uOttawa

L'Université canadienne
Canada's university

FACULTÉ DES ÉTUDES SUPÉRIEURES ET POSTOCTORALES

uOttawa

L'Université canadienne
Canada's university

FACULTY OF GRADUATE AND POSDOCTORAL STUDIES

Pawel Czechura
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Chemistry)
GRADE / DEGREE

Department of Chemistry
FACULTÉ, ÉCOLE, DÉPARTEMENT / FACULTY, SCHOOL, DEPARTMENT

Elucidating the Relationship between Hydration and Ice Recrystallization Inhibition with C-linked Antifreeze Glycoproteins

TITRE DE LA THÈSE / TITLE OF THESIS

Robert Ben
DIRECTEUR (DIRECTRICE) DE LA THÈSE / THESIS SUPERVISOR

CO-DIRECTEUR (CO-DIRECTRICE) DE LA THÈSE / THESIS CO-SUPERVISOR

EXAMINATEURS (EXAMINATRICES) DE LA THÈSE / THESIS EXAMINERS

Christopher Boddy

William Ogilvie

Kenneth Storey

Dennis Whitfield (National Research Council)

Gary W. Slater
Le Doyen de la Faculté des études supérieures et postdoctorales / Dean of the Faculty of Graduate and Postdoctoral Studies

ELUCIDATING THE RELATIONSHIP BETWEEN HYDRATION AND ICE RECRYSTALLIZATION INHIBITION WITH C-LINKED ANTIFREEZE GLYCOPROTEINS

PAWEL CZECHURA

Thesis submitted to the
Faculty of Graduate and Postdoctoral Studies
University of Ottawa
In partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

Ottawa-Carleton Chemistry Institute
University of Ottawa
Ottawa, Ontario
Canada

Your file Votre référence
ISBN: 978-0-494-61362-7
Our file Notre référence
ISBN: 978-0-494-61362-7

ABSTRACT

Antifreeze glycoproteins (AFGPs) are a subclass of biological antifreezes isolated from Atlantic and Antarctic Telost fish. At subzero temperature environments, these compounds have the ability to inhibit ice crystal growth and protect these organisms from cryoinjury and death. The native AFGPs show considerable promise as novel cryoprotectants, however, their limited bioavailability, cytotoxicity and inherent chemical and biological instability have precluded their widespread use. Consequently, rationally designed, non-toxic and stable AFGP analogues became an attractive synthetic challenge. While structural modifications made in previously synthesized C-linked AFGP analogues addressed these issues, further SAR studies were required to elucidate the key structural features necessary for potent inhibitors of ice recrystallization. The SAR studies described in this dissertation verify that hydration is related to ice recrystallization inhibition of the C-linked AFGP analogues.

The role of hydration in modulating solution conformation, molecular recognition and biological activity of oligosaccharides, proteins and nucleotides is widely recognized but is often neglected when investigating many biological processes such as the mechanism by which biological antifreezes inhibit the growth of ice. In chapter 3, we have investigated the relationship between carbohydrate configuration and recrystallization-inhibition (RI) activity of functional *C*-linked AFGP mimetics. The presented results suggest that the configuration of the carbohydrate moiety in *C*-linked AFGP analogues is extremely important and modulates recrystallization-inhibition activity. It seems likely that differences in hydration for each *C*-linked pyranose alter the compatibility of the carbohydrate moiety with the three-dimensional

hydrogen-bonded network of supercooled bulk water. Consequently, the energy associated with transferring a water molecule to the ice lattice changes and can result in inhibition of ice growth.

The study of hydration has been extended and in chapter 4 the first synthesis of *C*-linked AFGP analogues containing disaccharides has been presented. A method for estimating hydration indices for the glycopeptides was developed and used to rationalize recrystallization inhibition activity trends. The results signify that the relative hydration of the *C*-linked AFGP analogues directly correlates to the degree of their recrystallization-inhibition activity. Moreover, it was demonstrated that hydration of the AFGP mimetics can be estimated prior to their synthesis. This indicates that superior inhibitors of recrystallization may be engineered based on their predicted hydration.

The hydration model developed in chapter 4 was applied to explain RI results presented in chapter 5. It has been established that C-2 N-acetyl moiety of the carbohydrate and the L-alanine residues of the polypeptide backbone improve the RI activity of the C-linked AFGP analogues. Incorporating β-D-Gal-(1-->3)-D-GalNAc into our C-linked AFGP mimetics demonstrated that challenging custom synthesis of carbohydrates is not necessary to obtain a potent recrystallization inhibitor. Analogue containing β-D-Glc-(1-->3)-D-GalNAc proved that stereochemistry of the terminal carbohydrate is very important for RI activity and is optimal when it matches that of galactose. The C-2 N-acetyl moiety of the carbohydrate was shown to influence conformation, thus hydration and consequently recrystallization inhibition property of the C-linked AFGP analogues.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Prof. Robert N. Ben, for all his support, insightful ideas, enthusiasm, patience, critique and encouragement. His deep knowledge of chemistry, biochemistry as well as biology inspired me to appreciate synthetic organic chemistry and its applications. He was always a great colleague and mentor.

Also, I would like to thank my thesis committee members, Prof. Christopher Boddy, Prof. William Ogilvie, Prof. Kenneth Storey and Dr. Dennis Whitfield for reading this dissertation and providing me with essential suggestions for its improvement.

I had a pleasure to work with a number of amazing individuals and excellent scientists. I thank: Dr. Vincent Bouvet for helping me in the first few months, Roger Tam for good friendship and excellent teamwork, John Trant for his wisdom and kindness, Jackie Tokarew for delicious goodies brought to the lab, Frank Cease, Taz Cheema, Ivan Petrov, Wendy Campbell, Taline Boghossian, Jennifer Chaytor, Liz von Moos, Gloria Gong and everyone else that I ever worked with for great time of our lives we had as the Ben Group. I want to extend my gratitude to Dr. Mike Souweha for his friendship and honest advice, and Dr. Mathieu Leclare for all the extracurricular work he has done to help everyone in DRO 420. Further, I express my appreciation to the undergraduate students that contributed to my work: Aleksandra Paliga and Elena Dimitrijevic, and everyone else who contributed directly or indirectly to this work. A special thanks goes to Sandra Ferreira and my sister Krystyna Czechura that helped me correct this thesis.

Most of all, I would like to thank my parents and family for their unconditional support and love.

Thank You Everyone!

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF GRAPHS

Chapter 3

Chapter 4

Chapter 5

LIST OF SCHEMES

Chapter 3

Chapter 4

Chapter 5

LIST OF TABLES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

LIST OF ABBREVIATIONS

Ac	Acetyl
AFP	Antifreeze protein
AFGP	Antifreezeglycoprotein
Ala	Alanine
br	broad
Bn	Benzyl
Boc	tert-Butyloxycabonyl
Bz	Benzoyl
°C	degrees Celcius
CD	Circular dichroism
CH_2Cl_2	Dichloromethane
cm	Centimeter
d	doublet
dd	doublet of doublets
DBU	1,8-Diazabicyclo[5.4.0]undec-7-ene
DEPT	Distortion Enhancement of Polarization Transfer
DIPEA	Diisopropylethylamine
DIS	Dynamic Ice-Shaping
DMF	*N,N*-dimethylformamide
DMAP	4-dimethylamino pyridine
DMSO	Dimethylsulfoxide

dt	doublet of triplets
eq	equivalents
ESI	Electrospary ionization
Et	Ethyl
EtOAc	Ethyl acetate
Et_2O	Diethyl ether
Fmoc	9-Fluorenylmethoxycarbonyl
Gal	Galactose
GalNAc	N-acetyl-galactosamine
Glc	Glucose
Gly	Glycine
h	hour
HI	Hydration index
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HN	Hydration number
HOBT	*N*-Hydroxybenzotriazole
Hz	Hertz
IR	Infrared
kDa	kiloDalton
Lys	Lysine
m	multiplet
M+	Marent molecular ion
MALDI	Matrix-assisted laser desorption ionization

Man	Mannose
Me	Methyl
MeCN	Acetonitrile
MeOH	Methanol
MGS	Mean grain size
mL	milliliters
MLGS	Mean largest grain size
MS	Mass spectrometry or mass spectra or molecular sieves
MTT	Methylthiazolydiphenyl-tetrazolium bromide
NMR	Nuclear magnetic resonance
Orn	Ornithine
P	page
PBS	Phosphate buffered saline
PMC	Partial molar compressibility
pmm	parts per million
PMV	Partial molar volume
PPII	Polyproline type II
Pro	Proline
Pd/C	Palladium on carbon
q	quartet
RI	Recrystallization inhibition
s	singlet
SAR	Structure activity relationship

SEM	Standard error of mean
SPS	Solid phase synthesis
SPPS	Solid phase peptide synthesis
t	triplet
TBS	*tert*-Butyldimethylsilyl
TBDMS	*tert*-Butyldimethylsilyl
TFA	Trifluoroacetic acid
TH	Thermal hysteresis
THF	tetrahydrofuran
Thr	Threonine
TMSOTf	Trimethylsilyltrifluoromethane sulphonate
δ	chemical shift (in ppm)
ν	frequency (in cm^{-1})

CHAPTER 1

Introduction

1.1. Glycoproteins

Glycoproteins, a subclass of glycoconjugates, are proteins containing oligosaccharides covalently linked to selected amino acid residues.[1] They are important in many biological functions that include cell adhesion, cell differentiation, signal transduction, host-pathogen interactions, and immune response as well as their use as therapeutic agents (Table 1.1).[2, 3]

Table 1.1. Examples of glycoproteins and their function.[3]

Function	Glycoproteins
Structural molecule	Collagens
Lubricant and protective agent	Mucins
Transport molecule	Transferrin, ceruloplasmin
Immunologic molecule	Immunoglobins, histocompatibility antigens
Hormone	Chorionoic gonadotropin, thyroid-stimulating hormone (TSH)
Enzyme	Various (e.g. alkaline phosphatase)
Cell attachment-recognition site	Various proteins involved in cell-cell (e.g. sperm-oocyte), virus-cell, bacterium-cell, and hormone cell interactions
Antifreeze	Certain plasma proteins of coldwater fish
Interact with specific carbohydrates	Lectins, selectins (cell adhesion lectins), antibodies
Receptor	Various proteins involved in hormone and drug action
Affect folding of certain proteins	Calnexin, calreticulin
Regulation of development	Notch and its analogs, key proteins in development
Hemostasis (and thrombosis)	Specific glycoproteins on the surface membranes of platelets

Within any polypeptide there are only a few types of amino acids suitable for the attachment of a carbohydrate residue.[4] The amino acids must contain a reactive side chain such as amine, amide, thiol or hydroxyl group in order to be glycosylated. The most common of these amino acids are: aspargine, cystine, hydroxylysine, hydroxyproline, serine and threonine. The name of a particular class of glycoproteins, just like terminology of the glycosylation reaction, comes from the atom linking the amino acid and the saccharide (for example: nitrogen; N-linked glycoproteins, N-glycosylation). Hence, there are four classes of glycoproteins: N-linked, O-linked, S-linked and C-linked.[4b] C-linked glycoproteins are the most stable of the glycoproteins, as they are resistant towards acidic, basic and enzymatic degradation. In biological systems their abundance is low due to the difficulty of carbon-carbon bond formation.[5]

In cells N-glycosylation takes place in the endoplasmic reticulum while further carbohydrate modification and oxygen-, sulfur- and carbon-linked glycosylations are usually performed in the Golgi apparatus (Figure 1.1).[6, 4a] The linkages between amino acids and saccharides are assembled by glycosyltransferases, enzymes that act as catalysts in the transfer of a monosaccharide from an activated sugar phosphate to an acceptor molecule.[6] Since glycosylation is one of the most ubiquitous forms of posttranslational modification, the compartment where it occurs may vary depending on certain cellular condition and subsequently the localized enzyme concentration.[7]

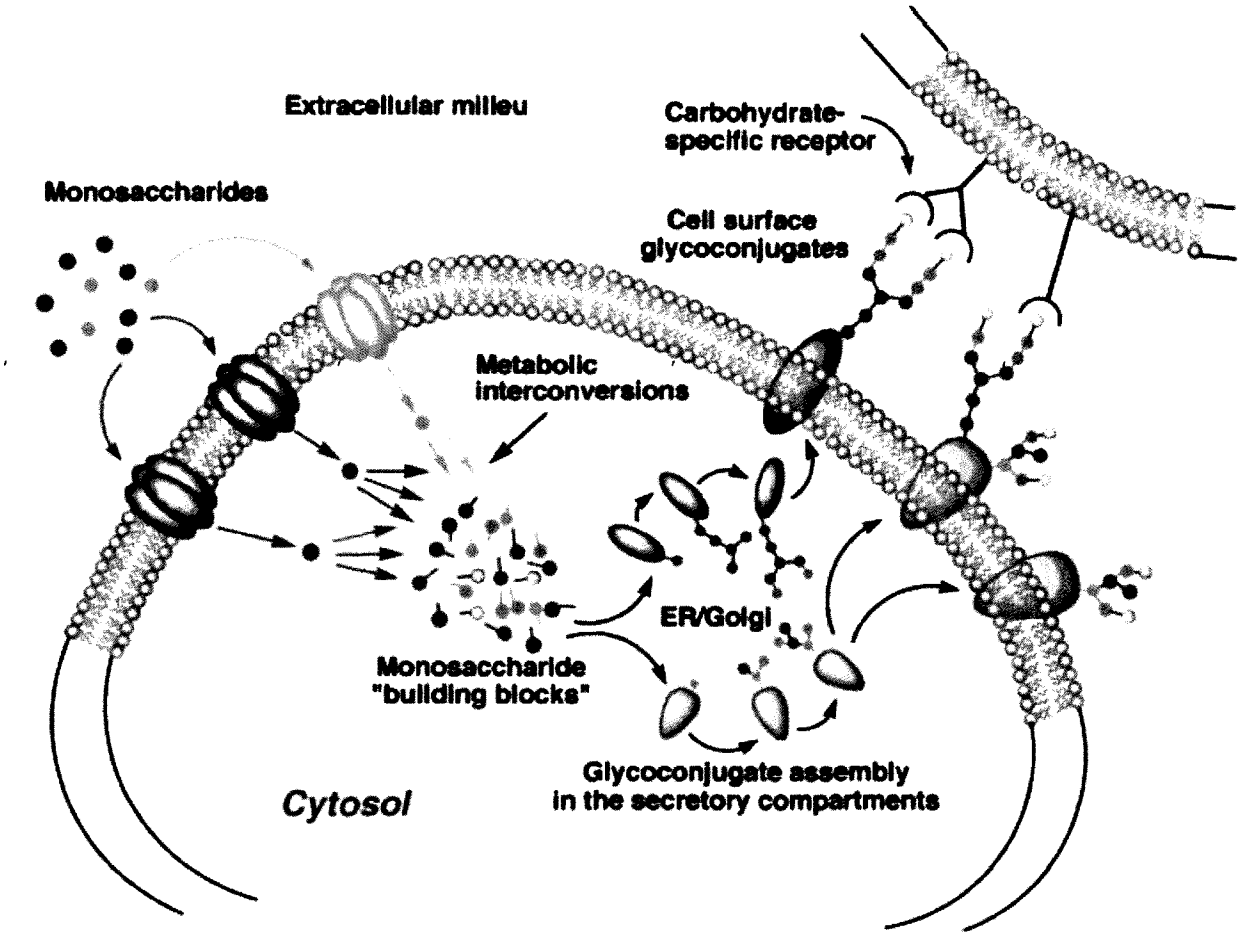


Figure 1.1. Glycoconjugate biosynthesis and cell surface recognition.[4a]

There is a wide range of structural modifications of glycoproteins which leads to a broad spectrum of molecules with an array of physiological properties. The variations in the peptide backbone are achieved by changes in length and/or changes in the amino acid sequence. Branching is also a possible source of diversity but it is limited due to the two directional nature of the peptide chain growth, either at the N- or the C- terminus. These structural modifications can alter flexibility, folding and consequently conformation of the glycoproteins. Nevertheless, the majority of glycoproteins complexity arises from the broad range of carbohydrate

modifications [1, 2b] Therefore, with respect to structural diversity, glycoproteins have the capacity to far exceed proteins and nucleic acids.

A variety of sugars ranging from monosaccharides to dendrimers and polysaccharides (also known as glycans) can be attached to a polypeptide. The most common for glycoproteins are oligosaccharides which are comprised of two to ten monosaccharide units. The most well-known monosaccharides are: D-galactose, D-glucose (and their amino and N-acetyl versions), D-mannose, D-talose and L-fucose. Each one of them exists in equilibrium between their open chain, furanose (5-membered ring) and pyranose (6-membered ring) forms (Figure 1.2).[8] The equilibrium population, although different for each sugar, is strongly biased towards the most stable form, the pyranose. Furanoses and pyranoses exist as diastereomers, called α and β anomers, that vary only in their configuration at the anomeric center (C1). The ratio of the α/β anomers differs for each sugar and is governed by the number of axial and equatorial hydroxyls as well as the strength of the anomeric effect. Anomeric effect is the tendency of the heteroatomic substituent at C1 of the carbohydrate to reside in the (axial) α position. This effect is attributed to the antiperiplanar donation of the electrons from the oxygen of the pyranose ring to the σ* antibonding orbital of the C1-substitutent (exocyclic) bond resulting in stereoelectronic stabilization (Figure 1.2).[9] Since configuration at the anomeric center can vary, α and β N-linked, O-linked, S-linked, and C-linked glycosidic bonds can be formed between C1 of a given sugar and an amino acid or another carbohydrate. Consequently, combinations of natural hexoses and their ability to form complex networks strongly influence the structural diversity of the glycoproteins and their broad range of functions.[2, 3, 10] Their unique combinations make up libraries of biological recognition sites, yet their simplest forms play key roles vital to human life.[11]

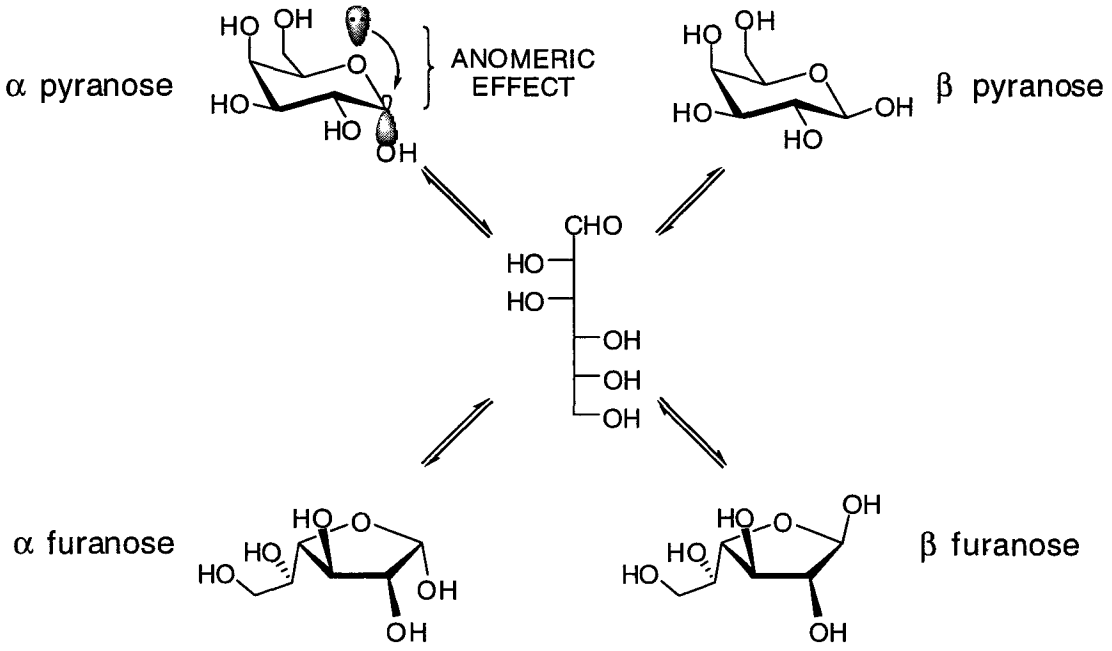


Figure 1.2. Illustration of anomeric effect and equilibrium mixture of cyclic and acyclic forms of D-galactose.

Sucrose is the most common sugar in plants and animal nutrition. It consists of α-D-glucopyranosyl-(1-->2)-β-D-fructofuranoside which under strong acidic conditions, or by an enzyme called glycoside hydrolase, is converted to its subunits glucose and fructose. Carbohydrates are one of the most important energy sources for mammals. They are easily converted to CO_2, H_2O and energy which is stored in a form of ATP or NADH. D-Glucose for example is broken down into two three-carbon pyruvate molecules yielding two ATP molecules and two high energy NADH molecules (Figure 1.3).

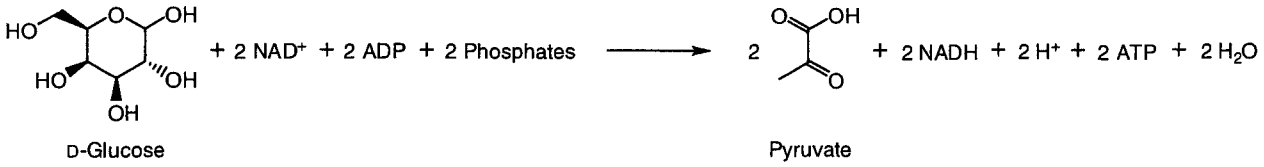


Figure 1.3. Overall reaction of glycolysis.[11]

Therefore, most glycoproteins can be considered as biodegradable polymers that can be recycled and also used as a source of energy. Lactose, β-D-galactopyranosyl-(1-->4)-D-glucopyranose, is a disaccharide found in milk and makes up 2-8% of its weight composition. It is broken down to D-galactose and D-glucose by an enzyme called lactase. Dairy farmers produce hundreds of million tons of milk yearly that impact many countries economic growth.

Another example of a biologically important disaccharide is β-D-galactopyranosyl-(1-->3)-2-acetamido-2-deoxy-D-galactopyranose (β-D-Gal-(1-->3)-D-GalNAc), a major constituent of blood group antigens.[1, 12] It is an essential feature the glycopeptide's core that is attached to the ABO(H) antigenic determinants (Figure 1.4). Subunit β-D-Gal-(1-->3)-D-GalNAc α-O-linked to serine or threonine is called T- or Thomsen-Friedenreich antigen, while the monosaccharidic N-acetyl-D-galactosamine version is termed Tn-antigen. The secreted human glycoproteins contain the T-antigens and the disaccharide entities that span the glycan foundation. The blood group active structure is in a form of glycolipids that are incorporated in the human erythrocyte membrane. T and Tn antigens are also known as cancer markers. Both are over-expressed in cancer cells and cause the metastatic behavior of tumors due to the antiadhesive shielding effects responsible for the lack of cell-cell or cell-matrix contacts. Thus, β-D-Gal-(1-->3)-D-GalNAc is an important disaccharide in many biological processes and is a

key building block of various glycan residues. Furthermore, it is an essential component of antifreeze glycoproteins (AFGPs); a focus of this dissertation.[13]

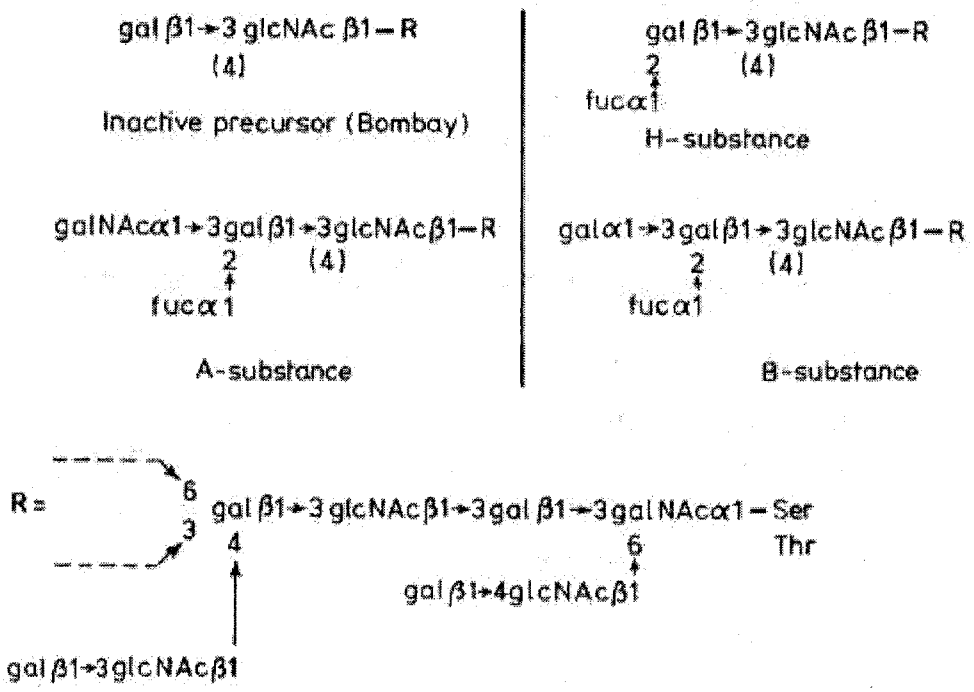


Figure 1.4. Secreted human glycoproteins with ABO(H) antigenic determinants.[1]

1.2. Biological Antifreezes

Biological antifreezes are peptide based compounds found in organisms such as fish, invertebrates and plants that inhabit subzero environments. They protect these organisms from cryoinjury, and sustain their survival. They were first isolated from Antarctic fish *Trematomus bernacchil* and *Trematomus borchgrevinki* by Devries and Wohlschlang in late 1960s and early 1970s.[13, 14] The isolated compounds that showed potent antifreeze activity were antifreeze glycoproteins (AFGPs). Their structure was not fully understood at that time but was later confirmed to be a peptide with 20% of D-galactosamine content. Since then a number of

antifreeze compounds were isolated, characterized and studied.[15] However, the most thoroughly studied compounds were that of teleost fish populating Antarctic and Arctic regions.

