme approach. *See* Sialosides
- Ovarian cancer
 - autologous peripheral blood stem cell transplantation (ASCT) treatment, 198–199
 - event free survival, 203, 204*f*
 - overall survival by relapse risk and treatment, 203, 204*f*
 - patient characteristics, 200*t*
 - See also* Breast cancer; Theratope (sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine

P

- Papulacandins, de novo synthesis, 7, 12, 13
- Pasteurella multocida*
sialyltransferase, α 2,3-linked sialosides, 109
- Peptides. *See* Glycosyl thio-carboamino peptides
- Peptidoglycan
cross-linked strands in cell wall, 55
solid-phase synthesis of branched, 65–66, 70f
- Photobacterium damsela* α 2,6-sialyltransferase, cloning and purification, 100, 102
- Polypeptide-based glycopolymers
characterization of
glycopolypeptides, 291–292
circular dichroism (CD) spectra of random coil and α -helical polypeptides before and after modification with saccharides, 291–292, 296f
design of glycopolypeptides, 298–299
direct enzyme linked assay (DELA), 292, 298f
effect of glycosylation, 299
effects of backbone composition, 300–302
effects of backbone conformation, 300
effects of saccharide spacing, 299–300
experimental, 290–292
gel permeation chromatography (GPC), 292, 295, 297
inhibition assays, 295, 296f
polypeptide sequence, molecular weight, degree of substitution and molecular mass, 294f
secondary structure of
glycopolypeptides, 293, 296f

synthesis of glycopolypeptides, 291, 292–293

Polysialic acid (PSA)

C-linked PSA target, 235f
retrosynthetic scheme, 228, 231
synthesis of PSA C-disaccharide, 228, 234

See also Anticancer vaccine candidates

Post-translational modification. *See* Glycosylation

Pronase digestion

ribonuclease B, 243–244
See also Glycosylation

Protein modification

glycosylations, 123–124
See also Glycopeptides

Pyranoses

enantio- and diastereoselective synthesis, 14, 16
palladium π -allyl chemistry, 6, 8, 9

R

Ribonuclease B

determination of glycosylation sites, 244–246
glycopeptides from pronase digestion of, 245f
oligosaccharides on ^{60}N of, 246f
pronase digestion, 243–244
See also Glycosylation

S

Saccharide spacing,
glycopolypeptides, 297–298

Secondary structure,
glycopolypeptides, 291, 294f

Sialic acids

α -anomeric forms of, 217f
biological roles, 217

- biosynthesis of CMP-sialic acid, 256f
 cell surface presentation of modified, 97
 C-linked polysialic acid (PSA) target, 235f
 derivatives of neuraminic acid (Neu5Ac, KDN, and KDO), 217
 linkage of neuraminic acid to glycoconjugates, 217–218
 mechanism and significance of structural diversity, 97–98
 nonulosonic acids, 97
 synthesis of PSA C-disaccharide, 228, 234
See also Anticancer vaccine candidates; Sialosides
Sialidases, high-throughput substrate specificity, 112–113
Sialosides
 α 2,6-linked, using 3-azidopropyl β -D-galactopyranose-(1 \rightarrow 4)- β -D-glucopyranoside (Lac β ProN₃) as acceptor, 104, 106t, 107t
 α 2,6-linked, using Gal β OMe as acceptor, 105, 110
 bacterial enzyme *Pasteurella multocida* sialyltransferase (PmST1), 111
 "Click Chemistry", 104, 111
 cloning and characterization of well reported *Escherichia coli* K-12 sialic acid aldolase, 102
 cloning and purifying *Photobacterium damsela* α 2,6-sialyltransferase (Pd2,6ST), 102, 104
 conjugation of functionalized, with terminal azido to other biomolecules, 104, 108, 109
 efficiency of chemoenzymatic synthesis of, derivatives, 101
 high-throughput substrate specificity assays, 114, 118
 high-throughput substrate specificity studies, 114–115
 one-pot three-enzyme approach, 98–102
 one-pot three-enzyme chemoenzymatic synthesis of, with natural and non-natural functionalities, 101, 103
 one-pot three-enzyme chemoenzymatic synthesis of α 2,3-linked, 112t, 113t
 one-pot three-enzyme chemoenzymatic synthesis of p-nitrophenyl (pNP)-tagged, libraries, 116, 117
 one-pot two-step enzymatic synthesis, 100, 101
 preparative synthesis of α 2,3-linked, with natural and non-natural sialic acids, 105, 111
 preparative synthesis of α 2,6-linked, with natural and non-natural sialic acids, 102, 104–105
 "Staudinger Ligation", 104, 111
 synthesis of α 2,6-linked sialyl-N-acetylglucosamine using one-pot multi-enzyme system, 98, 99
Sialyl Tn (STn)
 carbohydrate expressed tumor antigen, 198, 218
 immunological evaluation of O- and C-glycosides, 234t
 keyhole limpet hemocyanin (KLH) conjugates, 230f
 synthesis of C-glycoside analog of, 221, 223
 synthesis of double C-glycoside analog, 223, 226
 tumor antigen STn O-glycoside and its C-glycoside analog, 218, 221f
See also Anticancer vaccine candidates; Theratope (sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine

- Solid-phase synthesis, branched peptidoglycan derivative, 63–64, 68*f*
- Staudinger ligation
 - application, 266
 - mechanism, 263*f*
 - selective tagging of azide-containing glycoconjugates, 262
 - sialosides, 104, 111
- Structural diversity, sialic acids, 95–96
- Sugar oxazoline
 - modified, as substrates for transglycosylation, 132, 134, 138
- See also* Glycopeptides
- Sugars, bioorthogonal reactions on, 257*f*
- Sulfoxide and trichloroacetimidate glycosylation, hyaluronic acid trisaccharides, 38, 40, 41

T

- Theratope (sialyl TN-keyhole limpet hemocyanin (STn-KLH)) vaccine
 - ADCC (antibody-dependent cell-mediated cytotoxicity) cells, 212
 - antibody testing against ovine submaxillary mucin (OSM) and STn, 201
 - ASCT (autologous peripheral blood stem cell transplantation) patients, 198–199
 - breast and ovarian cancer ASCT patients by treatment group, 200*t*
 - cytotoxicity assay, 201
 - definitions, 202
 - EFS (event-free survival) by estrogen receptor status and treatment in breast cancer patients, 203, 206*f*
 - EFS by relapse risk and treatment in breast and ovarian cancer patients, 203, 204*f*
 - factors of EFS in vaccinated breast cancer patients, 203, 209
 - failure counts and percentages for EFS in vaccinated breast cancer patients, 210*t*
 - humoral and cellular immunity to, 209, 211
 - immune function testing, 203, 208*t*
 - immune reconstitution after ASCT, 211–212
 - immune responses in vitro, 201
 - methods, 199–202
 - multivariable hazard ratios for overall survival (OS) in vaccinated breast cancer patients, 209, 211*t*
 - OS by estrogen receptor status and treatment in breast cancer patients, 203, 207*f*
 - OS by relapse risk and treatment in breast and ovarian cancer patients, 203, 205*f*
 - outcome for breast cancer patients, 212
 - outcome of all patients, 203
 - proliferation assay, 201
 - sialyl Tn (STn) carbohydrate expressed tumor antigen, 198
 - statistical considerations, 202
 - study design, 199, 201
- Thio-carboamino peptides. *See* Glycosyl thio-carboamino peptides
- Thiol-maleimide conjugation, protein methodology, 258
- Transglycosylation
 - endoglycosidases (ENGases) catalyzing glycopeptide synthesis, 124–125
 - transition-state analog substrates for ENG-catalyzed, 125, 128, 132
- See also* Glycopeptides
- Trichloroacetimidate method
 - bacterial cell wall synthesis, 60

combination of sulfoxide and, for
hyaluronic acid trisaccharides,
38, 40, 41

U

Ulosonic acids
biological roles, 217
See also Sialic acids

V

Vaccines. *See* Anticancer vaccine
candidates; Theratope (sialyl TN-
keyhole limpet hemocyanin (STn-
KLH)) vaccine