The freezing point of seawater occupied by teleost fish is approximately -1.9°C.[16] This is a result of freezing point depression induced by the dissolved solutes. The osmolytes (sodium chloride, urea, and amino acids) present in fish allow for approximately -0.9°C freezing point depression. Extremely active biological antifreezes enable an additional freezing point depression bringing teleost fish freezing point to approximately -2.0°C. Their concentration is about 30 mg/mL, thus their mode of action is strictly non-colligative.[17] The term applies to a physical activity of a substance that depends on its identity and does not depend only on the ratio of the number of particles of solute and solvent in the solution. The colligative effect of this concentration would only result in 0.002°C freezing point depression despite the high molecular weight (2.6-33.0 kDa) of the glycopeptide. The potency of biological antifreezes is their ability to bind to ice crystals and alter their growth by creating a localized freezing point depression. The details of the effect will be discussed in subsequent section on physical properties of biological antifreezes.

Figure 1.5 compares the freezing point temperatures of a protein (lysozyme), galactose, sodium chloride, AFGP and AFGP with sodium chloride.[14b, 15b] As with the salt, a classical colligative effect is observed with the protein and galactose which are used to approximate the effect of AFGP. However, AFGP exhibits a non-colligative effect much greater than that of the protein and the carbohydrate. It is approximately 500 times more active then sodium chloride when compared on molar basis.[14b] This verifies AFGPs high potency as a biological antifreeze. The additive effect of freezing point depression is observed when AFGP is combined with

sodium chloride, which implies that there is no interaction between them. This solution is similar to internal fluids of the fish which contain salts and the biological antifreezes.

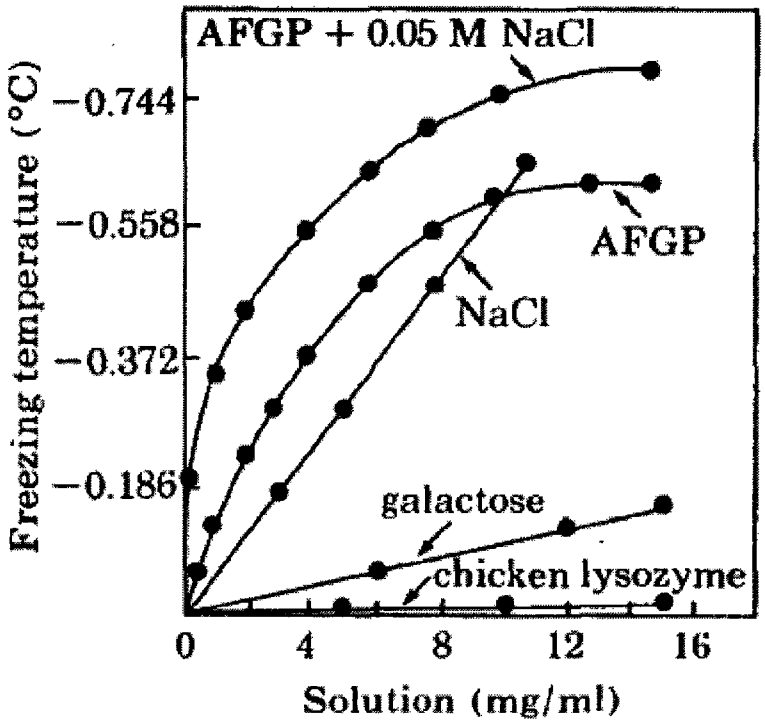


Figure 1.5. Freezing temperatures of colligative (lysozyme, galactose and sodium chloride) and non-colligative (AFGP) substances.[14b, 15b]

In vivo biological antifreezes are localized mainly in blood plasma, however they are also present in other tissues.[18] AFGPs were found to be secreted into the intestinal and ocular fluids at very low concentrations. Recent studies have shown that biological antifreezes are not only present in the extra cellular media but can also be internalized.[15h] The role of internalization is still unknown but possibly it is an important aspect of cryoprotection. Biological antifreezes are produced in the liver and distributed throughout the fluid compartments of an organism. Valerio

et al. have also isolated them from epithelium cells demonstrating that their synthesis can also be accomplished by other cell types.[19]

There are two classes of biological antifreezes: antifreeze proteins (AFPs) and antifreeze glycoproteins (AFGPs).[20] AFPs differ from one another in their primary, secondary and tertiary structures, while AFGPs possess low structural diversity and differ from one another mainly in their molecular weight. Both classes of compounds show similar physical properties.

1.2.1. Antifreeze Proteins (AFPs)

Antifreeze proteins are small proteins ranging from 3 to 24 kDa able to elicit antifreeze activity. All of them vary from one another in their amino acid composition, their folding and thus their primary, secondary and tertiary structures. Due to their diversity they were divided into four main groups: Type I, II, III and IV AFPs (Table 1.2).[15j]

Table 1.2. Classification and key structural features between AFPs.[15j]

Characteristic	Type I AFP	Type II AFP	Type III AFP	Type IV AFP
Mass (Da)	3300 – 4500	11000 – 24000	6500	12000
Key Properties	Alanine-rich α-helix	Disulfide bonded	β-sandwich	Alanine rich; helical bundle
Representative Structure				
Natural Source	Right-eyed flounders; sculpins	Sea raven; smelt; herring	Ocean pout; wolfish; eel pout	Longhorn sculpin

Type I AFPs were first reported by Duman and Devries, and Hew.[21] They are L-alanine rich proteins with mainly α helical structure which predominance increases to 85% as the temperature is decreased to 0°C.[22] Their molecular mass is in the range of 3 to 5 kDa. The L-alanine residues align on one side of the helix, while the other side is occupied by L-threonine hydroxyls. This unique structure gives type I AFPs an amphiphilic structure.[20]

Type II AFPs are also composed of mainly L-alanines, however they exhibits a significant number of L-cystine residues. L-cystine disulfide bonds formed force hydrophilic residues of L-threonine to one side of this 11-24 kDa globular protein giving them their unique structure.[23] Its secondary structure consists of α helices, β sheets and random coils.

Type III AFPs have no dominant amino acid residue. They are globular proteins with a molecular mass of approximately 6.5 kDa. Their β sheet structures are orthogonally arranged giving them a β sandwich structure.[24] They have only been isolated from Arctic and Antarctic pout.[25]

Type IV AFPs have been isolated in the late 1990s from Longhorn sculpin.[26] They are L-alanine and L-glutamine rich peptides and possess α helical bundle structure.[15i] Their molecular mass is about 12 kDa.

In recent years other types of AFPs have been isolated from moths, beetles and snow fleas.[21c, 15i, 27] These antifreeze proteins are 10 to 30 times more potent than those of the fish. Organisms which live on land battle temperatures approaching -30°C, thus their body needs more effective antifreezes. Owing to the large population of insect a number of novel AFPs will arise over the next few decades.

1.2.2. Antifreeze Glycoproteins (AFGPs)

Antifreeze glycoproteins are tripeptide repeating units of L-threonine-L-alanine-L-alanine α-O-linked to a disaccharide, β-D-Gal-(1-->3)-D-GalNAc. Unlike AFPs they have a lot of homology and mainly differ from one another in their length. They are grouped into eights fractions according to their molecular weight: AFGP1-8.[28] AFGP1 is the longest one with molecular mass of 33.7 kDa while AFGP8 is the shortest with molecular mass of 2.6 kDa. Based on molar concentrations, AFGP1 is more effective at lowering the freezing point of a solution then AFGP8. The advantage of using AFGP8 is their ease of chemical synthesis, as isolation of the compound is very labor intensive and expensive. Consequently, AFGPs are excellent candidates for chemical modification and analogue synthesis.

Structural variations within the AFGP family have been reported (Figure 1.6).[29] Some AFGPs have shown changes of the carbohydrate from N-acetyl-D-galactosamine to N-acetyl-D-glucosamine.[30] In low molecular weight AFGP6-8 substitution of L-threonine by L-arginine and L-alanine by L-proline has been observed.[31] Amino acid variations are rare but do occur. Several isoforms, different peptide sequences, of AFGP have been found in AFGP6 isolated from rock cod (*Gadus ogac*).[15h] This shows that the specific type of the AFGP contains a mixture of glycopeptides with a defined molecular mass. The three-dimensional structure of AFGPs is similar to that of type I AFP. The hydrophilic disaccharide units align on one side of the left handed helix while the hydrophobic methyls of the L-alanines and methyl groups of the N-acetyl-D-galactosamine on the other.[32] This arrangement gives AFGPs an amphiphilic structure, however some evidence shows that the conformation of the AGFPs in solution is made up of random coil and unfolded structure.[14a, 33]

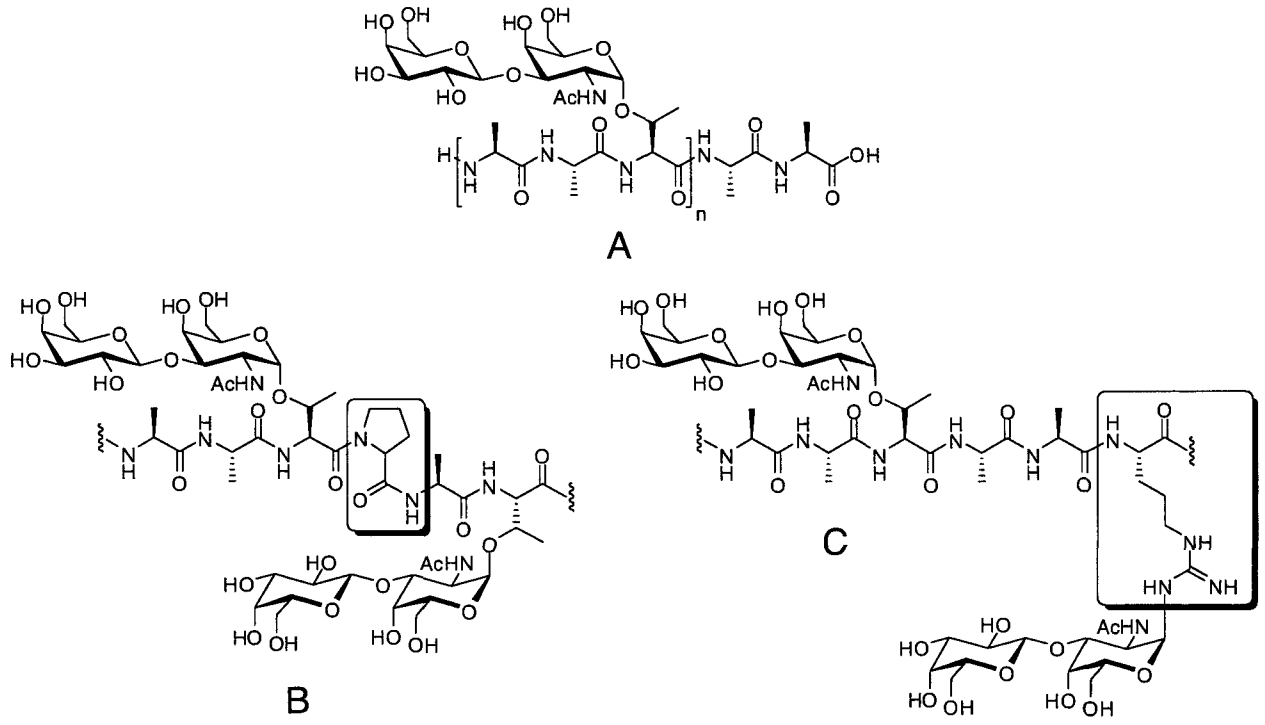


Figure 1.6. Structural variations of AFGP; A: most common AFGP, B: L-alanine replaced by L-proline, and C: L-threonine replaced by L-arginine.[29]

1.3. Physical Properties of Biological Antifreezes

General mechanism of action of biological antifreezes is based on an adsorption inhibition process.[34] The detailed mechanism of action is not fully understood however several hypothesis have been suggested. Raymond and Devries were the first to propose the idea that antifreeze molecules tightly bind to the surface of the growing ice.[14g] Consequently, the molecules of water can only be added to the ice lattice where the antifreeze molecules are not present. This results in a bulging ice front which growth is now thermodynamically unfavorable (Figure 1.7). The localized freezing point depression created is also known as the Kelvin effect.[35]

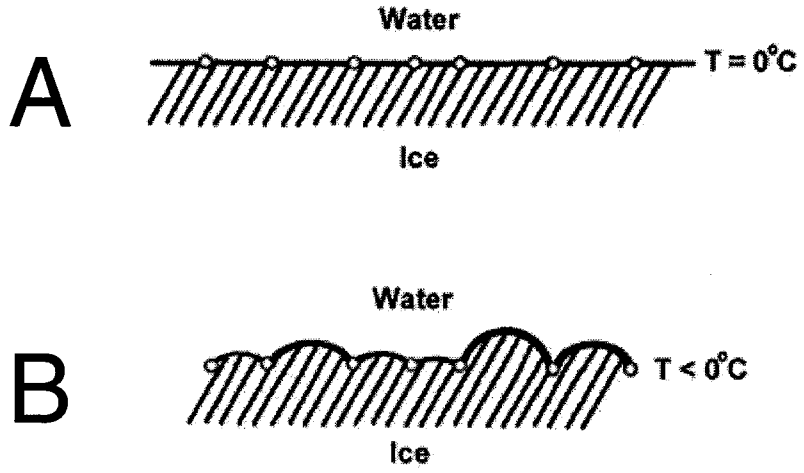


Figure 1.7. Ice crystal lattice with adsorbed antifreeze molecules at (A) T=0°C and (B) T<0°C illustrating bulging ice front (Kelvin effect).[35]

Knight and Devries proposed that the binding of the biological antifreezes to the surface of the ice is irreversible.[34a] The suggestion was based on the assumption that since there are a multiple number of hydrogen bonds formed between the antifreeze molecule and the ice, the probability of them being broken simultaneously is low.[36] Studies by Hew and Wen contradict the irreversible binding hypothesis.[37] Alternative mechanism proposed was based on the hydrophobic interaction of the antifreeze molecules and the ice though the van der Walls forces.

The binding mechanism was used to rationalize thermal hysteresis which is one of the two main physical properties of the biological antifreezes. The role of binding on the second physical property, recrystallization inhibition, is still a matter of debate. Nonetheless, studies associated both properties give important insights to further understanding of the mechanism of action of biological antifreezes.

1.3.1. Thermal Hysteresis (TH)

Thermal hysteresis is a freezing point depression and has very little effect on the melting point. It has been used to assess the antifreeze activity by measuring the freezing point depression temperature, a temperature at which ice starts to grow from a seed ice crystal.[38]

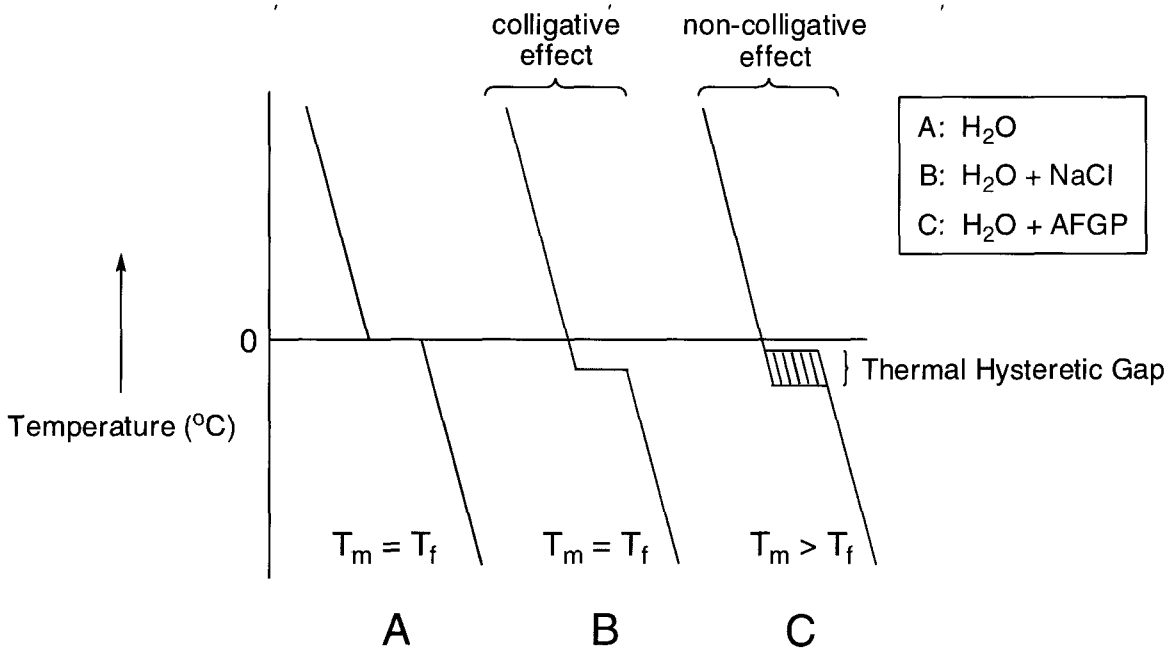


Figure 1.8. Quantitative diagram of thermal hysteresis.

Figure 1.8 is a quantitative diagram of thermal hysteresis. Melting and the freezing temperature of pure water occur at 0°C. When solutes such as sodium chloride are added to water, freezing and melting points are equally depressed due to the colligative effect. AFGPs in water have a different effect, the melting temperature stays approximately constant (it slightly drops due to the colligative effect of AFGPs) while freezing temperature drops significantly. The freezing point drop varies and it strictly depends on the type of the biological antifreeze. The difference

between the freezing and the melting point is referred to as thermal hysteretic gap. The biggest thermal hysteresis has been found in AFPs isolated from insects and possesses the gap of approximately 6°C.[27] The thermal hysteretic gap of AFGPs isolated from fish is approximately 2°C. The TH activity is a non-colligative effect unique to biological antifreezes due to their ability to bind to ice crystal lattice and temporarily stop its growth.

1.3.2. Dynamic Ice-Shaping

Dynamic Ice-Shaping is defined as a modulation of ice crystal morphology into different shapes.[39] As the ice crystal seed within the thermal hysteretic gap grows it may form different shapes. The shapes depend on the identity of the biological antifreeze, AFGPs for example form hexagonal bipyramidal structures (Figure 1.9).

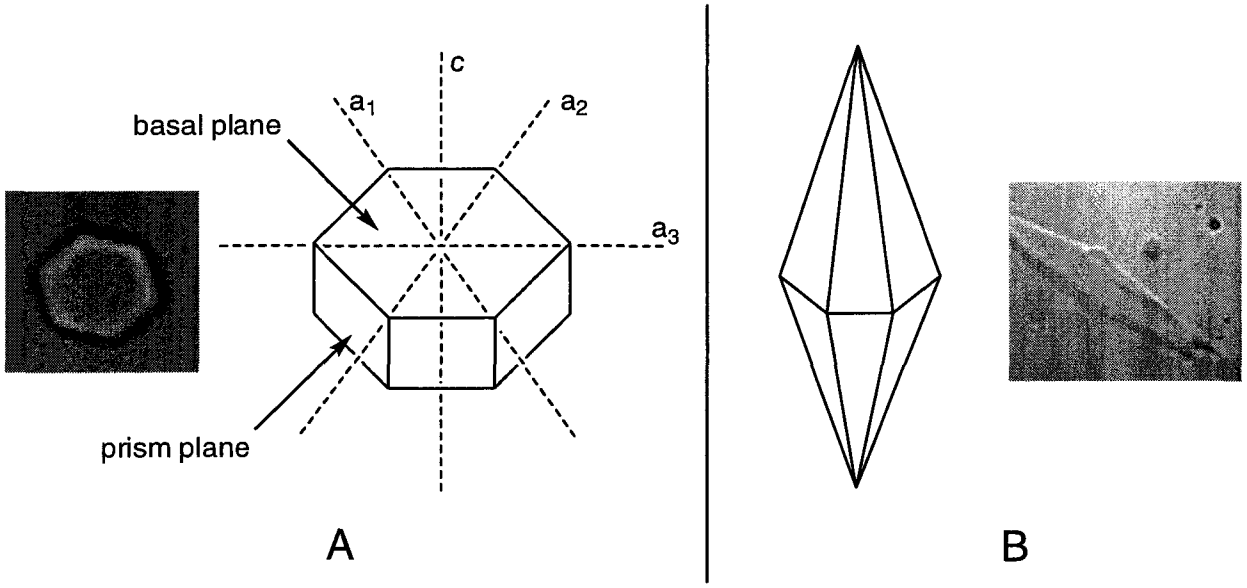


Figure 1.9. Dynamic ice shaping in the presence of AFGP; A: seeded ice crystal within the TH gap, B: hexagonal bipyramidal ice crystal.

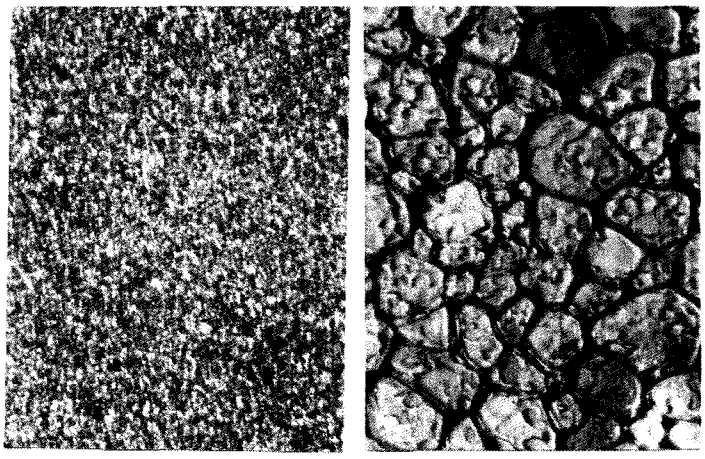
As pure water freezes the ice grows along the a-axis, which is parallel to its basal plane, forming a disk-like structure. The dynamics of ice growth in presence of biological antifreezes are altered. As ice grows any solutes present are excluded from the solid formed. This effect is exactly the same as when purifying a substance by a process of re-crystallization: as crystals grow the impurities are excluded from the uniformly formed matrix. Antifreeze molecules have the ability to adsorb and integrate into the ice crystal lattice. This was studied by Knight and co-workers using etching technique and proved that antifreeze molecules bind to the planes (a_1, a_2, a_3 and c) of the nucleating ice crystals thus block their growth. Different antifreeze molecules adsorb to different planes of the ice crystal.[36, 39] The unique binding depends on their folding and thus matching their binding moieties with the surface of the ice. This results in a specific ice crystal morphology that is halted while the temperature keeps dropping and forming the thermal hysteretic gap. As the temperature approaches the critical freezing point, the molecules of the antifreeze still inhibit the ice growth along the a-axis but not the c-axis which is now overgrown by the new layer of water molecules.[40] The growth on the basal plane (c-axis) continues forming hexagonal bipyramidal ice crystals. As the temperature decreases further these crystals burst into needle like ice-spicules.[14g]

1.3.3. Recrystallization Inhibition (RI)

Recrystallization is defined as the formation of larger crystals at the expense of smaller ones, and its mechanism has been dealt with extensively in the metallurgical literature.[41] With respect to ice, larger crystals are formed by the addition of bulk water molecules to a transitional domain between the ice/water interface known as the quasi-liquid layer (QLL) and subsequently from the QLL into the ice lattice.[42] The effect occurs to minimize the grain boundary energy by

increasing the surface area of the ice crystals. Biological antifreezes have the ability to inhibit this process, thus they possess recrystallization inhibition activity.

When a solution of a biological antifreeze is flash frozen at -78°C the ice crystals that are formed are relatively small (Figure 2.0).[43] This process of crystallization excludes the solute out of the crystal lattice creating highly concentrated boundary solution between the ice crystals. The antifreeze molecules present between the ice crystals bind to the ice grain boundaries, prevent their reorganization and consequently formation of large ice crystals.[44] It has been reported that concentrations as low as 10^{-12} M show effective recrystallization inhibition activity.[44a, 45]



A B

Figure 1.10. Photographs depicting the relative size of ice crystals in (A) a solution of AFGP 8 (10 mg/mL concentration) and (B) in the PBS control using the RI assay. The magnification is the same in both images.[43]

Another way to induce recrystallization inhibition is to decrease the diffusion of water between the ice boundaries.[44c] Knight et al. have shown that some peptides which possess no thermal hysteretic activity have the ability to inhibit ice recrystallization.[46] The recrystallization

inhibition activity was only observed in water when no other solutes were present. The presence of inorganic salts forced the antifreeze molecules into the bulk water thus excluding them from the QLL layer and as a result diminishing the effect. Their activity was dependent on and proportional to their concentrations, and is known as non-specific RI effect. This and other studies have demonstrated that thermal hysteresis and recrystallization inhibition act via two different mechanisms and therefore the two effects are not coupled.[43, 47]

1.4. Structure Activity Relationship (SAR) Studies

To date antifreeze activity in structure activity studies was determined by measuring thermal hysteresis.[20, 14b, 32, 48] This unique property was of main interest to the researchers due to its striking effect of selectively lowering the freezing point temperature. Studies have shown that some biological antifreezes which possess thermal hysteresis can cause ice crystal nucleation and hypothermic cell damage.[16, 49] Thus, in the past decade recrystallization inhibition and its potential applications have received more attention. The effect can be induced at very low antifreeze concentrations and shows promising results in cryopreservation of cells and tissues.

The importance of the hydrophobic and hydrophilic interactions is not fully understood, however the amphiphilic nature of the biological antifreeze molecules gives a lot of insight into the mechanism of action and the antifreeze activity.[20, 32] The molecular conformation of a particular AF(G)P has to be considered in order to postulate a rational binding interaction. It has been found that different types of antifreeze molecules bind to different planes of the ice lattice.[37b, 39, 50] Due to their diversity, AFPs adopt different conformations which suggest that their mode of binding to ice varies. The result is observed in formation of specific crystal shapes and patterns. The ability of biological antifreezes to bind to different planes of the ice lattice

relies on the matching of their topology to that of a particular plane of the ice.[37b] Thus, hydrophobic and hydrophilic interactions may be responsible for the two different mechanisms of TH and RI. [20, 32, 51]

Type I AFP is the most studied of the AFPs due to their relatively simple α-helical structure.[21] This structure is very similar to that of AFGP and this suggests that hydrophilic moieties are aligned on one side of it while the hydrophobic on the other. The frequency of the residues of the α-helix turn matches different planes of the ice lattice.[22] Type I AFP was considered to bind to ice though hydrogen bonding of the L-threonine residues.[52] On the contrary, further studies indicated that hydrophobic interactions are responsible for the ice binding.[20, 51b] Structure activity relationship studies were performed on type I AFP isolated from Winter flounder and the results show that β-methyl groups in L-threonine residues are necessary for antifreeze activity. This was proven by Hayment et al who used site-directed mutagenesis and changed the L-threonine amino acid to L-serine.[20] This had no effect on the overall conformation as well as the hydrogen bonding but it resulted in diminished thermal hysteresis and no interaction with ice lattice as indicated by ice hemisphere test. They also replaced L-threonine with L-valine and observed no effect on the antifreeze activity. Thus, these SAR studies proved that β-methyl groups play a more important role, then the hydroxyls, in binding to ice. Consequently, this suggests that the hydrophobic sites of the protein are the ones responsible for adsorption and interaction with the ice lattice.

Type III AFPs were studied by Yang et al.[53] The group was successful in determining the structure of type III AFP isolated from Ocean pout by high resolution X-ray crystallography. The planarity of the surface was analyzed and is composed of 67% of non-polar area. The flat hydrophobic region is thought to bind to ice lattice though the van der Waals interactions. Due to

the limited understanding of the AFP type II and IV structure, no details regarding their mechanism of action has been reported.

Structure activity relationship studies of AFGPs were performed to point out important features necessary for achieving thermal hysteresis. AFGPs vary mainly in length and it has been shown that AFGPs with molecular weight higher then 13kDa had approximately three times the thermal hysteresis then the shorter ones. The longer the glycoprotein is, the better the thermal hysteretic activity. [14b, 14d, 28, 48b]

The disaccharide moiety of the AFGP was modified and studied by Feeney and co-workers.[14b] Removal of the disaccharide and the hydroxyl group of the L-threonine via β elimination resulted in complete loss of thermal hysteresis. Peracetylation of the carbohydrate or its oxidation also resulted in loss of the TH activity. The oxidation of the C6 hydroxyl groups to aldehydes had no influence on TH, thus implying they are not important structural features. However, their substitution by negatively charged groups such as carboxylates or sulphates showed a diminished activity.

In 2004, the importance of the disaccharide and the hydrophobic interactions between ice and AFGPs were further addressed in a structure activity study by Nishimura and coworkers.[32] The group has synthesized several different analogues and analyzed them for antifreeze activity. The linkage between the L-threonine and the sugar was changed from α to β resulting in no TH activity. The same effect was observed when the N-acetyl moiety of the N-acetyl-D-galactosamine was changed to a hydroxyl, forming a β-D-Gal-(1-->3)-D-Gal disaccharide. Coupling a sialic acid to the C3 hydroxyl of the galactose showed no thermal hysteresis. The loss of TH activity may be due to a possible conformational change of the carbohydrate and impairment of AFGPs ice binding ability, as described by Nishimura. Analogues containing N-

acetyl-D-galactosamine and N-acetyl-D-lactosamine exhibited a weak TH. This implies that N-acetyl-D-galactosamine is a crucial component necessary for antifreeze activity. As in previous studies, β methyl group of the L-threonine, when changed to hydrogen, proved to be an essential feature for TH activity. It is possibly related to retaining a specific conformation of the hydrophilic residue, the disaccharide. When a monomer of the AFGP (β-D-Gal-(1-->3)-α-D-GalNAc-[Thr-Ala-Ala]) was tested, it only exhibited some dynamic ice shaping. As type I AFPs, AFGPs also form a simple amphiphilic α helix. Hence, it is believed that the hydrophobic methyl groups of the L-alanines are also necessary for the antifreeze activity. Hydrophobic and hydrophilic interactions therefore account for the effective antifreeze activity of the AFGPs. A detailed mechanism of action is still unknown; speculations were only made whether the hydrophilic or hydrophobic residues bind to the surface of the ice. [20, 32, 36, 51]

1.5. Applications

Biological antifreezes show a vast potential of applications in different areas. Their benefits lie in their ability to: (1) disrupt the ice crystal growth within the TH gap, (2) inhibit the growth of ice crystals under prolonged period of time, and (3) stabilize cell membranes against osmotic shock caused by extracellular ice growth.[55] Their disadvantage in cryoprotection is the dynamic ice shaping associated with TH. The formation of needle shaped ice crystals and their burst at subfreezing temperature causes mechanical cell damage. However, even this property has found its application in the area of cryosurgery. Another disadvantage discovered by Ben et al. is cytotoxicity of the AFGP8 toward mammalian cell lines (Figure 2.1).[56]

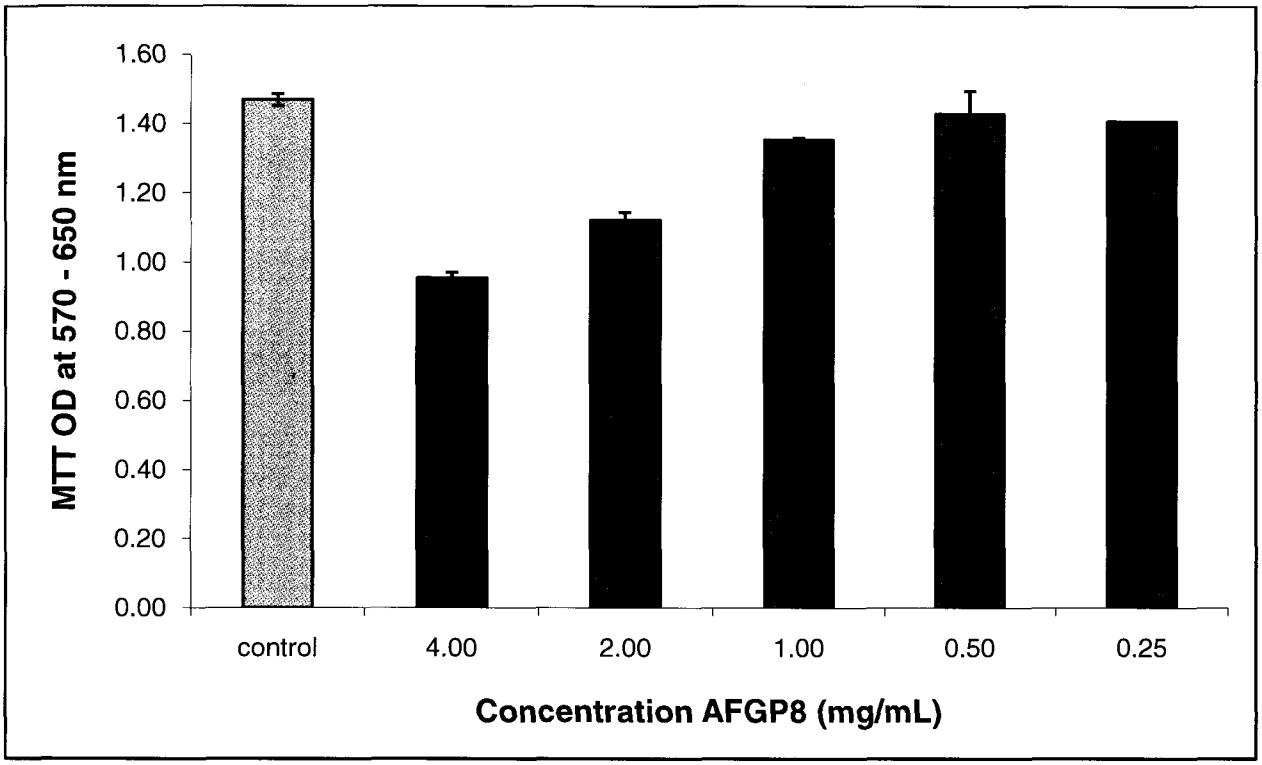


Figure 1.11. AFGP8 cytotoxicity MTT assay in Human cells. The MTT is reduced in mitochondria to formazan by healthy cells. The result is observed by a color change from yellow to purple, and is measured by optical density of the solution.[56]

1.5.1. Cryosurgery

Cryosurgery is a technique that employs extreme cold to destroy abnormal or diseased tissue. It has been used for treatment of a number of skin diseases ranging from moles to skin cancers.[57] The destructive force of freezing temperatures is caused by ice crystal formation that ruptures the tissue of interest. The technique however is not 100% effective as reoccurrence of the symptoms does exist. Rubinsky at el recently applied cryosurgery to tumor destruction. In vivo tumors were injected with type I AFP which exhibits strong dynamic ice shaping.[58] The researchers observed an increased efficiency of cryosurgery with mortality close to 100%. The

burst of the ice spicules caused by the antifreeze protein showed an amplified effect of mechanical damage of the tumor cells.

1.5.2. Cryopreservation

Cryopreservation is defined as a technique used to preserve cells, tissues or organs at sub-zero temperatures.[58c, 59] Their viability or biological function is always reduced by cryopreservation process due to physical cell rupture, necrosis or cold induced apoptosis. Owing to the stress induced by the extracellular and intracellular ice formation as well as the fluctuating cell volumes, the most common cell death associated with cryosurgery is mechanical damage to the external cell membrane by the ice crystals and/or osmotic stress.[16]

Ice can be formed intracellularly and extracellularly.[16, 49] Intracellular ice formation is always lethal to cells and no studies have shown that the damage caused can be prevented. Extracellular ice formation leads to: (1) dehydration of cells caused by the osmotic pressure gradient across the cell membrane, and (2) physical injury induced by the pressure of large ice crystals on the cell membrane.

As the temperature is lowered, the dynamic lipid bilayer forms two phases: liquid and crystal phase. During this thermotropic phase transition, the mismatch in the packing of the lipids is created and leads to the leakage of the intracellualar content and ultimately cell death.[60] When the temperature reaches the freezing point, extracellular ice formation occurs first. This results in an increase in solute concentration and therefore pressure outside of the cell. The cell compensates for the osmotic pressure increase by regulating the flow of water though the lipid bilayer. The rate of the osmotic flux is directly proportional to the supercooling and when the rate of the extracellular ice growth is fast, the cell may rupture due to the hypertonic stress and

water depletion. In addition to the aforementioned effects, physical stress induced by the growth of the ice crystals also disrupts the lipid bilayer membrane. As the temperature is lowered even further the whole sample freezes. During sample storage or thawing, recrystallization occurs resulting in formation of large crystals that also apply mechanical stress on the cell membrane. Biological antifreezes have the ability to protect cells at hypothermic and cryogenic temperatures and minimize the abovementioned effects of ice growth though TH and RI activity. Two hypotheses were made regarding cell integrity and membrane stabilization to reduce the damage caused by hypothermia. They state that biological antifreezes can: (1) block the potassium and calcium ion channels, and (2) stabilize cell membrane during the thermotropic transition phase.[61]

The effects of biological antifreezes and their cryoprotective effects vary based on type of cells being preserved, type and concentration of the antifreeze used, cooling rates and storage temperatures.[60c] Cell death has been observed due to the DIS and the advantage of the effect has been explored by cryosurgery. Cell toxicity has been induced by some of the biological antifreezes, therefore a proper match has to be made between cryoprotectant and a cell line of interest.

1.5.3. Food preservation

The quality of frozen food is lost due to the adverse effect of ice formation and ice crystal growth at subfreezing temperatures.[15f, 62] Texture and nutrition loss of frozen meats during multiple freezing/thawing cycles are caused by the large ice crystals formed during recrystallization process.[63] Dehydration and water holding capability are also affected by the ice growth. Stabilization of cell membranes and recrystallization inhibition ability of biological antifreezes makes them excellent candidates for cryopreservation of frozen food. Their use is

limited by their difficulty of availability. AFPs and AFGPs can be extracted or synthesized only in small amounts, thus their utility strictly depends on their cost-benefit. Although successful results have been observed in the effect of biological antifreezes on frozen foods, biomedical applications such as cryopreservation are the most profitable of the applications.

1.6. References

1. Hughes, R. C., *Glycoproteins*, Chapman and Hall, New York, USA, 1983.
2. (a) Dwek, R. A., *Chem. Rev.* **1996**, 96, 683. (b) Varki, A., *Glycobiol.* **1993**, 3, 97.
3. Murray R. K., Granner, D. K., Mayers, P. A., Rodwell, V. W., *Harper's Illustrated Biochemistry*, McGraw-Hill Co., USA, 2006.
4. (a) Bertozzi, C. R., Kiessling, L. L., *Science* **2001**, 2357. (b) Marcaurelle, L. A., Bertozzi, C. R., *Chem. Eur. J.* **1999**, 5, 1384. (c) Thayer, D. A., Yu, H. N., Galan, M. C., Wong C. H., *Angew. Chem. Int. Ed.* **2005**, 44, 4596.
5. (a) Vilegenthart, J. F. G., Casset, F. *Curr. Opin, Struct. Biol.* **1998**, 8, 565. (b) Hofsteenge, J., Muller, D. R., de Beer, T., Loffler, A., Richter, W. J., Vilegenthart, J. F. G., *Biochemistry*, **1994**, 33, 13524. (c) de Beer, T., Vilegenthart, J. F. G., Loffler A., Hofsteenge, J., *Biochemistry*, **1995**, 34, 11785.
6. Czlapinski, J. L., Bertozzi, C. R., *Curr. Opin. Chem. Biol.* **2006**, 10, 645.
7. Graffenried, C. L., Bertozzi, C. R., *Curr. Opin. Cell Biol.* **2004**, 26, 356.
8. Stick, R. V., *Carbohydrates: The Sweet Molecules of Life*, Academic Press, San Diego, California, USA, **2001**.
9. (a) Lemieux, R., *Pure. Appl. Chem.* **1971**, 25, 527. (b) Lemieux, R., Can. J. Chem. 1969, 47, 442. (c) Deslongchamps, P., *Stereoelectronic Effects in Organic Chemistry*, Pergamon Press Ltd., Oxford, **1983**.
10. Pratt, M. R., Bertozzi, C. R., *Chem. Soc. Rev.* **2005**, 34, 58.
11. Moran, L. A., Scrumgeour, K. G., Horton, H. R., Raymond, S. O., Rawn, J. D., *Biochemistry*, Prentice-Hall Inc., **1994**.
12. Beak M. G., Roy, R., *Revs. Molec. Biotech.* **2002**, 90, 291.

13. DeVries, A. L., Wohlschlag, D. E., *Science* **1969**, 163, 1073.

14. (a) DeVries, A. L., Komatsu, S. K., Feeney, R. E., *J. Biol. Chem.* **1970**, 245, 2901. (b) Komatsu, S. K., DeVries, A. L., Feeney, R. E., *J. Biol. Chem.* **1970**, 245, 2909. (c) Scholander, P. F., Maggert, J. E., *Cryobiology* **1971**, 8, 371. (d) DeVries, A. L., Vandenheede, J., Feeney, R. E., *J. Biol. Chem.* **1971**, 246, 305. (e) Duman, J. G., DeVries, A. L., *Cryobiology*, **1972**, 9, 469. (f) Lin, Y., Duman, J. G., DeVries, A. L., *Biochem. Biophys. Res. Commun.* **1972**, 46, 87. (g) Raymond, J. A., DeVries, A. L., *Cryobiology* **1972**, 9, 541.

15. (a) DeVries, A. L., *Ann. Rev. Phys.* **1983**, 45, 245. (b) Feeney, R. E., Burcham, T. S., Yeh, Y., *Ann. Rev. Biophys. Biophys. Chem.* **1986**, 15, 59. (c) DeVries, A. L., *Methods Enzymol.* **1986**, 127, 293. (d) Hew, C. L., Yang, D. S. C., *Eur. J. Biochem.* **1992**, 203, 33. (e) Davies, P. L., Sykes, B. D., *Curr. Opin. Struct. Biol.* **1997**, 7, 828. (f) Feeney, R. E., Yeh, Y., *Trends Food Sci. Technol.* **1998**, 9, 102. (g) Harding, M. M., Ward, L. G., Haymet, A. D. J., *Eur. J. Biochem.* **1999**, 264, 653. (h) Ewart, K. V., Lin, Q., Hew, C. L., *Cell. Mol. Life Sci.* **1999**, 55, 271. (i) Jia, Z., Davies, P. L., *Trends. Biol. Sci.* **2002**, 27, 101. (j) Hardling, M. M., Anderberg, P. I., Haymet, A. D. J., *Eur. J. Biochem.* **2003**, 270, 1381.

16. Costanzo, J. P., Lee, J. E., DeVries, A. L., Wang, T., Layne, J. R. *The FASEB Journal*, **1995**, 9, 351.

17. Morris, H. R., Thompson, M. R., Osuga, D. T., Ahmed, A. L., Chan, S. M., Vandenheede, J. R., Feeney, R. E., *J. Biol. Chem.* **1978**, 253, 5155.

18. (a) O`Grady, S., Clark, A., DeVries, A. L., *Exp. Zool.* **1982**, 220, 179. (b) Alghren, J. A., Cheng C. C., Schrag, J. D., DeVries, A. L., *J. Exp. Biol.* **1988**, 137, 549.

19. (a) Valerio, P. F., Kao, M. H., Fletcher, G. L., *Can. J. Zool.* **1990**, 68,1065. (b) Valerio, P. F., Kao, M. H., Fletcher, G. L., *J. Exp. Biol.* **1992**, 164, 135.

20. Haymet, A. D. J., Ward, L. G., Harding, M. M., *J. Am. Chem. Soc.* **1999**, 121, 941.

21. (a) Duman, J. G., DeVries, A. L., *Nature*, **1974**, 247, 237. (b) Duman, J. G., DeVries, A. L., *Comp. Biochem. Physiol.* **1976**, B54, 375. (c) Ewart, K. V., Lin, Q., Hew, C. L., *Cell. Mol. Life Sci.* **1999**, 55, 271.

22. Yang, D. S., Sax, M., Chakrabartty, A., Hew, C. L., *Nature* **1988**, 333, 232.

23. (a) Slaughter, D., Fletcher, G. L., Ananthanarayanan, V. S., Hew, C. L., *J. Biol. Chem.* **1981**, 256, 2022. (b) Ng, N. F., Trinh, K. Y., Hew, C. L., *J. Biol. Chem.* **1986**, 261, 15690.

24. Sonnichsen, F. D., Sykes, B. D., Chao, H., Davies, P. L., *Science*, **1993**, 259, 1154.

25. DeLuca, C. I., Davies, P. L., Ye, Q., Jia, Z., *J. Mol. Biol.* **1998**, 275, 515.

26. (a) Deng, G. J., Andrews, D. W., Laursen, R. A., *FEBS Lett.* **1997**, 402, 17. (b) Deng, G. J., Laursen, R. A., *Biochim. Biophys. Acta* **1998**, 1388, 305.

27. Graham, L. A., Davies, P. L., *Science*, **2005**, 310, 461.

28. Feeney, R. E., *Am. Sci.* **1974**, 62, 712.

29. Harding, M. M., Anderberg, P. I., Haymet, A. D. J., *Eur. J. Biochem.* **2003**, 270, 1381.

30. Wohrmann, A. P. A., *Mar. Ecol. Prog. Series* **1996**, 130, 47.

31. (a) Geoghegan, K. F., Osuga, D. T., Ahmed, A. I., Yeh, Y, Feeney, R. E., *J. Biol. Chem.* **1980**, 255, 663. (b) Hew, C. L., Slaughter, D., Fletcher, G. L., Joshi, S. B., *Can. J. Zool.* **1981**, 59, 2186, (c) O'Grady, S. M., Schrag, J. D., Raymond, J. A., DeVries, A. L., *J. Exp. Zool.* **1982**, 224, 177. (d) Fletcher, G. L., Hew, C. L., Joshi, S. B., *Can. J. Zool.*

1982, 60, 348. (e) Bucham, T. S., Osuga, D. T., Rao, B. N. N., Bush, C. A., Feeney, R. E., *J. Biol. Chem.* **1986**, 261, 6384.

32. Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishmura, S., *Angew. Chem. Int. Ed.*, **2004**, 43, 856.

33. (a) Tomimatsu, Y., Scherer, J., Yeh, Y., Feeney, R. E., *J. Biol. Chem.* **1976**, 251, 2290 (b) Berman, E., Allerhand, A., DeVries, A. L., *J. Biol. Chem.* **1980**, 255, 4407.

34. (a) Knight, C. A., DeVries, A. L., *J. Cryst. Growth* **1994**, 143, 301. (b) Ben, R. N., *Chembiochem*, **2001**, 2, 161.

35. Wilson, P., *Cryoletters*, **1993**, 14, 31.

36. Knight, C. A., Driggers, E., DeVries, A. L., *Biophys. J.* **1993**, 64, 252.

37. (a) Hew, C. L., Yang, D. S. C., *Eur. J. Biochem.* **1992**, 203, 33. (b) Wen, D., Laursen, R. A., *Biophys. J.* **1992**, 63, 1659.

38. Fletcher, G. L., Hew, C. L., Davies, P. L., *Annu. Rev. Physiol.* **2001**, 63, 359.

39. Knight, C. A., Cheng, C. C., DeVries, A. L., *Biophys. J.* **1991**, 59, 409.

40. Wilson, P. W., Beaglehole, D., DeVries, A. L., *Biophys. J.* **1993**, 64, 1878.

41. (a) Pronk, P.; Infante Ferreira, C. A.; Witkamp, G. J. *J. Crystal Growth* **2005**, 275, e1355. (b) Huige, N. J. J.; Thijsenn, H. A. C. *J. Crystal Growth* **1972**, 13/14, 483. (c) Inada, T.; Modak, P. R.; *Chem. Eng. Sci.* **2006**, 61, 3149. (d) Kingery, W. D. *J. Appl. Phys.* **1959**, 30, 301.

42. (a) Karim, O. A., Haymet, A.D.J., *J. Chem. Phys.* **1988**, 89, 6889. (b) Knight, C. L., *Nature* **2000**, 406, 249. (c) Sadtchenko, V., Ewing, G.E., *J. Chem. Phys.* **2002**, 116, 4686.

43. Horwath K. L., Easton C. M., Poggioli Jr., G. J. P., Myers K., Schnorr I. L. *Eur. J. Entomol.* **1996**, 93, 419.

44. (a) Knight, C. L., Hallet, J., DeVries, A. L., *Cryobiology* **1988**, 25, 55. (b) McKown, R. L., Warren, G. J., *Cryobiology* **1991**, 28, 474. (c) Yeh, Y., Feeney, R. E., McKown, R. L., Warren, G. J., *Biopolymers* **1994**, 34, 1495.

45. Knight, C. A., DeVries, A. L., Oolman, L. D., *Nature* **1984**, 308, 295.

46. Knight, C. A., Wen, D. Y., Laursen, R. A., *Cryobiology* **1995**, 32, 23.

47. Sidebottom, C., Buckley, S., Pudney, P., Twigg, S., Jarman, C., Holt, C., Telford, J., McArthur, A., Worrall, D., Hubbard, R., Lillford, P., *Nature* **2000**, 406, 256.

48. (a) Ahmed, A. I., Yeh, Y., Osuga, D. T., Feeney, R. E., *J. Biol. Chem.* **1976**, 251, 3033. (b) DeVries, A. L., *Science* **1971**, 171, 1152. (c) Osuga, D. T., Feather, M. S., Shah, M. J., Feeney, R. E., *J. Prot. Chem.* **1989**, 8, 519.

49. (a) Pegg, D. E., *Semin, Reprod. Med.* **2002**, 20, 5. (b) Meryman, H. T., *Annu. Rev. Biophys. Bioeng.* **1974**, 3, 341.

50. Chou, K. C., J. *Mol. Biol.* **1992**, 233, 509. (b) Jorgensen, H., Mori, M., Matsui, H., Kanaoka, M., Yanagi, H., Yanusaki, Y., Kikuzono., *Protein Eng.* **1993**, 6, 19.

51. (a) Wierzbicki, A., Taylor, M. S., Knight, C. A., Madura, J. D., Harrington, J. P., Sikes, C. S., *Biophys. J.* **1996**, 71, 8. (b) Chao, H., Houston Jr., M. E., Hodges, R. S., Kay, C. M., Sykes, B. D., Loewen, M. C., Davies, P. L., Sonnichsen, F. D., *Biochemistry* **1997**, 36, 14652.

52. DeVries, A. L., Lin, Y., *Biochim. Biophys. Acta* **1977**, 495, 388.

53. Yang, D. S., Hon, W. C., Bubanko, S., Xue, Y., Seetharaman, J., Hew, C. L., Sicheri, F., *Biophys. J.* **1998**, 74, 2142.

55. (a) Rubinsky, B., Arav, A., DeVries, A. L., *Cryobiology* **1992**, 29, 69. (b) Storey, K. B., Bischof, J., Rubinsky, B., Am. J. Physiol. **1992**, 263, R185.

56. Liu, S., *Ph.D. Thesis*, University of Ottawa, Ottawa, **2006**.

57. (a) Maiwand, M. O., Asimakopoulos, G., *Technol. Cancer Res. Treat.* **2004**, 3, 151. (b) Kuflik, E. G., *Dermatol. Surg.* **2004**, 30, 297.

58. (a) Koushafar, H., Rubinsky, B., *Urology* **1997**, 49, 421. (b) Pham, L., Dahiya, R., Rubinsky, B., *Cryosurgery* **1999**, 38, 169. (c) Wang, J. H., *Cryobiology* **2000**, 41, 1.

59. Tablin, F., Oliver, A. E., Walker, N. J., Crowe, L. M., Crowe, J. H., J. *Cell. Physiol.* **1996**, 168, 305.

60. (a) Mazur, P., *J. Gen. Physiol.*, **1963**, 47, 347. (b) Clerc, S. G., Thompson, T. G., *Biophys. J.* **1995**, 58, 2333. (c) Quinn, P. J., *Cryobiology* **1995**, 22, 128.

61. (a) Hincha, D. K., DeVries, A. L., Schmitt, J. M., *Biochim. Biophys. Acta* **1993**, 1146, 258. (b) Hays, L., Feeney, R. E., Crowe, L. M., Crowe, J. H., Oliver, A. E., *Proc. Natl. Acad. Sci. USA* **1996**, 93, 6835.

62. Griffith, M., Ewart, K. V., *Biotech. Adv.* **1995**, 13, 375.

63. (a) Payne, S. R., Sandford, D., Harris, A., Young, O. A., *Meat Sci.* **1994**, 37, 429. (b) Payne, S. R., Young, O. A., *Meat Sci.* **1995**, 41, 147.

CHAPTER 2

Goals and Objectives

2.1. Previous Precedent for Rational Design of AFGP Analogues

AFGPs show considerable promise as novel cryoprotectants as they are able to inhibit the recrystallization of ice.[1] Recent reports also suggest that they may stabilize cell membranes[2] during thermotropic phase transitions associated with freezing/thawing. Unfortunately several problems have precluded the widespread use of these compounds. Firstly, the native AFGP are O-linked glycoproteins and hence the anomeric carbon-oxygen bond is sensitive to cleavage under a variety of biological and chemical conditions.[3] For instance, under acidic conditions the glycosidic bond is readily cleaved forming an oxocarbenium ion which can be further converted to its corresponding disaccharide by reacting with water, while under basic conditions the O-linked threonine can be cleaved via β-elimination (Figure 2.1). Secondly, Ben et al. have shown that native AFGP8 exhibits significant cytotoxicity in human embryonic liver cell lines at physiological and cryogenic temperatures.[4] Thirdly, quantities of AFGPs are limited as they must be isolated from biological sources.[5] Thus, to address the abovementioned issues, series of synthetic AFGP analogues needs to be synthesized. The analogues should: (1) possess a potent RI and lack TH activity, (2) be structurally stable, (3) be non-cytotoxic, and (4) be easily accessible via conventional chemical routes.

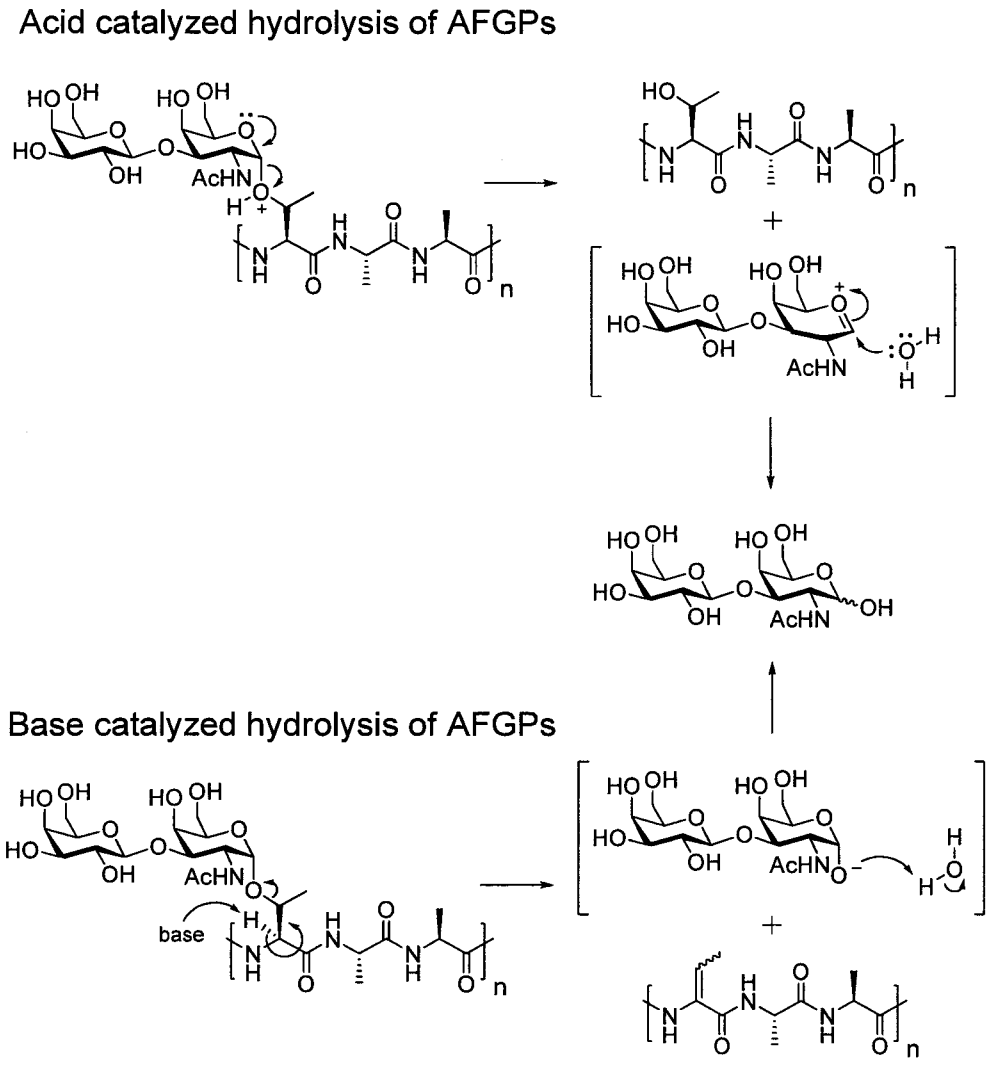


Figure 2.1. Acidic and basic hydrolysis of AFGPs.[3b]

Several structurally significant features of native AFGPs have been identified as essential for antifreeze protein-specific activity.[6] These include: (1) carbohydrate (α-O-liniked β-D-Gal-(1-->3)-D-GalNAc residue), (2) β-methyl group of the threonine residue, and (3) the Thr-Ala-Ala tripeptide backbone (Figure 2.2). In order to determine which of these structural features are dominant contributors to RI activity, structure-activity-relationships (SARs) need to be performed.

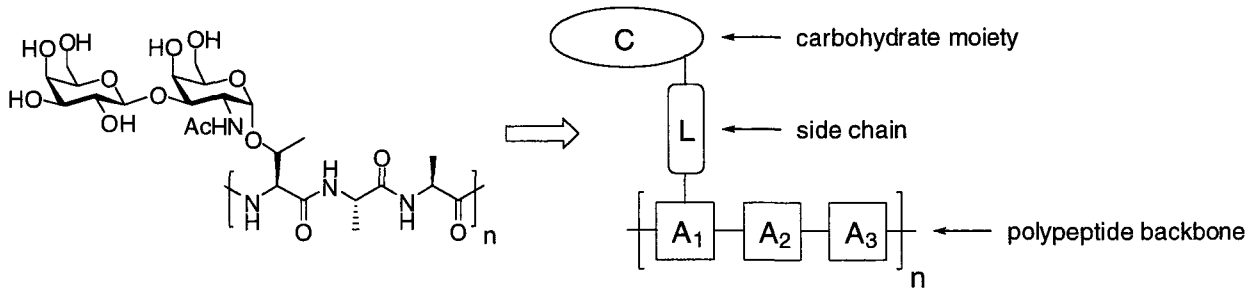


Figure 2.2. Three major modification constituents of AFGPs.

Prior SAR work has demonstrated that carbohydrate stereochemistry is important for TH activity,[6, 7] however, the influence in RI activity is not known. This work verified that the relative stereochemistry at C3 and C4 of the terminal galactose residue was important as well as the N-acetyl group of the GalNAc residue.

Given that the synthesis of O-linked carbohydrates can be quite challenging, much effort has been devoted to the preparation of glycoprotein mimetics.[8] Typically, carbon-linked (C-linked) or sulfur-linked (S-linked) derivatives are commonly employed as both analogues have been shown to exhibit physical and biological properties comparable to that of the O-linked oligosaccharides. In our laboratory, we have chosen to utilize a C-linked glycoconjugate instead of an O-linked glycoconjugate. This will greatly facilitate synthesis and impart additional chemical and biological stability in these analogues. Furthemore, it is expected that these analogues will greatly contribute to the current understanding of how a potent inhibitor of ice recrystallization functions. Consequently, a number of C-AFGP analogues will be prepared in which the carbohydrate and primary amino acid sequence will be structurally modified.

2.1.1. Investigating the Importance of the Carbohydrate in the O-linked AFGP Analogues

To date, the Ben laboratory have synthesized and assessed approximately 50 different O-linked and C-linked AFGP analogues composed of only four repeating tripeptide sequences.[4, 9] This is the minimum length required to observe TH activity.[9] In one particular series of compounds, the carbohydrate component was modified using: D-galactose, D-glucose, D-mannose and N-acetyl-D-galactosamine monosaccharides as an alternative to the native disaccharide.[9a] (Figure 2.3). These analogues were subsequently assessed for both TH and RI activity and these results are detailed below.

TH Activity of O-linked AFGP analogues:

Unfortunately, no TH activity was observed for any of the analogues, however, DIS was observed only for analogues $[T(Gal)AA]_4G$ and $[T(Man)AA]_4G$. These findings suggest that the terminal galactose is essential interacting with the ice lattice. We hypothesize that the loss of TH maybe due to the lower number of hydroxyl groups present on the molecule which prevents a strong interaction with the ice lattice.[9]

RI Results

Approximately 66% of RI activity of analogue $[T(GalNAc)AA]_4G$ was lost compared to AFGP8. This finding verifies that D-galactose plays a significant role in RI activity. At glycopeptide concentration of 5.54×10^{-06} M, an increase in RI was observed from D-glucose (MGLS=0.0087±0.0002) being the worst, following D-mannose (MGLS=0.0070±0.0002), D-galactose (MGLS=0.0056±0.0002) and N-acetyl-D-galactosamine (MGLS=0.0049±0.0002). Analogue $[T(GalNAc)AA]_4G$ exhibited the most potent RI of the analogues, comparable to that

of AFGP8 (0.0016±0.0001). In this study it is interesting to note that while the level of RI activity is much lower than the native system containing a disaccharide, significant RI can be achieved with a monosaccharide.

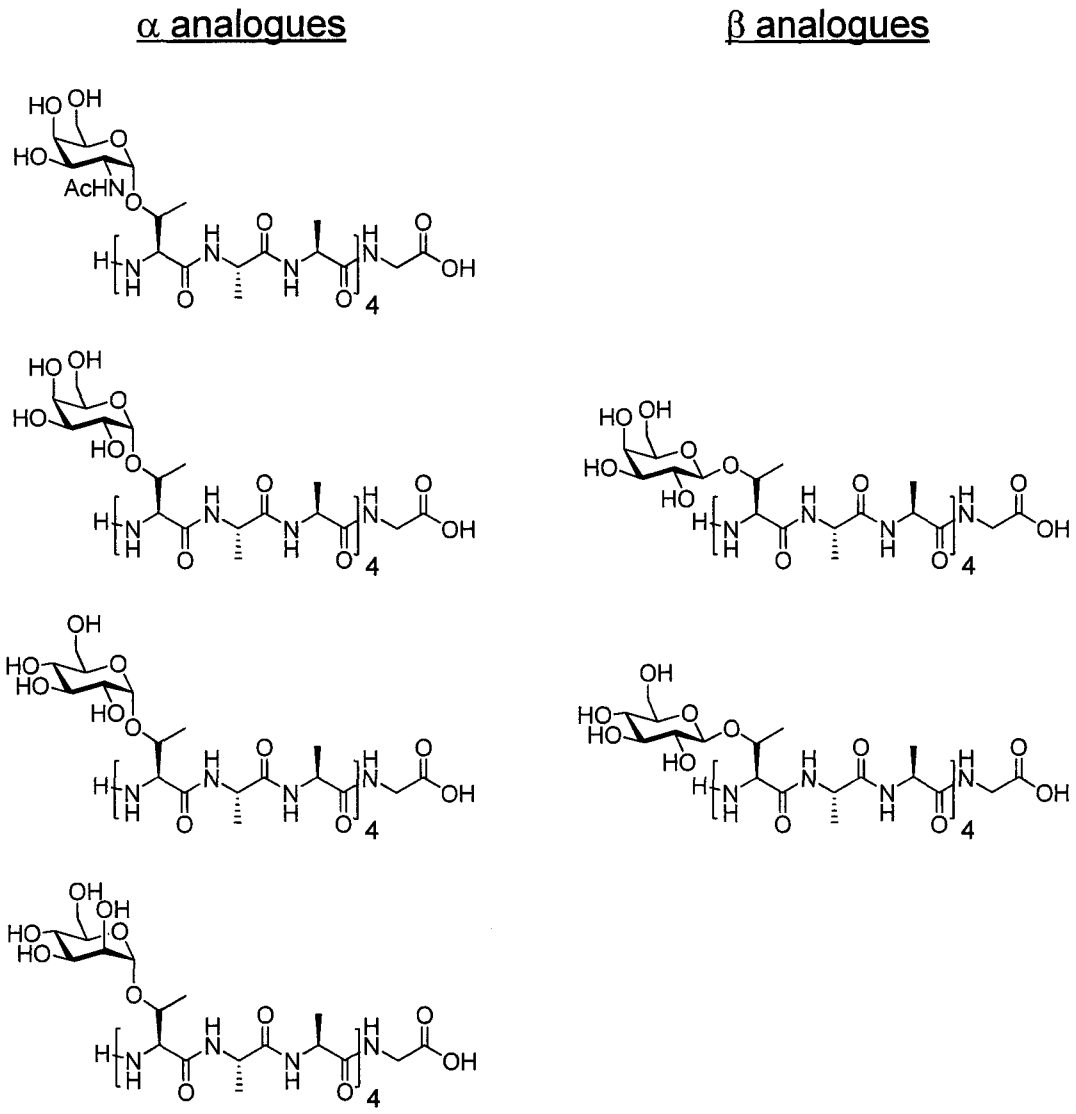


Figure 2.3. A library of synthesized O-linked [Thr(monosaccharide)-Ala-Ala]$_4$Gly analogues.[9a]

In summary, this class of analogues demonstrates that O-linked monosaccharide AFGP analogues bearing a galactose residue exhibit weak interactions with the ice lattice and also possess moderate RI activity.

2.1.2. Design of the C-linked AFGP Analogues

2.1.2.1. The Side Chain

The Ben laboratory has also studied many different C-linked AFGP analogues during the last eight years. In these analogues, the C-linked pyranose was attached to the polypeptide backbone via an amide bond with a lysine or ornithine residue. This structural modification originated from the fact that in low molecular mass AFGPs threonine residues are occasionally replaced by arginine.[10] Substitution of the amino acids has no significant loss of the antifreeze activity. The guanidine structure of the arginine side chain was simulated by the amide bond (Figure 2.4).[11]

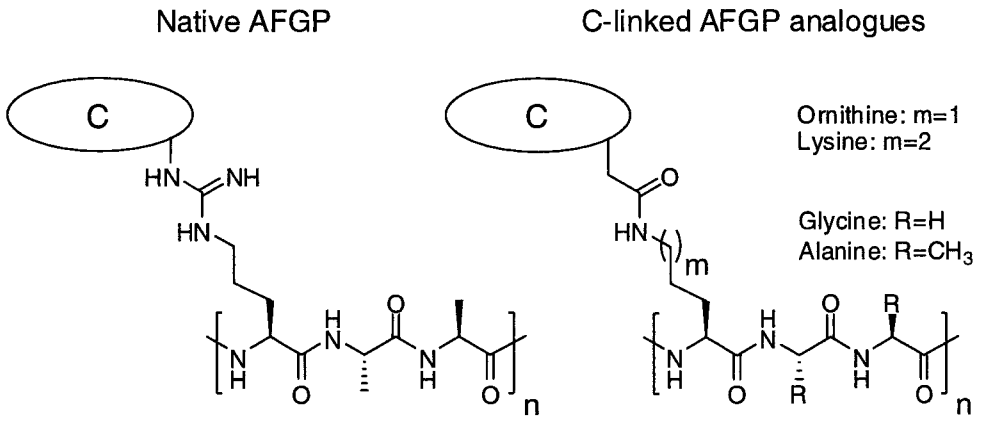


Figure 2.4. Rational design of AFGP analogues: mimicking the arginine residue of the native AFGPs with an amide bond synthesized from ornithine or lysine.

In addition, C-linked AFGP analogues that are structurally similar to a C-linked serine derivative have also been prepared and studied (Figure 2.5).[4, 12] In this series of analogues, the number of methylenes was varied from 2 to 4. The β-methyl group of the L-theronine residue has been reported to be an important feature for the TH activity in native AFGP.[6] However, the lack of this β-methyl group gives the analogues more freedom of rotation as compared to AFGP8 and our previously designed C-linked series which include the rigid amide bond. These two general classes of analogues form the platform for subsequent modifications of the carbohydrate and polypeptide backbone of all the C-linked AFGP analogues prepared in our laboratory.

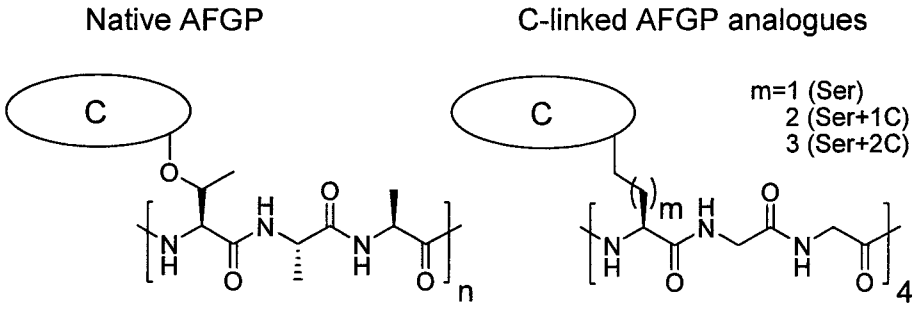


Figure 2.5. Rational design of AFGP analogues: mimicking the threonine residue of the native AFGPs with a C-linked serine series.[4, 12]

2.1.2.2. The Carbohydrate

The RI results generated from the O-linked AFGP analogue library (Figure 2.3) verified that the disaccharide can be replaced by monosaccharides without significant loss of activity.[9a] It was therefore reasonable to implement the same changes in our C-linked AFGP analogue series. Consequently, N-acetyl-D-galactosamine, D-galactose, D-glucose and D-mannose were chosen as the monosaccharides in our C-linked AFGP mimetics.[4, 9, 11-13]

In addition, a divalent derivative (Figure 2.6) was also prepared to investigate the effect of carbohydrate concentration on antifreeze protein-specific activity.[9b] This analogue possesses a total of eight carbohydrate residues with two per L-lysine.

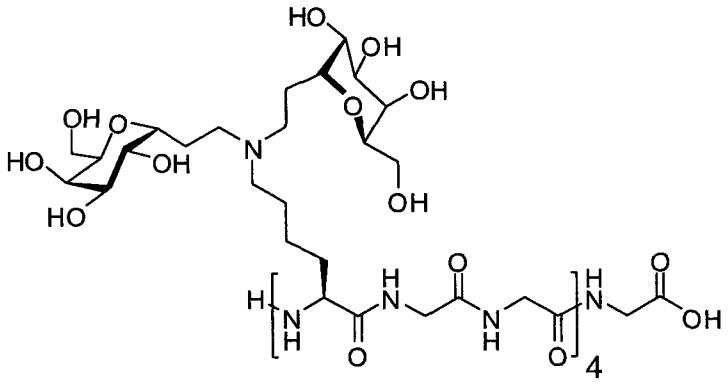


Figure 2.6. Divalent C-linked analogue [Lys(Gal_2)-Gly-Gly]$_4$Gly.[9b]

2.1.2.3. The Peptide Backbone

The final component to be altered in our C-linked AFGP analogues was the peptide backbone. L-Alanine residues were a logical choice as they are readily present in the structure of the native AFGPs. Furthermore, methyl groups of the alanines were proved to be important in hydrophobic interactions and binding to ice crystal lattice. Secondly, L-proline was incorporated into the polypeptide of the C-linked AFGP analogues. In the structure of the native AFGPs, it has been found that the first L-alanine after the L-threonine was occasionally replaced by L-proline.[14]

In addition to altering the amino acids, glycoproteins with varied peptide length were designed in order to determine their effect on antifreeze activity.[9b, c] It is well known that the higher molecular weight AFGPs possess a more potent antifreeze activity.[6] The antifreeze

activity is defined in terms of RI as well as TH and the mechanisms by which they act are thought to be uncoupled,[15] therefore the varying glycopeptide length may have different effect on the two properties. Ultimately, optimal glycopeptide length will be determined for potent recrystalliation inhibitors.

2.1.3. Insight into the SAR of C-linked AFGP Analogues

The following paragraphs in this section will discuss trends and results of the synthesized C-linked AFGP analogues summarized in Table 2.1.[4, 9, 11-13]

The first subset of C-linked AFGP analogues synthesized incorporated two L-glycine residues and one L-lysine residue bound to a C-linked galactose (Entries 1-3, Table 2.1) The resulting glycopolymers, $[Lys(Gal)\text{-}Gly\text{-}Gly]_nGly$, varied in length from n=3 to n=9 units. To rule out any non-specific effects on their activity the carbohydrate concentration was kept constant. Although $[Lys(Gal)\text{-}Gly\text{-}Gly]_9Gly$ was tested with peptide concentration three times lower then $[Lys(Gal)\text{-}Gly\text{-}Gly]_3Gly$ it still exhibited higher RI activity. The results verify that a direct correlation between molecular weight and RI activity exists.

To further investigate the importance of the carbohydrate concentration, a divalent analogue (Entry 4) was synthesized and tested for antifreeze activity. As indicated in Table 2.1, the divalent analogue $[Lys(Gal_2)\text{-}Gly\text{-}Gly]_4Gly$ had RI activity comparable to that of $[Lys(Gal)\text{-}Gly\text{-}Gly]_9Gly$ suggesting that the carbohydrate concentration with respect to the peptide backbone is very important for RI activity. In essence, doubling the carbohydrate concentration, results in a two-fold increase in RI activity.

Table 2.1. Dynamic Ice Shaping, Thermal Hysteresis and Recrystallization Inhibition of C-linked AFGP analogues previously synthesized by Ben et al. [4, 9, 11-13]

Entry	Analogues	n	DIS	TH	RI[(a)]
1	$[Lys(Gal)\text{-}Gly\text{-}Gly]_nGly$	3	None	–	insignificant
2	$[Lys(Gal)\text{-}Gly\text{-}Gly]_nGly$	6	Hexagon	0.06°C	**
3	$[Lys(Gal)\text{-}Gly\text{-}Gly]_nGly$	9	Hexagon	0.06°C	***
4	$[Lys(Gal_2)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	***
5	$[Lys(Gal)\text{-}Gly\text{-}Gly]_nGly$	4	Hexagon	–	*
6	$[Lys(Glc)\text{-}Gly\text{-}Gly]_n$ Gly	4	None	–	*
7	$[Lys(Man)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	*
8	$[Lys(Gal)\text{-}Ala\text{-}Ala]_nGly$	4	None	–	insignificant
9	$[Lys(Gal)\text{-}Pro\text{-}Gly]_nGly$	4	None	–	***
10	$[Lys(Glc)\text{-}Pro\text{-}Gly]_nGly$	4	None	–	*
11	$[Lys(Man)\text{-}Pro\text{-}Gly]_nGly$	4	None	–	*
12	$[Orn(Gal)\text{-}Gly\text{-}Gly]_nGly$	4	Hexagon	–	*****
13	$[Orn(Glc)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	**
14	$[Orn(Man)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	*
15	$[Orn(GalNAc)\text{-}Ala\text{-}Ala]_nGly$	4	None	–	****
16	$[Ser(GalNAc)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	**
17	$[Ser(Gal)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	*****
18	$[(Ser+1C)(Gal)\text{-}Gly\text{-}Gly]_nGly$	4	None	–	**
19	$[(Ser+2C)(Gal)\text{-}Gly\text{-}Gly]_nGly$	4	Hexagonal	–	**

a: * - least potent RI (MGLS=0.010-0.014mm^2), ***** - most potent RI (MGLS=0.000-0.004), for **, *** and **** increments of 0.002mm^2 MGLS were used to verify RI potency.

It is interesting to note that TH activity also correlates to glycopeptides length. For instance, [Lys(Gal)-Gly-Gly]$_n$Gly, where n=6 and 9 (Entries 2 and 3), displayed TH, while [Lys(Gal$_2$)-Gly-Gly]$_4$Gly (Entry 4) did not, suggesting that there is a minimum length of the glycopolymer necessary to observe TH activity. To examine the length at which no TH will be observed yet the analogue will possess RI activity, [Lys(Gal)-Gly-Gly]$_4$Gly was synthesized (Entry 5). The number of tripeptide units was between that of 3 which showed no TH activity and that of 6 which exhibited TH. [Lys(Gal)-Gly-Gly]$_4$Gly possessed weak RI while TH was not detected, however hexagonal DIS was observed. Since the analogue's length was the same as that of AFGP8, a length of 4 units seemed reasonable for the design of subsequent C-linked AFGP analogues.

To improve RI of [Lys(monnosaccharide)-Gly-Gly]$_4$Gly analogues and elucidate the importance of the hydroxyl group stereochemistry, a series of analogues with different carbohydrates was synthesized (Entries 6-8). None of the two analogues showed any significant improvement in RI activity, yet both had no TH or DIS. Modifying the peptide backbone so it more resembles the native system and synthesizing [Lys(Gal)-Ala-Ala]$_4$Gly did not bring any further success.

Subsequent to this, other modifications were made to the polypeptide backbone and L-proline residues were incorporated to furnish C-linked AFGP analogues [Lys(Gal)-Pro-Gly]$_4$Gly, [Lys(Glc)-Pro-Gly]$_4$Gly, [Lys(Man)-Pro-Gly]$_4$Gly (Entries 9-11). It was expected that conformational changes in the solution conformation of these analogues might result, and that this may affect antifreeze activity. However, no significant changes were observed in [Lys(Glc)-Pro-Gly]$_4$Gly and [Lys(Man)-Pro-Gly]$_4$Gly. Conversely, [Lys(Gal)-Pro-Gly]$_4$Gly exhibited a significant increase in RI with no evidence of DIS or TH. The most significant result for this

subset of analogues was that the relative stereochemistry of the carbohydrate moiety was a significant factor in activity and that galactose appeared to be the "best" carbohydrate.

To fully investigate this relationship, C-linked AFGP analogues possessing different carbohydrate moieties with the same backbone (Entries 12-14) were prepared. As expected, [Orn(Gal)-Gly-Gly]$_4$Gly exhibited potent RI activity, comparable to that of the natural AFGPs. The analogue showed hexagonal DIS however no TH was observed. RI for [Orn(Glc)-Gly-Gly]$_4$Gly was moderate, while for [Orn(Man)-Gly-Gly]$_4$Gly it was weak, with no TH activity detected for both analogues.

The study of the influence of N-acetyl-D-galactosamine on RI activity is still incomplete but [Orn(GalNAc)-Ala-Ala]$_n$Gly (Entry 15) shows promising results and suggesting that the C2 N-acetyl group may be important. Using the second C-serine based platform for our C-AFGP analogues, a series of compounds were prepared to further investigate the importance of the C2 NHAc group (Entries 16-19). The number of carbon-carbon bonds of the "serine" was varied from 3 to 5, and the monosaccharides used were D-galactose and N-acetyl-D-galactosamine. The galactose analogue, [Ser(Gal)-Gly-Gly]$_4$Gly, showed superior RI activity compared to any previously synthesized C-AFGP analogues. The results suggest that incorporation of the N-acetyl moiety does not result in improved RI in this series of analogues.

In summary, simple C-linked AFGP analogues were designed by modifying the carbohydrate component, side chain, and the peptide backbone. The synthesized mimetics were tested for antifreeze activity and the best analogues, C-linked [Orn(Gal)-Gly-Gly]$_4$Gly and [Ser(Gal)-Gly-Gly]$_4$Gly, exhibited no TH and RI activity comparable to that of the native AFGP8. Moreover, the C-linked AFGP analogues showed superior stability with custom-tailored physical antifreeze activity. This is an important result as it suggests the rational design

of novel AFGP molecules suitable for medical, commercial and industrial applications is feasible. However, before this goal can be realized, additional SAR work must be performed.

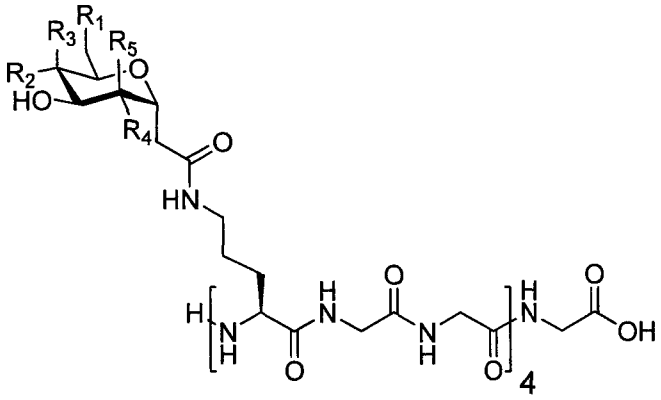
2.2. Research Goals and Objectives

Previously developed AFGP analogues gave important insights into the future design of novel cryoprotecting agents. Their selective ability to inhibit ice recrystallization without strong interaction with ice lattice (no TH) makes them excellent candidates for protection of cells at subzero temperatures. It is well known that native AFGPs form ice spicules at temperatures below the TH gap.[6, 7] The sharp ice crystals can damage the cells leading to their death, making TH activity an undesirable property for a cryoprotectant.[16] Thus one of the goals of our research program is to inhibit ice crystal growth with suitable and non-toxic recrystallization inhibitors.

The structural modifications made in previously synthesized C-linked AFGP analogues addressed the importance of the side chain, polypeptide backbone but especially the carbohydrate. The RI activity was shown to be sensitive to changes in the stereochemistry of the monosaccharides of the C-linked AFGP mimetics. While analogues containing D-galactose were determined to be the most potent inhibitors of ice recrystallization, to date no model has been proposed that can explain this effect. Consequently, the objective of this dissertation is to design a new generation of C-linked analogues to further probe SAR trends, and help to better understand how carbohydrate stereochemistry relates to RI activity. Perhaps these SAR studies will further elucidate the key structural features necessary for potent inhibitors of ice recrystallization. The goals of my research are:

Goal (1) – *To investigate the influence of carbohydrate stereochemistry on RI activity and determine whether it is related to hydration of the carbohydrate moiety.*

It is well accepted that hydration of proteins plays an important role in modulating protein function.[17] Similarly, hydration or solvation of carbohydrates and oligosaccharides tempers their biological activity.[18] Subsequently, Galema et al. studied the hydration of many commercially available hexoses and correlated it to carbohydrate stereochemistry.[19] To test whether the hydration has an effect on RI activity and it can be related to carbohydrate configuration, the C-linked [Orn(carbohydrate)-Gly-Gly]$_4$Gly series will be expanded to incorporate D-talose, D-lactose and D-melibiose (Figure 2.7). If there is a true relationship between hydration and RI, it will allow one to predict the influence of the carbohydrate on RI activity and revolutionize the future design of AFGP analogues.



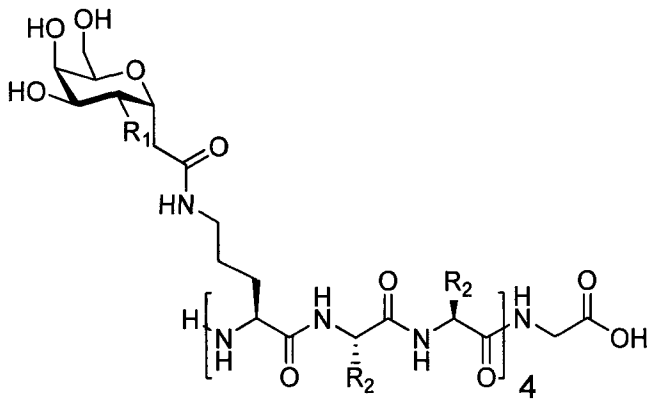
[Orn(Tal)-Gly-Gly]$_4$Gly: R_1=OH, R_2=H, R_3=OH, R_4=H, R_5=OH
[Orn(Lactose)-Gly-Gly]$_4$Gly: R_1=OH, R_2=β-D-galactopyranosyl-(1-4)-, R_3=H, R_4=OH, R_5=H
[Orn(Melibiose)-Gly-Gly]$_4$Gly: R_1=α-D-galctopyranosyl-(1-6)-, R_2=OH, R_3=H, R_4=OH, R_5=H

Figure 2.7. Target C-linked [Orn(Tal)Gly-Gly]$_4$Gly, [Orn(Lactose)-Gly-Gly]$_4$Gly and [Orn(Melibiose)-Gly-Gly]$_4$Gly analogues.

Goal (2) – *To assess the importance of the C-2 N-acetyl moiety on RI activity and its potential effect on solution conformation.*

While the role of the C-2 acetamide group has been already addressed, more studies need to be devoted to understand its function in more detail. Studies by Nishimura suggested that the methyls of the C-2 acetamide are important in forming the hydrophobic portion of the AFGPs.[6] In addition, NMR studies indicate that the nitrogen proton of the amide forms hydrogen bonds with the carbonyl residues of the peptide backbone.[20, 21] This hydrogen bonding is strong enough to influence the conformation of the glycopeptide. The C-2 N-acetyl group has proven to be an important motif for TH,[6] however, no reports have been found regarding its effect on RI. In our C-linked [Ser(monosaccharide)-Gly-Gly]$_4$Gly analogues N-acetyl-D-galactosamie showed a decrease in RI over its counterpart containing D-galactose. On the contrary, [Orn(GalNAc)-Ala-Ala]$_4$Gly analogue exhibited a fairly potent RI activity.

In order to complete the investigation of the effect of C-2 acetamide group on RI, [Orn(GalNAc)-Gly-Gly]$_4$Gly and [Orn(Gal)-Ala-Ala]$_4$Gly were designed (Figure 2.8). Both analogues will be synthesized and their antifreeze activities will be compared to that of [Orn(GalNAc)-Ala-Ala]$_4$Gly and [Orn(Gal)-Gly-Gly]$_4$Gly.



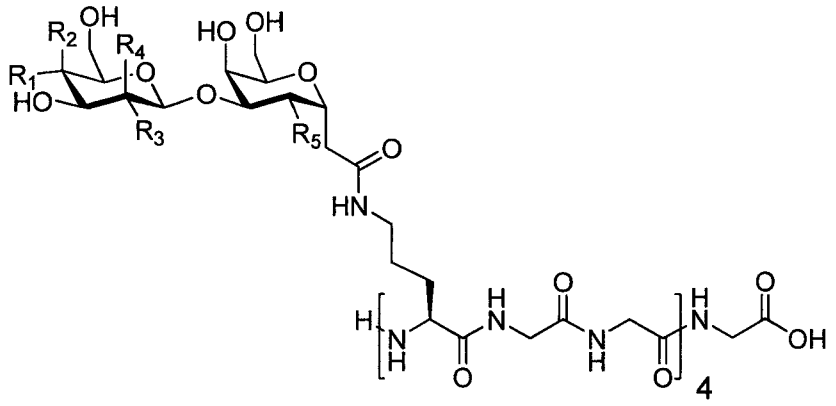
[Orn(GalNAc)-Gly-Gly]$_4$Gly: R_1=NHAc, R_2=H
[Orn(Gal)-Ala-Ala]$_4$Gly: R_1=OH, R_2=CH_3

Figure 2.8. Target C-linked [Orn(GalNAc)Gly-Gly]$_4$Gly and [Orn(Gal)-Ala-Ala]$_4$Gly analogues.

Goal (3) – *To examine the effect of the C-linked AFGP disaccharide derivatives on hydration and RI activity.*

One of the major structural changes implemented in the synthesis of the C-linked AFGP analogues was the change of the disaccharide to the monosaccharide. As evident in a series of our O-linked analogues and the native AFGP8, the β-D-Gal-(1-->3)-D-GalNAc plays a key role in TH activity.[6, 9a] The disaccharide mode of action is still unknown and only speculations have been made that (a) its configuration is important for TH and (b) that it plays a specific role in setting up the conformation of the glycopeptide essential for ice binding and as a result TH activity. AFGP8 also exhibits a very potent RI activity which seems to be also attributed to the disaccharide as implied by the data from our O-linked series. Since no disaccharide bearing C-linked analogues have been synthesized, the above hypothesis has not been verified. However, it can be suggested that if RI is concentration dependent, as observed by comparing the C-linked

[Lys(Gal_2)-Gly-Gly]$_4$Gly and [Lys(Gal)-Gly-Gly]$_4$Gly, then AFGP8 RI strength relies on the presence of the β-D-Gal-(1-->3)-D-GalNAc disaccharide. In order to address the importance of the disaccharide on antifreeze activity C-linked [Orn(disaccharide)-Gly-Gly]$_4$Gly series have been designed (Figure 2.9).



[Orn(Gal-GalNAc)-Gly-Gly]$_4$Gly: R_1=H, R_2=OH, R_3=OH, R_4=H, R_5=NHAc
[Orn(Glc-GalNAc)-Gly-Gly]$_4$Gly: R_1=OH, R_2=H, R_3=OH, R_4=H, R_5=NHAc
[Orn(Man-GalNAc)-Gly-Gly]$_4$Gly: R_1=OH, R_2=H, R_3=H, R_4=OH, R_5=NHAc
[Orn(Gal-Gal)-Gly-Gly]$_4$Gly: R_1=H, R_2=OH, R_3=OH, R_4=H, R_5=OH

Figure 2.9. Target C-linked [Orn(Gal-GalNAc)Gly-Gly]$_4$Gly, [Orn(Glc-GalNAc)-Gly-Gly]$_4$Gly, [Orn(Man-GalNAc)-Gly-Gly]$_4$Gly and [Orn(Gal-Gal)Gly-Gly]$_4$Gly analogues.

2.3. References

1. (a) Maiwand, M. O., Asimakopoulos, G., *Technol. Cancer Res. Treat.* **2004**, 3, 151. (b) Kuflik, E. G., *Dermatol. Surg.* **2004**, 30, 297. (c) Wang, J. H., *Cryobiology* **2000**, 41, 1. (d) Tablin, F., Oliver, A. E., Walker, N. J., Crowe, L. M., Crowe, J. H., J. *Cell. Physiol.* **1996**, 168, 305. (e) Griffith, M., Ewart, K. V., *Biotech. Adv.* **1995**, 13, 375. (f) Payne, S. R., Sandford, D., Harris, A., Young, O. A., *Meat Sci.* **1994**, 37, 429. (g) Payne, S. R., Young, O. A., *Meat Sci.* **1995**, 41, 147.
2. (a) Rubinsky, B., Arav, A., DeVries, A. L., *Cryobiology* **1992**, 29, 69. (b) Storey, K. B., Bischof, J., Rubinsky, B., *Am. J. Physiol.* **1992**, 263, R185.
3. (a) Marcaurelle, L. A., Bertozzi, C. R., *Chem. Eur. J.* **1999**, 5, 1384. (b) Boons, G. J., Hale, K. J., Organic Synthesis with Carbohydrates, Blackwell Science, Malden, USA, **2000**.
4. Liu, S., *Ph.D. Thesis*, University of Ottawa, Ottawa, **2006**.
5. (a) Tsuda, T., Nishimura, S. I., *Chem. Commun.* **1996**, 24, 2779. (b) Matsushita, T., Hunou, H., Kurogochi, M., Shimizu, H., Nishimura, S. I., *Org. Lett.* **2005**, 7, 877. (c) Tachibana, Y., Monde, K., Nishimura, S. I., *Marcomolecules* **2004**, 37, 6771. (d) Tachibana, Y., Matsubara, N., Nakajima, F., Tsuda, T., Tsuda, S., Monde, K., Nishimura, S. I., *Tetrahedron. Lett.* **2002**, 58, 10213. (e) Tseng, P. H., Jiiang, W. T., Chang, M. Y., Chen, S. T., *Chem. Eur. J.* **2001**, 7, 585.
6. Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishmura, S., *Angew. Chem. Int. Ed.*, **2004**, 43, 856.
7. Komatsu, S. K., DeVries, A. L., Feeney, R. E., *J. Biol. Chem.* **1970**, 245, 2909.

8. (a) Sparks, M. A., Williams, K. W., Whitesides, G. M., *J. Med. Chem.* **1993**, 36, 778. (b) Kutterer, K. M. K., Barnes, M. L., Arya, P. *J. Comb. Chem.* **1999**, 1, 28.

9. (a) Bouvet, V., *Ph.D. Thesis*, University of Ottawa, Ottawa, **2006**. (b) Murphy, A., *Ph.D. Thesis*, State University of Binghamtom, Binghamton, **2004**. (c) Eniade, A., Purushotham, M., Ben, R. N., Wang, J. B., Horwath, K., *Cell Biochem. Biophys.* **2003**, 38, 115.

10. (a) Morris, H. R., Thompson, M. R., Osuga, D. T., Ahmed, A. I., Chan, S. M., Vandenheede, J. R., Feeney, R. E. *J. Biol. Chem.* **1978**, 253, 5155. (b) Hew, C. L., Slaughter, D., Fletcher, G., Shashikant, J. B., *Can. J. Zool.* **1981**, 59, 2186. (c) Raymond, J. A., Lin, Y., DeVries, A. L., *J. Exp. Zool.* **1975**, 193, 125. (d) Harding, M. M., Anderberg, P. L., Haymet, A. D. J. *Eur. J. Biochem.* **2003**, 270, 1381.

11. Ben, R. N., Eniade, A., Hauer, L., *Org. Lett.* **1999**, 1, 1759.

12. Liu, S., Ben, R. N., *Org. Lett.* **2005**, 7, 2385.

13. (a) Eniade, A., Ben, R. N., *Biomacromolecules* **2001**, 2, 557. (b) Eniade, A., Murphy A. V., Landreau, G., Ben, R. N., *Bioconjgate Chem.* **2001**, 12, 817.

14. (a) Geoghegan, K. F., Osuga, D. T., Ahmed, A. I., Yeh, Y, Feeney, R. E., *J. Biol. Chem.* **1980**, 255, 663. (b) O'Grady, S. M., Schrag, J. D., Raymond, J. A., DeVries, A. L., *J. Exp. Zool.* **1982**, 224, 177. (c) Fletcher, G. L., Hew, C. L., Joshi, S. B., *Can. J. Zool.* **1982**, 60, 348. (d) Bucham, T. S., Osuga, D. T., Rao, B. N. N., Bush, C. A., Feeney, R. E., *J. Biol. Chem.* **1986**, 261, 6384.

15. (a) Horwath K. L., Easton C. M., Poggioli Jr., G. J. P., Myers K., Schnorr I. L. *Eur. J. Entomol.* **1996**, 93, 419. (b) Sidebottom, C., Buckley, S., Pudney, P., Twigg, S., Jarman,

C., Holt, C., Telford, J., McArthur, A., Worrall, D., Hubbard, R., Lillford, P., *Nature* **2000**, 406, 256.

16. (a) Koushafar, H., Rubinsky, B., *Urology* **1997**, 49, 421. (b) Pham, L., Dahiya, R., Rubinsky, B., *Cryosurgery* **1999**, 38, 169. (c) Wang, J. H., *Cryobiology* **2000**, 41, 1.

17. Frauenfelder, H., Fenimore, P. W., McMahon, B. H., *Biophys. Chem.* **2002**, 98, 35-48.

18. Franks, F., *Pure Appl. Chem.* **1987**, 59, 1189.

19. (a) Galema, S. A., Høiland, H., *J. Phys. Chem.* **1991**, 95, 5321. (b) Galema, S.A., Engberts, J. B. F. N., Høiland, H., Førland, G.M. *J. Phys.Chem.* **1993**, 97, 6885. (c) Galema, S. A., Howard, E., Engberts, J. B. F. N.; Grigera, J. R. *Carbohydr. Res.* **1994**, 265, 215.

20. Mimura, Y., Yamamoto, Y., Inoue, Y., Chujo, R., *Int. J. Biol. Macromol.* **1992**, *14*, 242.

21. Tam, R. Y., Rowley, C. N., Petrov, I., Zang, T., Afagh, N. A., Woo, T., Ben, R. N., unpublished results.

CHAPTER 3

Elucidating the Relationship between Carbohydrate Stereochemistry, Hydration and Recrystallization Inhibition

3.1. Introduction

The importance of hydration or solvation of proteins, peptides, oligosaccharides and various glycoconjugates has been well recognized.[1] However, determining the degree of hydration of a particular molecule is not a trivial matter. While recent advances in nuclear magnetic resonance (NMR) and other spectroscopic techniques aid in quantifying hydration, hydration effects are still overlooked when examining the fundamental nature of many biological processes.[2, 3]

Antifreeze glycoproteins (AFGPs) are found in many species of deep sea Teleost fish inhabiting sub-zero environments. They protect organisms from cryo-injury and death by inhibiting the growth of seeded ice crystals *in vivo*.[4] These compounds have attracted a great deal of interest both fundamentally and commercially over the last forty years and have been shrouded in controversy. For instance, the mechanism of action by which protection is conferred is regarded as an adsorption-inhibition process where the binding of AFGP to the ice surface results in a localized freezing point depression according to the Kelvin (Gibbs-Thompson) relation.[5] The binding event has been thought to be irreversible, but recent results have questioned this.[6, 7] Local energy transfer (or energy coupling) has complimented the irreversible nature of antifreeze binding to ice and has begun to link the mechanism of action of AFPs to AFGPs.[5, 8] The structural moieties responsible for binding to ice remain unknown but current hypotheses either invoke hydrophilic[9] or hydrophobic interactions.[10] Further controversy arises

from the fact that AFGPs display two modes of antifreeze activity: thermal hysteresis (TH) and recrystallization-inhibition (RI). While the mechanism of TH is reasonably well-understood,[4, 11] the manner in which AFGPs function as potent inhibitors of recrystallization is not.

The relevance of AFGP hydration and its importance at the ice-protein interface has not been directly addressed. However, hydration has been implicated as an important factor preventing the dimerization of individual AFP molecules, thereby ensuring that molecules function independently at the ice protein interface.[5] Similarly, previous conformational analyses of AFGP8 and AFGP1-5[8] have suggested these molecules are highly hydrated. In contrast, AFGP8 (the lowest molecular mass fraction of AFGP) has been shown to form higher order aggregates in solution,[12] but the issue of solvation was not examined despite the fact that this aggregation results in increased thermal hysteresis (TH) activity and likely has important implications *in vivo* as AFGP8 is the only AFGP to be found in tissues outside the bloodstream. AFGP8 has also been shown to function cooperatively as an inhibitor of recrystallization[8] but again, the question of solvation was not addressed.

As part of our on-going studies into the rational design of *C*-linked antifreeze glycoprotein analogues possessing custom-tailored antifreeze activity, we postulate that carbohydrate hydration might be an important factor in the activity of our analogues. To investigate this hypothesis, we prepared a series of *C*-linked AFGP analogues incorporating D-galactose, D-mannose, D-glucose and D-talose hexoses. The structure of these analogues was based upon *C*-linked AFGP analogues synthesized previously in our laboratory that demonstrated significant RI activity with little or no TH activity.[13, 14]

3.2. Glycopeptide Synthesis

The synthesis of glycoconjugates has received a lot of attention in the past decades due to the advances made in the field of glycobiology.[15] Many synthetic methods were developed to overcome the challenges associated with the glycopeptide synthesis and to maximize its efficiency and precision.[16] While the search for more reliable and facile protocols is still underway the main two methods utilized up to date are: solution phase peptide synthesis, and more common solid phase peptide synthesis (SPPS).[16a-e] Solution phase synthesis is based on protection of the amino acids' carboxyl terminus with either acid and/or base resistant protecting group. The N-terminus is then selectively deprotected and coupled to the free carboxylic group of the subsequent amino acid. The cycle of deprotecting/coupling continues until the desired sequence and length of the peptide is achieved. The main disadvantages associated with this method are its low efficiency, and lengthy work-up as well as purification procedures. In longer peptide syntheses the repetition of isolation and purification of the peptide intermediates can become labor intensive. This protocol has thus evolved and solid phase peptide synthesis has been developed. The method employs solid phase support on which the chemical transformations such as amide bond ligation, post-synthetic modifications, work-up and purification are performed. The method was first developed and pioneered by Bruce Merrifield in synthesis of polypeptides and earned him a Nobel Prize in 1984.[17]

The widely used SPPS utilizes either of the two common protocols: Boc (tert-butyloxocarbonyl) or Fmoc (9-fluorenyl-methoxycarbonyl) chemistry.[17a, 18] In the Boc procedure, to remove the tert-butyloxocarbonyl protecting group and obtain a free amine ready for the next coupling, strong acids such as trifluoroacetic acid (TFA) have to be used. Moreover, to cleave the peptide from the resin at the end of the synthesis, the beads have to be incubated in

a highly dangerous hydrofluoric acid. Since HF cleavage has to be performed using specialized equipment, this method is generally disfavored. In the Fmoc strategy, the 9-fluorenyl-methoxycarbonyl protecting group on the N-terminus can be cleaved using a mild base (usually 20% piperidine in DMF). The end product is detached from the solid resin using TFA. The Fmoc protocol being safer and less problematic became the most routine technique used. For the aforementioned reasons, Fmoc protocol will be used in the synthesis of our antifreeze analogues.

The AFGP analogues can be synthesized using two synthetic strategies.[13] First one is the convergent synthesis in which the peptide backbone is first synthesized (Figure 3.1). Consequently, the glycosylation sites are deprotected and the carbohydrate moieties are coupled. The glycopeptide is then cleaved from the beads, fully deprotected and purified. The second strategy is the linear synthesis in which the amino acids are assembled stepwise (Figure 3.2). The amino acid bearing the carbohydrates or the building block unit is prepared first and subsequently employed in the SPPS sequence. After cleavage from the beads, the glycopeptide is fully deprotected and purified. Since the linear synthesis in our laboratory proved to be more efficient, this is the method of choice for the target analogues in this dissertation.

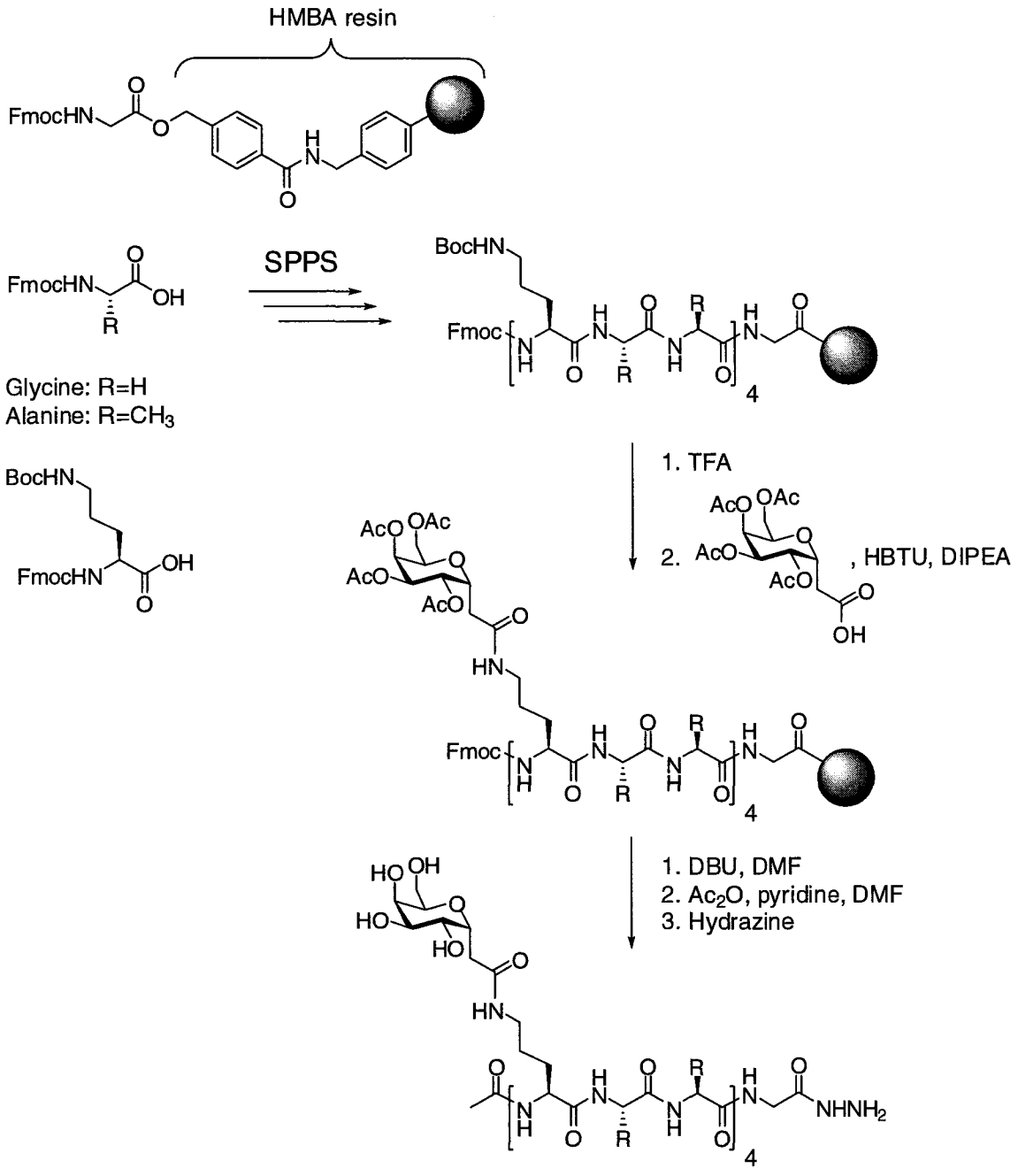


Figure 3.1. Convergent synthesis of the *C*-linked AFGP analogues.[13]

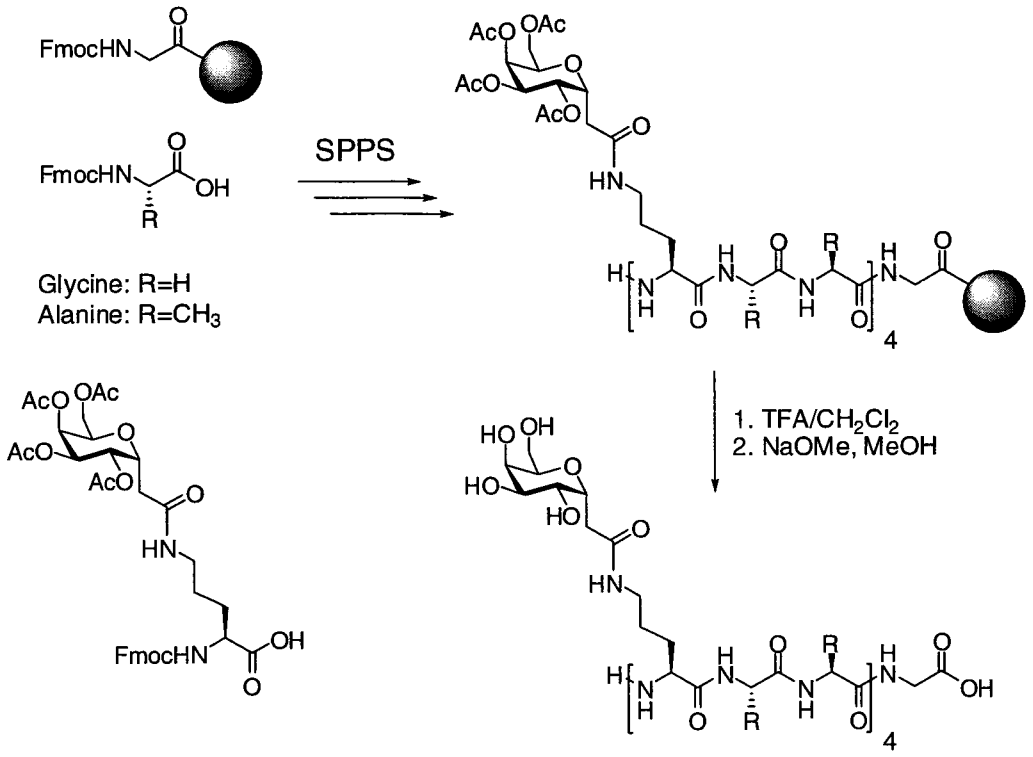
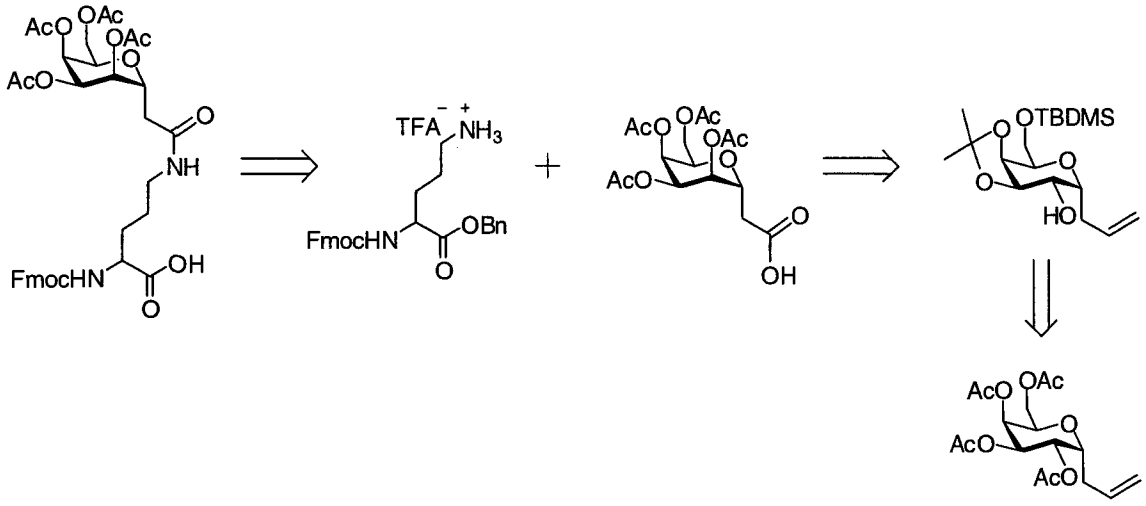


Figure 3.2. Linear synthesis of the *C*-linked AFGP analogues.[13]

3.3. Retrosynthetic Analysis of the *C*-linked [Orn(Tal)-Gly-Gly]$_4$Gly Analogue

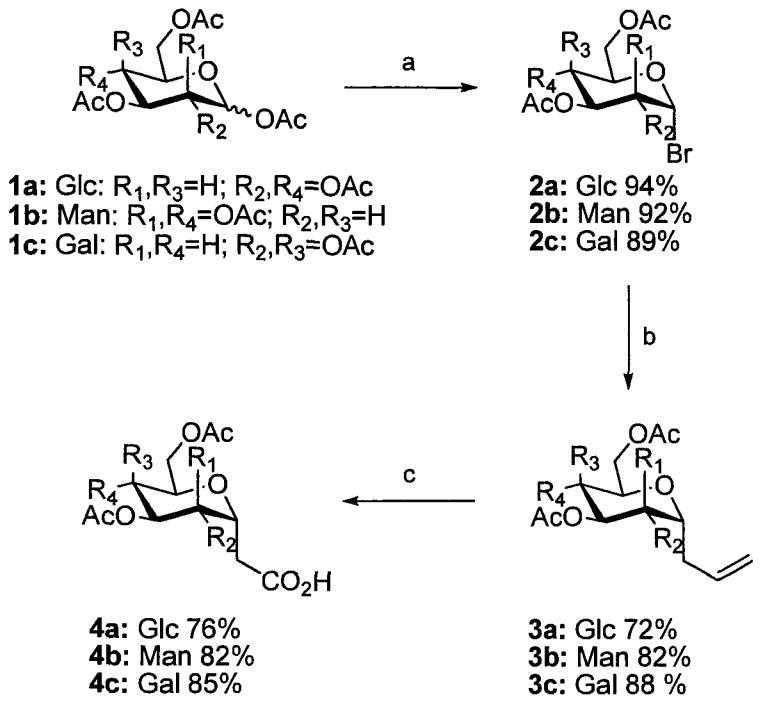
Preparation of the *C*-linked ornithine-talose analogue is important as we anticipate that it will exhibit no RI activity and verify the link between carbohydrate stereochemistry, hydration and RI activity. This building block can be synthesized by coupling a TFA salt of ornithine and a carboxylic acid derivative of talose (Scheme 3.1). The *C*-linked oxidized talose derivative can be obtained via: inversion of the C-2 hydroxyl of galactose, following deprotection of the acid labile protecting groups, subsequent acetylation of the free hydroxyls and oxidation of the allyl moiety.[19] The corresponding galactose derivative can be prepared by protecting group manipulation from allyl 2,3,4,6-tetra-O-acetyl-α-D-galactopyranoside.



Scheme 3.1. Retrosynthetic analysis of the *C*-linked talose-ornithine building block.

3.4. Synthesis of the *C*-linked AFGP Analogues 19-22

The four *C*-linked AFGP analogues, **19-22,** were prepared in a linear manner whereby the *C*-linked pyranose was coupled to an orthogonally protected amino acid residue and assembled into the desired glycopolymer using automated solid phase synthesis as outlined previously. The D-glucose, D-galactose and D-mannnose series were prepared by Ben et al. as outlined in Scheme 3.2 and Scheme 3.4, however, several of these compounds were resynthesized for this study.[14, 20, 21]



1a: Glc: R_1,R_3=H; R_2,R_4=OAc
1b: Man: R_1,R_4=OAc; R_2,R_3=H
1c: Gal: R_1,R_4=H; R_2,R_3=OAc

2a: Glc 94%
2b: Man 92%
2c: Gal 89%

3a: Glc 72%
3b: Man 82%
3c: Gal 88 %

4a: Glc 76%
4b: Man 82%
4c: Gal 85%

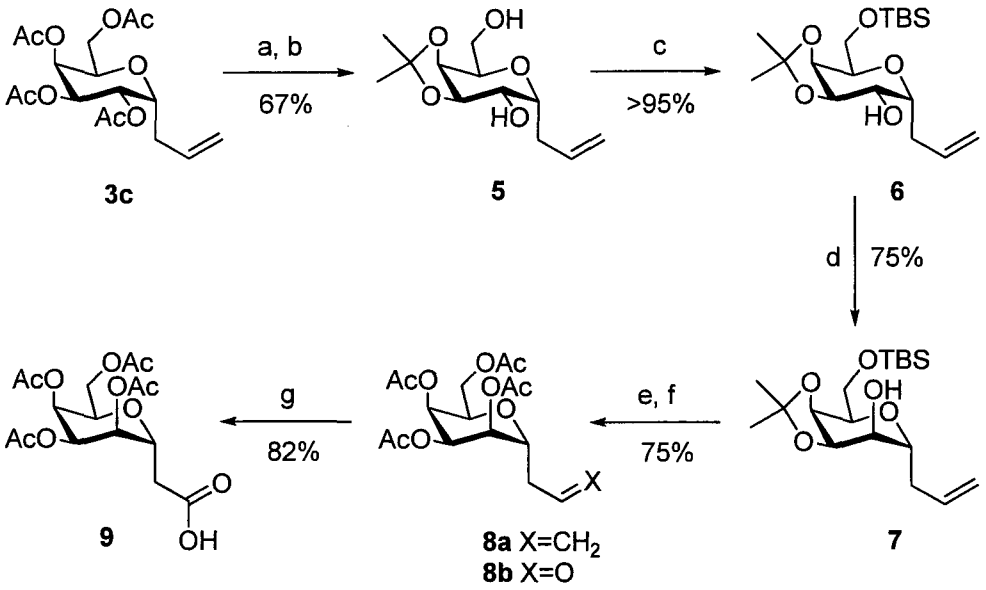
Conditions: (a) HBr/HOAc; (b) $(nBu)_6Sn_2$, allylphenylsulfone, benzene, hv; (c) $RuCl_3 \cdot H_2O$, $NaIO_4$, $CCl_4/CH_3CN/H_2O$

Scheme 3.2. Synthesis of the carbohydrate component of the *C*-linked AFGP analogues **19-22**.

Commercially available peracetylated pyranoses **1a-c** were converted to glycosyl bromides **2a-c** (89-94% yield) by treatment with hydrogen bromide in glacial acetic acid. Photochemical-mediated allylation furnished *C*-allylated pyranose derivatives **3a-c** possessing the α-configuration.[22] These derivatives were oxidized to produce the desired carboxylic acids **4a-c** in 76-85% yield.

The *C*-allylated D-talose derivative was prepared from *C*-linked D-galactose derivative **3c**. As illustrated in Scheme 3.3, the acetate protecting groups were removed using sodium in methanol solution. The C3 and C4 hydroxyl groups were selectively protected as a dimethyl acetal (70% yield) and the C6 hydroxyl group was protected as a TBS-ether (>95% yield).

Inversion of configuration at C2 was accomplished by a Swern Oxidation and subsequent reduction of the ketone with $NaBH_4$ to produce **7** in 75% yield (two steps).[19] Deprotection and acetylation followed by ozonolysis resulted in **8b** (75% yield, three steps), which was further oxidized to give **9** (82% yield).

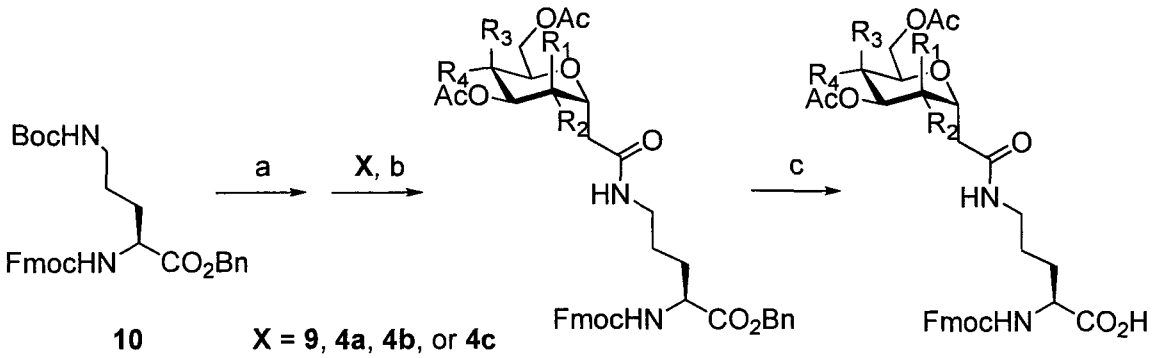


Conditions: (a) NaOMe, MeOH; (b) $(CH_3)_2C(OCH_3)_2$, *p*-TsOH, CH_3CN; (c) TBDMSCl, DMAP, py; (d) $(CO)_2Cl_2$, DMSO, Et_3N, then $NaBH_4$, CH_3OH; (e) 80% HOAc-H_2O, then Ac_2O, py-CH_2Cl_2; (f) O_3, PPh_3, CH_2Cl_2, -78°C; (g) $NaClO_2$, KH_2PO_4, 2-methyl-2-butene, ^tBuOH-^iPrOH-H_2O.

Scheme 3.3. Synthesis of the *C*-linked talose derivative **9**.

Compounds **4a-c 9** were then coupled to orthogonally protected ornithine derivative **10**[14] using O-benzotriazole-N,N,N',N'-tetramethyl-uronium-hexafluoro-phosphate (HBTU) and Hünig's base to afford benzyl protected building blocks **11-14** in 87-93% yields. The benzyl esters were carefully removed by chemoselective hydrogenolysis to furnish the requisite glycoconjugate building blocks **15-18** for preparation of the polymers (Scheme 3.4). The *C*-

linked AFGP analogues were prepared in an automated manner using an APEX 396 and standard Fmoc-based chemistry.[23] The glycopolymers were removed from the Wang resin by treatment with a 50% TFA solution in dicholormethane following ether precipitation. The acetate protecting groups on the carbohydrates were removed using Na/methanol solution. Purification by C-18 reversed-phase HPLC furnished the desired *C*-linked AFGP analogues **19-22** in yields ranging from 40-62%.

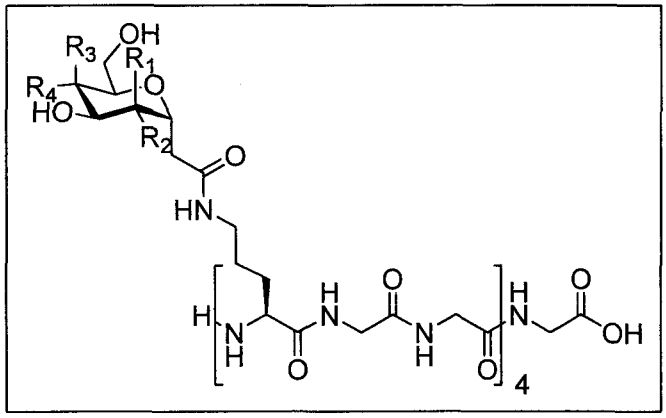


10 **X = 9, 4a, 4b, or 4c**

11: Gal: 83%
12: Glc: 85%
13: Man: 93%
14: Tal: 87%

15: Gal: 84%
16: Glc: 87%
17: Man: 82%
18: Tal: 85%

Automated SPPS



19: $[O(Gal)GG]_4G$: R_1=H, R_2=OH, R_3=OH, R_4=H: 55%
20: $[O(Glc)GG]_4G$: R_1=H, R_2=OH, R_3=H, R_4=OH: 43%
21: $[O(Man)GG]_4G$: R_1=OH, R_2=H, R_3=H, R_4=OH: 40%
22: $[O(Tal)GG]_4G$: R_1=OH, R_2=H, R_3=OH, R_4=H: 62%

Conditions: (a) TFA/CH_2Cl_2; (b) HBTU, DIPEA, CH_2Cl_2; (c) 10% Pd-C, H_2, EtOH-EtOAc.

Scheme 3.4. Synthesis of glycopolymers **19-22**.

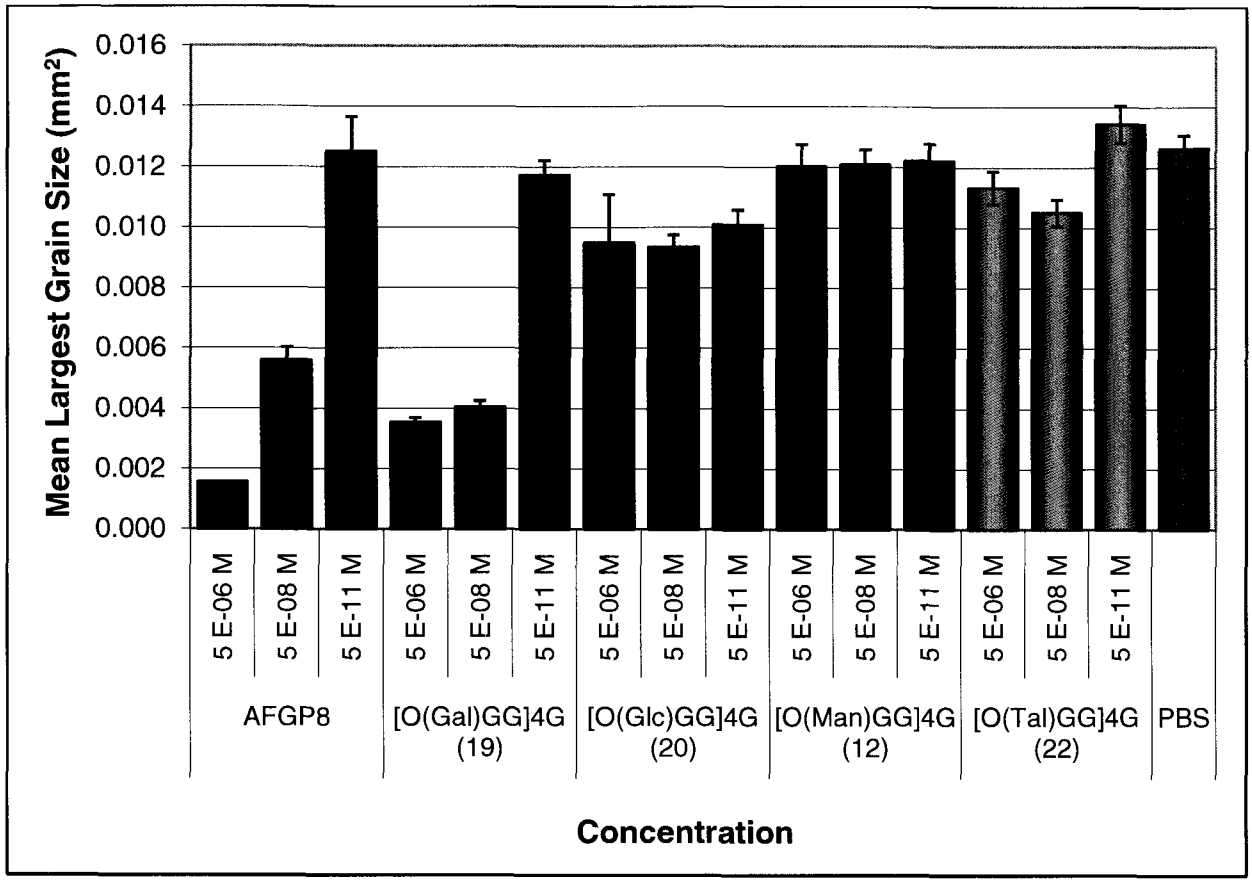
3.5. Results: Assessment of Antifreeze Activity and Conformational Analysis of *C*-linked AFGP Analogues 19-22

Thermal Hysteresis (TH):

C-Linked AFGP analogues **19-22** did not show any TH activity. However, analogue **19** did exhibit weak dynamic ice shaping indicating that this compound had the ability to interact with the ice lattice. This result was consistent with previous results obtained and outlined in Chapter 2.

Recrystallization-Inhibition (RI):

Analogues **19-22** were assessed for their ability to function as inhibitors of ice recrystallization (Graph 3.1). All samples were compared to a solution of phosphate buffered saline (PBS) which was used as a negative control and a solution of AFGP 8 isolated from *Gagus ogac* (generously provided by AquaBounty Farms). Each solution was tested at three different concentrations in order to rule out any non-specific RI effects.[24, 25] The size of ice crystals was analyzed using mean elliptical method. The Y-axis in Graph 3.1 represents the mean largest grain size (MLGS) where small bars are representative of potent recrystallization-inhibition activity. The D-talose and D-mannose analogues (**22** and **21**) did not show any RI activity and had MLGS values consistent with PBS, the negative control. D-Glucose analogue **20** exhibited only very moderate RI activity while D-galactose analogue **19** was the most potent analogue with the largest MLGS value of 0.00354 mm^2 at 5.54×10^{-6} M.



Graph 3.1. RI activity of *C*-Linked AFGP Analogues **19-22**.

The activity of analogue **19** is significantly greater than *C*-linked AFGP analogue containing L-lysine which has been previously reported by our laboratory.[13] This is despite the fact that the lysine analogue is comprised of six repeating tripepide units and not four. Interestingly, analogue **19** is not as active as the *C*-serine AFGP analogue previously reported by our laboratory.[25] Overall, analogue **19** possesses approximately one half the RI activity as native AFGP8.

Conformational Analysis:

The solution conformation of *C*-linked AFGP analogues **19-22** was examined by circular dichroism (CD) spectroscopy.[12] Deconvolution of the raw spectral data was accomplished using CD Pro with Selcon III and I-Basis 4 as a representative database.[26-28] The CD specta is presented in Figure 3.3.

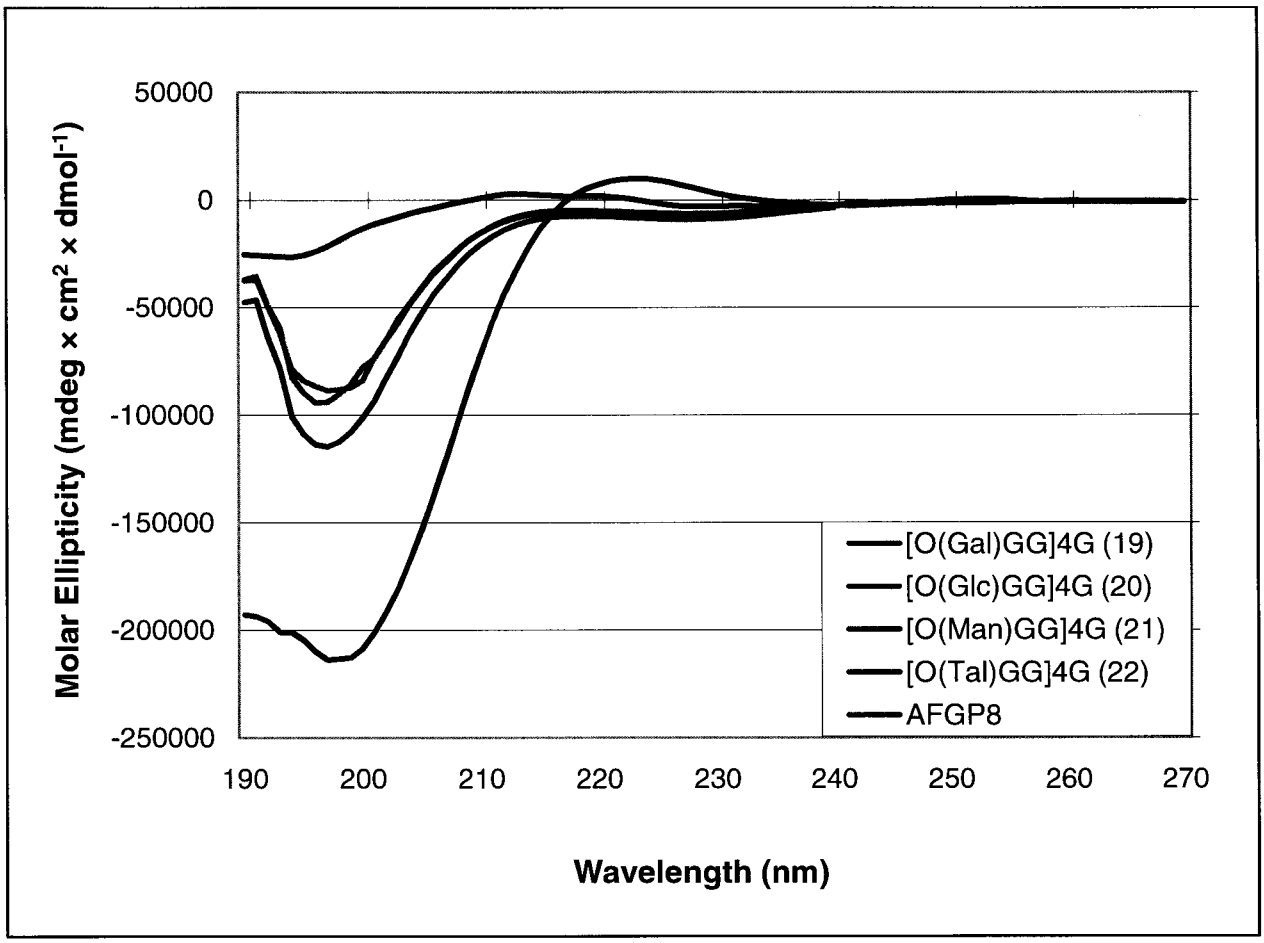


Figure 3.3. Circular dichroism spectra of analogues **19-22**.

Overall, the solution conformations of analogues **19-22** is consistent with a large percentage of random coil structure with varying degrees of other more ordered secondary structures (Table 3.1). These data correlate well with earlier conformational analyses of AFGP8 using CD spectroscopy.[12, 29]

Table 3.1. CD deconvolution data of AFGP analogues **19-22** and AFGP8 using CD Pro with Selcon III and IBASIS 4.

Compound	Regular α helix	Distorted α helix	Regular β strand	Distorted β strand	Turn	Random coil	Sum
$[O(Gal)GG]_4G$ (**19**)	-0.011	0.160	0.031	0.092	0.256	0.504	1.032
$[O(Glc)GG]_4G$ (**20**)	-0.012	0.172	-0.004	0.073	0.232	0.568	1.029
$[O(Man)GG]_4G$ (**21**)	0.010	0.164	0.008	0.007	0.225	0.540	0.954
$[O(Tal)GG]_4G$ (**22**)	0.020	0.049	0.084	0.101	0.248	0.529	1.031
AFGP8	0.049	0.114	0.164	0.094	0.237	0.378	1.036

3.6. Discussion: Interpretation of Antifreeze Activity and Conformational Analysis of Analogues 19-22

Graph 3.1 depicts an interesting result where the stereochemical relationship of the pyranose ring is clearly important for recrystallation-inhibition (RI) activity. The relative orientation between the OH groups at C2 and C4 appears to be crucial. For instance, when the C4-OH group is axial and the C2-OH group is equatorial (as in the galactose analogue **19**) the largest amount of RI activity is observed.

This result compliments previous structure-function studies which suggested that the cis-3,4 hydroxyl group in the terminal D-galactose residue[9] is essential for thermal hysteresis activity. Recent studies by Corzana et al.,[31] have suggested a correlation between the surrounding water shells of glycosylated serine and threonine residues with TH activity, however, identification of any structural features of AFGP8 for RI activity has yet to be elucidated.

Given the stereochemical sensitivity of the carbohydrate moiety to RI activity and the Corzona precedent,[31] we hypothesized that an explanation for the increased RI activity of analogue **19** relative to analogues **20**, **21** and **22** might lie in the hydration of these individual carbohydrate residues. An understanding of the structural characteristics essential to a potent inhibitor of recrystallization would greatly facilitate the rational design of novel cryoprotectants with custom-tailored activity. Such an approach is attractive as cell damage due to recrystallization is the single major cause of cell death during cryopreservation.[32, 33]

The hydration of carbohydrates has been studied extensively over the last three decades.[1] Various hypotheses have been proposed to rationalize hydration characteristics of carbohydrates and protein and the subsequent influence of hydration on bulk water. The "hydration layer" refers to water that encompasses the carbohydrate and/or protein and is often bound very tightly. In contrast, the region outside of this layer is very dynamic, consequently less ordered, and is referred to as "bulk water".[34] Hypotheses that rationalize hydration characteristics include: hydration number,[35-38] anomeric effect,[39] ratio of axial versus equatorial hydroxyl groups,[40, 41] hydrophobic index,[42] hydrophilic volume[43] and compatibility with bulk water based upon the position of the next-nearest-neighbour hydroxyl groups.[44, 45] Galema et al. have studied the hydration of various monosaccharides using molecular dynamics simulations, kinetic experiments, as well as density and ultrasound measurements.[46, 47] Consequently, the partial

molar volumes and isentropic partial molar compressibilities of many commercially available hexoses have been determined. In general, these values relate to the volume of space occupied by a molecule upon solvation by water. This data reveals that the "compatibility" of a carbohydrate with the three-dimensional hydrogen-bonded network of water is governed by carbohydrate stereochemistry. Monosaccharide pyranoses with variable hydroxyl stereochemistry at C2 and C4 showed a relatively wide range of isentropic partial molar compressibilities. In other words, the 'fit' of monosaccharide pyranoses into the three-dimensional hydrogen bonded network is highly variable and dependent upon the C2 and C4 relative stereochemistry. The molar compressibilities of D-talose, D-mannose, D-glucose and D-galactose are shown in Figure 3.4 where hexoses with low molar compressibilitity values exhibit the poorest fit with the three-dimensional hydrogen-bonded network of water.[46]

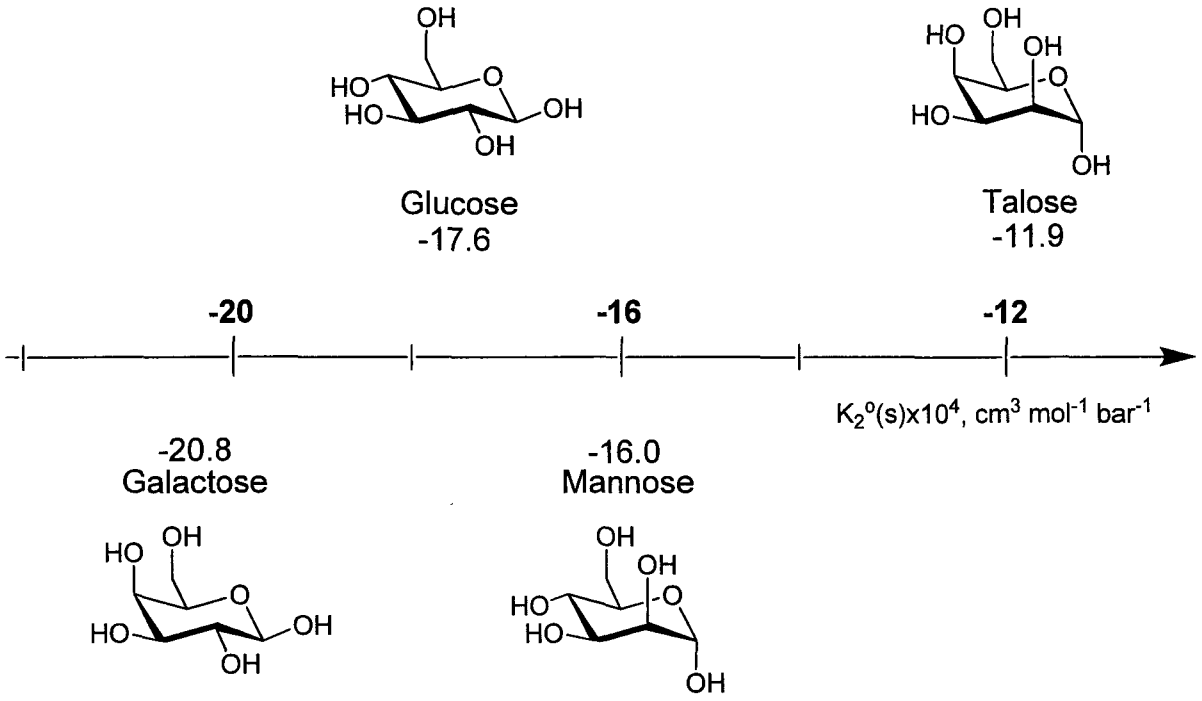


Figure 3.4. Isentropic partial Molar Compressibilities of Monosaccharides in Aqueous Solution at 298 K ($K_2°(s)$ x 10^4, cm^3 mol^{-1} bar^{-1}).[46, 47] Values are given for dominant conformer in solution.

A comparison of Graph 3.1 to Figure 3.4 reveals a strong correlation between ice recrystallization inhibition activity and molar compressibility of the monosaccharide hexoses. The fact that the AFGP analogues are *C*-linked deserves special comment. Recent studies have shown that *C*-linked carbohydrate derivatives can adopt many more conformations than their *O*-linked analogues.[48] However, much precedent exists to indicate that *C*-linked carbohydrates are able to adopt key conformations *in vitro* ensuring they are just as active or in some instances more active than their *O*-linked counterparts.[49-51] Furthermore, conformational analyses of *C*-linked and *O*-linked oligosaccharides suggest that the conformations between these two substrates are identical as intramolecular hydrogen bonds play a significant role in dictating

conformation.[51, 52] Intramolecular hydrogen bonding cooperativity also has an effect on the structuring of water, and has been studied extensively by Dashnau et al.[53] This study indicated that the stereochemistry of the C4 OH is crucial in forming intramolecular hydrogen bonds. In addition, the stereochemistry at C4 and C2 serves to modulate the magnitude of the response for hydration.[47, 53] The stereochemistry at anomeric and C3 position has been shown to not affect the hydration[53] and thus, the hydration characteristics between an *O*-linked pyranose and a *C*-linked pyranose should not differ.

Our data (Graph 3.1.) indicates that the compatibility of a hexose is inversely proportional to recrystallization-inhibition activity. For example, talose-containing AFGP analogue **19** exhibits no RI activity (Graph 3.1.), however, based upon the fact that D-talose would have the best 'fit' into the three-dimensional hydrogen-bonded network of water compared to the other hexoses, this analogue would demand the least re-organization of bulk water and thus require the least amount of energy for solvation. The D-mannose (**21**) and D-glucose (**20**) analogues exhibit only slight RI activity and the molar compressibility values correlate nicely with these results. In contrast, D-galactose analogue **19** has the lowest molar compressibility value and hence, would exhibit the poorest fit into the three-dimensional hydrogen-bonded network of water. Consequently, the energetic cost to incorporate this analogue into the three-dimensional hydrogen-bonded network of water would be the largest.

Due to this apparent correlation between carbohydrate stereochemistry and hydration, we believe that recrystallization-inhibitors such as biological antifreezes or our *C*-linked AFGP analogues may function by disturbing the highly ordered structure of supercooled water. This in turn would increase the energy associated with a water molecule transferring from bulk water to the pre-ordered water layer or "quasi-liquid layer" (QLL), and subsequently from the QLL to the

ice lattice. The QLL is a transitional domain existing at the ice/water interface.[34, 54] While the thickness of the QLL has been shown to be temperature dependant, the characteristics of this region more closely resemble that of the ice lattice (i.e. less dynamic and more ordered than bulk water) at temperatures less than –1 °C.[34] The net result of disrupting the ordered structure of supercooled water would be an inhibitory effect on ice growth.

The mechanism of recrystallization has been dealt with extensively in the metallurgical literature.[55] There are two generally accepted theories describing the mechanism of recrystallization: Ostwald ripening and agglomeration. Ice recrystallization is defined as the formation of larger crystals at the expense of smaller ones, and likely occurs when the QLLs of two crystals are in contact with each other (via agglomeration mechanism). [55c, d] The non-uniform crystal shapes from our RI experiments suggest that recrystallization is occurring predominately via agglomeration.[55c] According to IR studies by Sadtchenko, the thickness of the QLL is inversely proportional to temperature, and at -6 °C, the thickness of the quasi-liquid layer is ~1 nm, which equates to about only three monolayers of water.[34] The consequence of such a thin layer means that a carbohydrate residue with low molar compressibility will have a more pronounced effect on the ordering of the bulk water/QLL interface and ultimately ice growth.

The conformation of our analogues (**19-22**) was examined by CD to ensure that the substitution of carbohydrates did not drastically affect the secondary structure of the glycopeptide (Figure 3.3, Table 3.1). All of these glycoconjugates possess a large percentage of random coil structure (> 50 percent) and β-turn (~25%), with varying degrees of other secondary structure. These conformational characteristics are consistent with previous analyses of AFGP 8[12, 29] and indicate that the absence of RI activity for analogues **21** and **22** is not due to a change in secondary structure. It is feasible that a high degree of flexibility in the polypeptide may be an

important factor for RI activity allowing the glycoprotein to move easily through the concentrated solution of solute during the freezing process. Further studies to investigate this hypothesis are necessary.

3.7. Conclusion

In summary, it was demonstrated that the stereochemistry of the carbohydrate moiety in *C*-linked AFGP analogues is an extremely important factor for recrystallization-inhibition activity. Furthermore, the molar compressibility of the *C*-linked pyranose in the glycoconjugate is inversely proportional to the degree of recrystallization-inhibition activity. The relative stereochemistry of the hydroxyl groups at C2 and C4 plays an important role in determining the compatibility of the carbohydrate component with the three-dimensional hydrogen-bonded network of bulk water at the ice-protein interface. These results suggest that potent inhibitors of recrystallization may ultimately function by disrupting the ordering of bulk water present at the QLL-bulk water interface, thus making addition of a water molecule from supercooled bulk water to the QLL unfavorable. This will result in inhibition of ice growth. Moreover, CD analysis of the *C*-linked AFGP analogues used in this study indicates that all of the glycoproteins contain a large amount of random coil structure in solution and that the decreased RI activity associated with analogues **19-22** is not due to a change in secondary structure. These results emphasize the importance of additional studies to further elucidate the role of hydration in carbohydrates, peptides and proteins and its subsequent effects on conformation, recognition and activity in biological systems. In the next chapter 4, the effect of carbohydrates hydration on RI will be further explored using *C*-linked AFGP analogues containing disaccharide residues.

3.8. References

1. Franks, F., *Pure Appl. Chem.* **1987**, 59, 1189.
2. Quiocho, F. A., *Ann. Rev. Biochem.* **1986**, 55, 287.
3. Lemieux, R. U., *Chem. Soc. Rev.* **1989**, 18, 347.
4. Yeh, Y., Feeney, R. E., *Chem. Rev.* **1996**, 96, 601.
5. Nguyen, D. H., Colvin, M.E., Yeh, Y., Feeney, R. E., Fink, W. H., *Biopolymers* **2004**, 75, 109.
6. Kuroda, T., *Proceedings of the 4th Topical Conference on Crystal Growth Mechanisms, Japan* **1991**, 157.
7. Ba, Y., Wongskhaluang, J., Li, J., *J. Am. Chem. Soc.* **2003**, 125, 330.
8. Nguyen, D. H., Colvin, M. E., Yeh, Y., Feeney, R. E., Fink, W. H., *Biophys J.* **2002**, 82, 2892.
9. Komatsu, S. T., DeVries, A. L., Feeney, R. E., *J. Biol. Chem.* **1970**, 245, 2909.
10. Sönnichsen, F. D., Sykes, B. D., Davies, P.L., *Protein Sci.* **1995**, 4, 460.
11. Wilson, P. W., *Cryobiology* **1994**, 31, 406.
12. Bouvet, V. R., Lorello, G. R., Ben, R. N., *Biomacromolecules* **2006**, 7, 565.
13. Eniade, A., Purushotham, M., Ben, R. N., Wang, J. B., Horwath, K., *Cell Biochem. Biophys.* **2003**, 38, 115.
14. (a) Murphy, A. V., Ph.D. Thesis, SUNY Binghamton, Binghamton **2004**. (b) Eniade, A., Murphy, A. V., Landreau, G., Ben, R. N., *Bioconj. Chem.* **2001**, 12, 817.
15. (a) Dwek, R. A., *Chem. Rev.* **1996**, 683. (b) Bertozzi, C. R., Kiessling, L. L., *Science* **2001**, 2357.
16. (a) Herzner, H., Reipen, T., Schultz, M., Kunz, H., *Chem. Rev.* **2000**, 100, 4495. (b) Seitz, O. ChemBioChem. 2000, 1, 214. (c) Paulsen H., *Angew. Chem. Int. Ed.* **1990**, 29,

823. (d) Kunz, H., *Pure Appl. Chem.* **1993**, 65, 1223. (e) Meldal, M., Bock, K., *Glycoconj. J.* **1994**, 11, 59. (f) Witte, K., Seitz, O., Wong, C. H., *J. Am. Chem. Soc.* **1998**, 120, 1979. (g) Roberge, J. Y., Beebe, X., Danishefsky, S. J., *J. Am. Chem. Soc.* **1998**, 120, 3915. (h) Marcaurelle, L. A., Bertozzi, C. R. *J. Am. Chem. Soc.* **2001**, 123, 1587. (i) Laczko, I., Hollosi, M., Urge, L., Ugen, K. E., Weiner, D. B., Mantsch, H. H., Thurin, J., Otvos Jr., L., *Biochemistry*, **1992**, 31, 4282. (j) Shao, N., Xue, J., Guo, Z., *J. Org. Chem.* **2003**, 68, 9003.

17. (a) Merrifield, R. B. J., *Am. Chem. Soc.* **1963**, 85, 2149. (b) Lerner, R. A., *Science* **2006**, 313, 57. (c) Gierasch, L. M., *Pept. Sci.* **2006**, 84, 433. (d) Jones, H. J., *J. Pept. Sci.* **2007**, 13, 363.

18. (a) Carpino, L. A., Han, G. Y., *J. Org. Chem.* **1972**, 37, 3404. (b) Fields, G. B., Noble, R. L., *Int. J. Pept. Protein Res.* **1990**, 35, 161.

19. Bundle, D. R., Eichler, E., Lowary, T. L., *Can. J. Chem.* **2002**, 80, 1112.

20. Ben, R. N., Eniade, A.A., Hauer, L., *Org. Lett.* **1999**, 1, 1759.

21. Eniade, A., Ben, R. N., *Biomacromolecules* **2001**, 2, 557.

22. Pontén, F., Magnusson, G., *J. Org. Chem.* **1996**, 61, 7463.

23. Fields, G. B., Fields, C. G., *J. Am. Chem. Soc.* **1991**, 113, 4202.

24. Knight, C. A., Hallett, J., DeVries, A. L., *Cryobiology* **1988**, 25, 55.

25. Liu, S., Ben, R. N., *Org. Lett.* **2005**, 7, 2385.

26. Greenfield, N. J., *Anal. Biochem.* **1996**, 35, 1.

27. Sreerama, N., Venyaminov, S. Y., Woody, R. W., *Anal. Biochem.* **2000**, 287, 243.

28. Sreerama, N., Woody, R.W., *Anal. Biochem.* **2000**, 287, 252.

29. Raymond, J. A., Radding, W., DeVries, A. L., *Biopolymers* **1977**, 16, 2575.

30. Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishimura, S. I., *Angew. Chem. Int. Ed.* **2004**, 43, 856.

31. Corzana, F., Busto, J. H., Jimenez-Oses, G., de Luis, M. G., Asensio, J. L., Jimenez-Barbero, J., Peregrina, J. M., Avenoza, A., *J. Am. Chem. Soc.* **2007**, 129, 9458.

32. Wang, T., Zhu, Q., Yang, X., Layne, J. R., DeVries, A. L. , *Cryobiology* **1994**, 31, 185.

33. Petzel, D. H., DeVries, A. L., *Cryobiology* **1979**, 16, 585.

34. Sadtchenko, V., Ewing, G. E., *J. Chem. Phys.* **2002**, 116, 4686.

35. Stokes, R. H., Robinson, R. A., *J. Phys. Chem.* **1966**, 70, 16.

36. Suggett, A., Ablett, S., Lillford, P. J., *J. Solution Chem.* **1976**, 5, 17.

37. Tait, M. J., Suggett, A., Franks, F., Abblett, S., Quikenden, P. A., *J. Solution Chem.* **1972**, 1, 131.

38. Uedaira, H., *J. Solution Chem.* **1985**, 14, 7.

39. Kabayama, M. A., Patterson, D., Piche, L., *Can. J. Chem.* **1958**, 36, 557.

40. Franks, F., *Cryobiology* **1983**, 20, 335.

41. Suggett, A., *J. Solution Chem.* **1976**, 5, 33.

42. Miyajima, K., Machida, K., Nagagaki, M., *Bull. Chem. Soc. Jpn.* **1985**, 58, 2595.

43. Walkinsaw, M. D., *J. Chem. Soc., Perkin Trans. 2* **1987**, 1903.

44. Danford, M. D., *J. Am. Chem. Soc.* **1962**, 84, 3965.

45. Warner, D. T., *Nature* **1962**, 196, 1055.

46. Galema, S. A., Hoiland, H., *J. Phys. Chem.* **1991**, 95, 5321.

47. (a) Galema, S. A., Howard, E., Engberts, J. B. F. N., Grigera, J. R., *Carbohydr. Res.* **1994**, 265, 215. (b) Galema, S.A., Engberts, J. B. F. N., Høiland, H., Førland, G. M., *J. Phys. Chem.* **1993**, 97, 6885.

48. Espinosa, J. F., Montero, E., Vian, A., Garcia, J. L., Dietrich, H., Schmidt, R. R., Martin-Lomas, M., Imberty, A., Canada, F. J., Jimenez-Barbero, J., *J. Am. Chem. Soc.* **1998**, 120, 1309.

49. Wang, J., Kovac, P., Sinay, P., Glaudemans, C. P. J., *Carbohydr. Res.* **1998**, 308, 191.

50. Wang, Y., Babirad, S. A., Kishi, Y., *J. Org. Chem.* **1992**, 57, 468.

51. Ravishankar, R., Surolia, A., Vijayan, M., Lim, S., Kishi, Y., *J. Am. Chem. Soc.* **1998**, 120, 11297.

52. Ma, B., Schaefer III, H. F., Allinger, N. L., *J. Am. Chem. Soc.* **1998**, 120, 3411.

53. Dashnau, J. L., Sharp, K. A., Vanderkooi, J. M., *J. Phys. Chem. B* **2005**, 109, 24152.

54. Karim, O. A., Haymet, A. D. J., *J. Chem. Phys.* **1988**, 89, 6889.

55. (a) Pronk, P., Infante Ferreira, C. A., Witkamp, G. J., *J. Crystal Growth* **2005**, 275, e1355. (b) Huige, N. J. J., Thijsenn, H. A. C., *J. Crystal Growth* **1972**, 13/14, 483. (c) Inada, T., Modak, P. R., *Chem. Eng. Sci.* **2006**, 61, 3149. (d) Kingery, W. D., *J. Appl. Phys.* **1959**, 30, 301.

CHAPTER 4

Towards Engineering RI Activity of the C-linked AFGP Analogues

4.1. Introduction

In the previous chapter, [O(monosaccharide)GG]$_4$G AFGP analogues **19-22** were synthesized and assessed for RI activity.[1] It was discovered that the fit of the carbohydrate of the glycopeptide into the three-dimensional hydrogen-bonded network of water is inversely proportional to RI activity. Furthermore, structural changes in this series of analogues confirmed that stereochemistry of the carbohydrate hydroxyl groups at C2 and C4 have a significant effect on hydration and consequently recrystallization inhibition. While the effect on RI of many hexose sugars have been extensively studied in our laboratory, this chapter will examine the influence of disaccharide C-linked AFGP analogues.

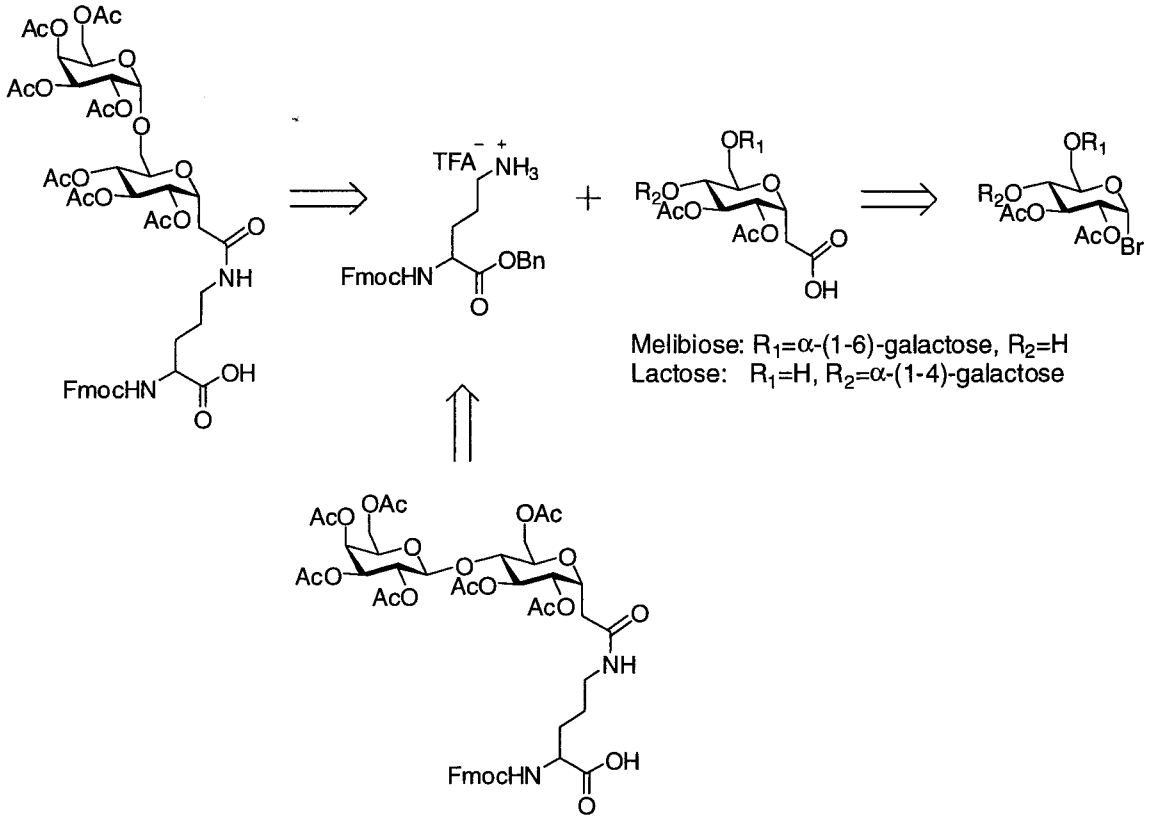
Lactose and melibiose are commercially available disaccharides for which the hydration parameters have been determined.[2] From the list of carbohydrates studied by Galema et al., these are the only disaccharides comprised of two hexoses. Consequently we sought to prepare lactose (β-D-Gal-(1-->4)-D-Glc) and melibiose (α-D-Gal-(1-->6)-D-Glc) C-linked AFGP analogues possessing the [O(carbohydrate)GG]$_4$G backbone. The antifreeze activity of these analogues and will determine whether there is a need for generation of additional analogues containing various di- and oligo-saccharides. To date no C-linked AFGP analogue containing a more complex carbohydrate than a monosaccharide has been synthesized.

In chapter 3, partial molar compressibility (PMC) values were used to describe the hydration however, recent work in our laboratory has led to the concept of a hydration index.[3] We have calculated hydration indices for several monosaccharides and disaccharides studied by

Galema, and correlated these to RI activity. Subsequently, it is hypothesized that RI activity of the C-linked AFGP analogues may also be related to hydration indices of their carbohydrates. To test this theory, RI activity of [Orn(Lac)-Gly-Gly]$_4$Gly and [Orn(Meli)-Gly-Gly]$_4$Gly analogues and those described previously [Orn(Gal)-Gly-Gly]$_4$Gly (**19**), [Orn(Glc)-Gly-Gly]$_4$Gly (**20**), [Orn(Man)-Gly-Gly]$_4$Gly (**21**), and [Orn(Tal)-Gly-Gly]$_4$Gly (**22**), will be correlated to partial molar compressibility values and hydration indices of carbohydrates. This will allow us to examine how these two parameters compare in predicting recrystallization inhibition. Moreover, the results in this chapter will reveal how by estimating hydration indices of glycopeptides we might be able to engineer AFGP analogues with superior RI activity.

4.2. Retrosynthetic Analysis of C-linked [Orn(Lac)-Gly-Gly]$_4$ and [Orn(Meli)-Gly-Gly]$_4$ Analogues

Melibiose and lactose ornithine building blocks can be obtained by assembly of the amide linkage between the TFA salt of ornithine and carboxylic acid derivative of the desired disaccharide (Scheme 4.1.). Melibiose and lactose carboxylic acids can be prepared by bromination of their acetylated derivatives, following allylation and subsequent oxidation.

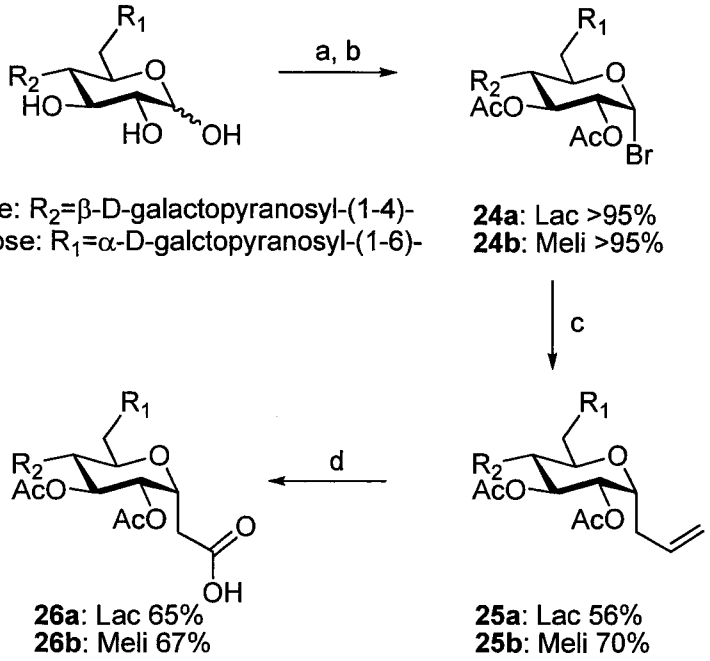


Scheme 4.1. Retrosynthetic analysis of the C-linked melibiose-ornithine and lactose-ornithine building blocks.

4.3. Synthesis of the C-linked AFGP Analogues 31-32

The lactose and melibiose disaccharide series were prepared as outlined in Scheme 4.2. The disaccharides **23a, b** were acetylated using acetic anhydride, pyridine and catalytic DMAP in CH_2Cl_2. Without further purification, peracetylated disaccharides were diluted in CH_2Cl_2 and converted to glycosyl bromides **24a, b** (>95% yield) by treatment with hydrogen bromide in glacial acetic acid. Photochemically-mediated allylation furnished *C*-allylated disaccharide derivatives **25a, b**.[4] The α/β ratio of the anomers for melibiose was 4/1 and for lactose 3/1,

however only the desirable α-configuration was isolated and characterized. These derivatives were then oxidized[5] to produce the desired carboxylic acids **26a, b** in ~65% yield.

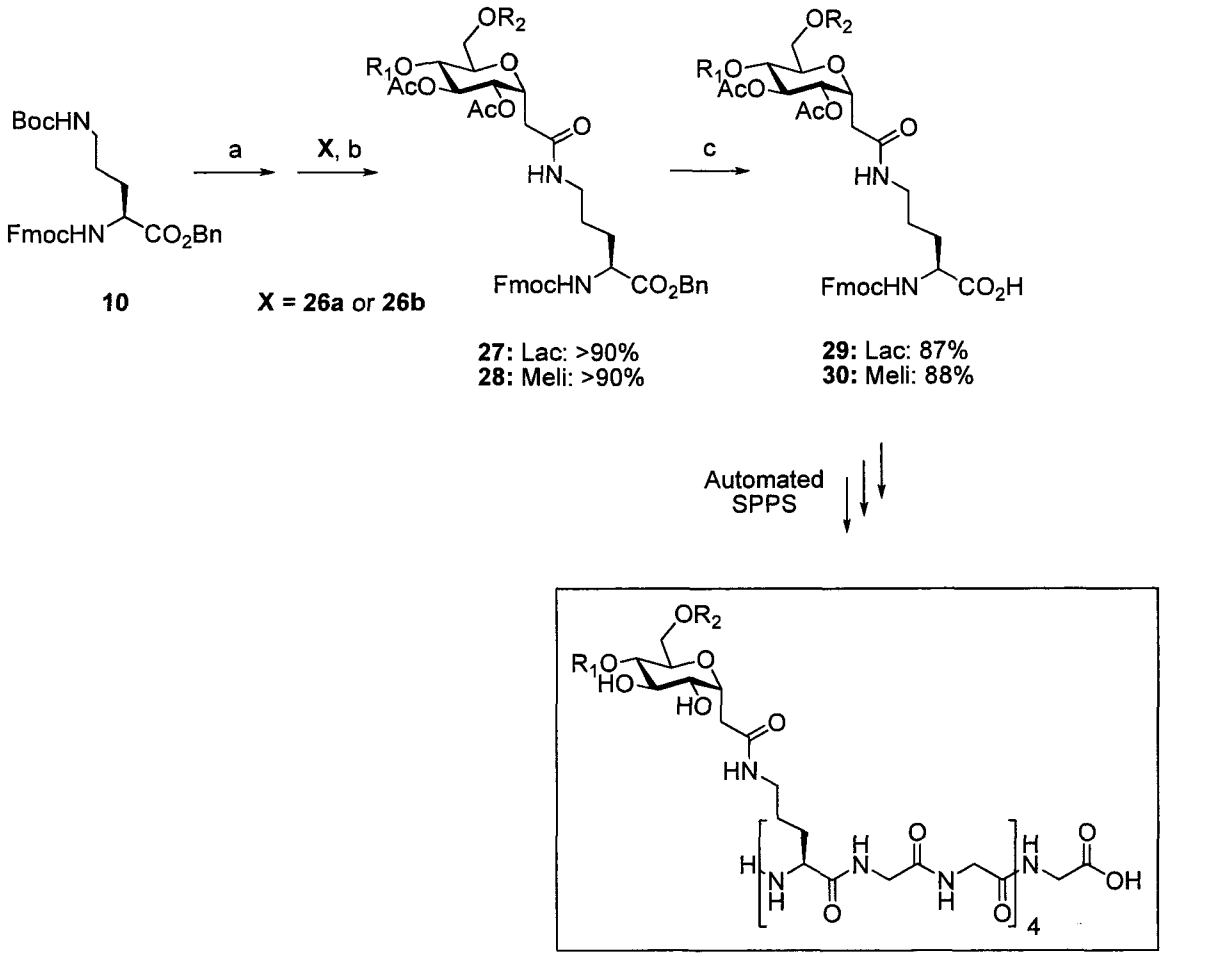


Conditions: (a) Ac_2O, py, cat. DMAP, CH_2Cl_2; (b) HBr/HOAc, CH_2Cl_2; (c) $(nBu)_6Sn_2$, allylphenylsulfone, benzene, hν; (d) $RuCl_3{\cdot}H_2O$, $NaIO_4$, $CCl_4/CH_3CN/H_2O$

Scheme 4.2. Synthesis of the C-linked melibiose and lactose derivatives **26a, b**.

Compounds **26a, b** were then coupled to orthogonally protected ornithine derivative **10** using HBTU and DIPEA to afford benzyl protected building blocks **27** and **28** in >90% yield. The benzyl esters were carefully removed by treatment with palladium-on-carbon and hydrogen gas to provide the essential glycoconjugate building blocks **29-30** (Scheme 4.3.). Subsequently, the *C*-linked AFGP analogues were prepared via automated SPPS using an APEX 396 and standard Fmoc-based chemistry.[6] The glycopolymers were removed from the Wang resin by

treatment with a 50% TFA solution in dichloromethane following ether precipitation. The acetate protecting groups on the carbohydrates were removed using Na/methanol solution. Purification by C-18 reversed-phase HPLC furnished the desired *C*-linked AFGP analogues **31-32** in 65 and 62% yield respectively.



10 **X = 26a** or **26b**

27: Lac: >90%
28: Meli: >90%

29: Lac: 87%
30: Meli: 88%

31: [O(Lac)GG]$_4$G: R_1=β-D-galactopyranosyl-(1-4), R_6=H: 65%
32: [O(Meli)GG]$_4$G: R_1=H, R_6=α-D-galactopyranosyl-(1-6): 62%

Conditions: (a) TFA/CH_2Cl_2; (b) HBTU, DIPEA, CH_2Cl_2; (c) 10% Pd-C, H_2, EtOH-EtOAc.

Scheme 4.3. Synthesis of the C-linked AFGP analogues **31** and **32**.

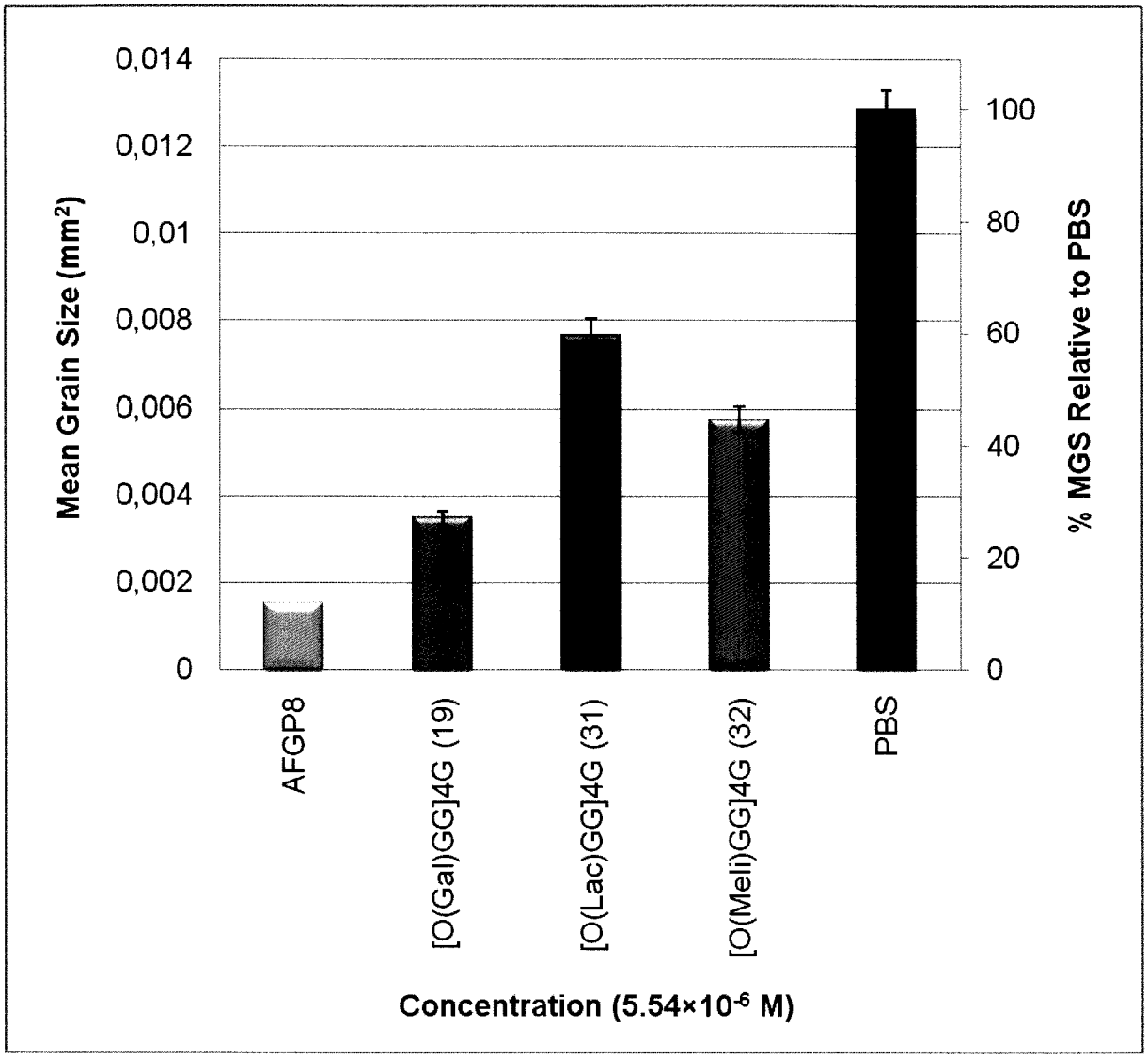
4.4. Results: Assessment of Antifreeze Activity and Conformational Analysis of C-linked AFGP Analogues 31 and 32

Thermal Hysteresis (TH):

C-Linked AFGP analogues **31** and **32** did not show any TH activity or DIS. This result is consistent with other C-lined AFGP analogues previously synthesized in our laboratory which possess custom tailored antifreeze activity (i.e. no TH, but potent RI activity).

Recrystallization-Inhibition (RI):

Analogues **31** and **32** were assessed for their ability to inhibit ice recrystallization (Graph 4.1.). All samples were compared to a solution of phosphate buffered saline (PBS) which was used as a negative control and a solution of AFGP8 isolated from *Gagus ogac* (generously provided by AquaBounty Farms). Ice crystals size was analyzed using domain recognition software (DRS) program.[7] The Y-axis in Graph 4.1 represents the mean grain size (MGS) that was normalized to MLGS where small bars are representative of potent recrystallization-inhibition activity. The X-axis represents concentration of 5.54×10^{-6} M at which significant RI activity was previously observed.

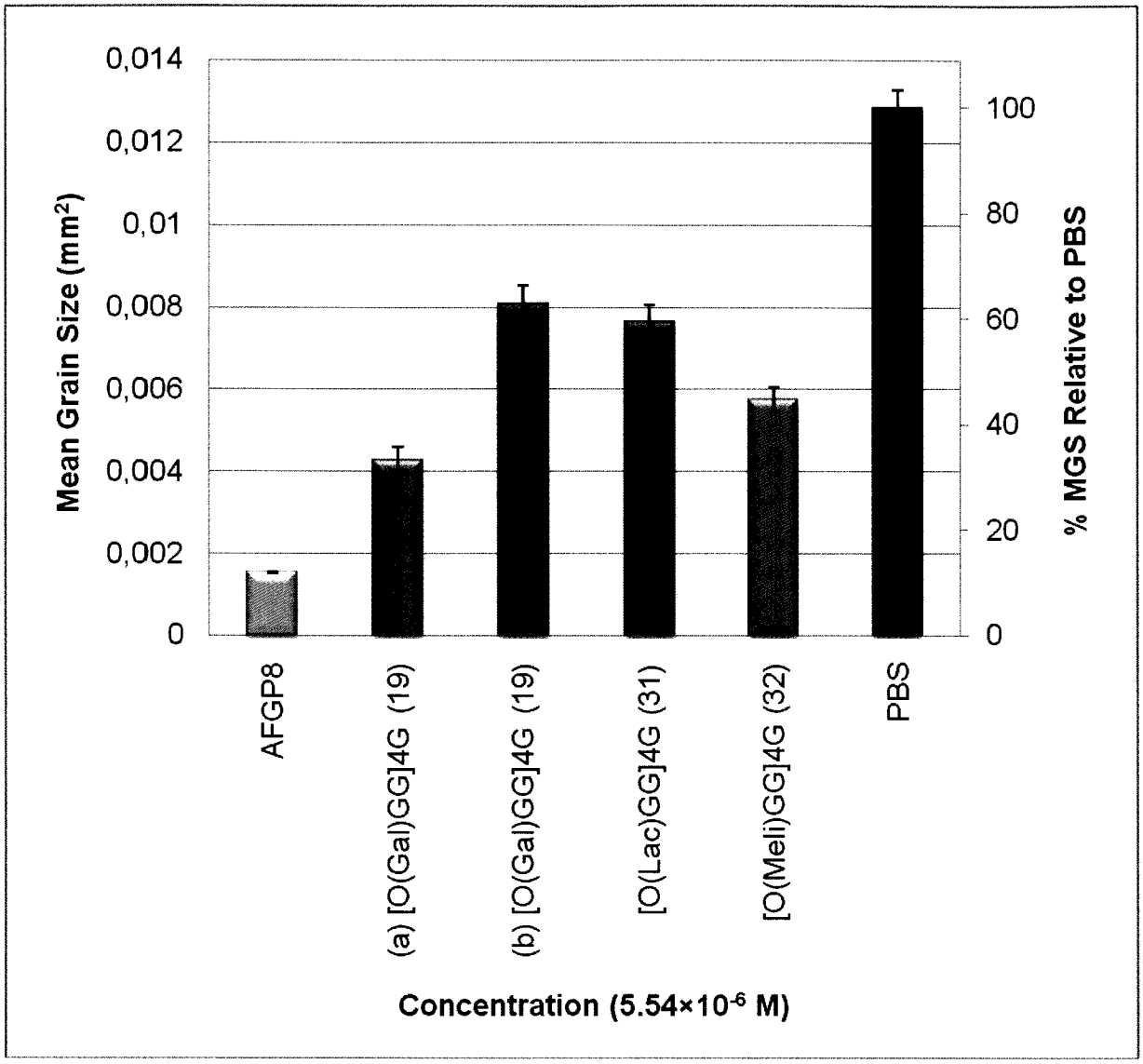


Graph 4.1. RI activity of C-Linked AFGP analogues **19**[8], **31** and **32**.

As expected, RI activity of melibiose analogue **32** was better than that of lactose analogue **31**, but surprisingly the RI activity of these two analogues did not exceed that of $[O(Gal)GG]_4G$ analogue **19**. Further, the data point for analogue **19** was measured via mean elliptical method, while the MGS for analogues **31** and **32** was evaluated using DRS program which has been reported to give more accurate results.[7] Thus, in order to have a directly

comparable data point for analogue **19** the raw data was reanalyzed using the DRS program (Graph 4.2, (a) $[O(Gal)GG]_4G$). The RI activity of **19** was determined to be ~10% lower than when analyzed via mean elliptical method (Graph 4.1). Subsequently, to verify the RI results a more recently synthesized batch of analogue **19** was tested twice and analyzed using the DRS program (Graph 4.2, (b) $[O(Gal)GG]_4G$), and its MGS value was determined to be 0.0081 ± 0.0004 mm^2. Suprisingly, the difference in RI activity between the two samples of $[O(Gal)GG]_4G$ analogue (Graph 4.2, (a) $[O(Gal)GG]_4G$ and (b) $[O(Gal)GG]_4G$) was now approximately 50%. There are many factors that might account for this discrepancy and hence it will be it will be necessary to synthesize, test and analyze other samples of analogue **19** to validate this data point. Nevertheless, data point '(b) $[O(Gal)GG]_4G$' presented in Graph 4.2 will be used to compare the activity of analogue **19** to subsequent series that are described in this thesis because it better matches trends in RI predicted via the hydration index model which will be discussed imminently.

The data presented in Graph 4.2 shows that lactose analogue **31** exhibited slightly better RI activity when compared to analogue **19**, and melibiose analogue **32** was the most potent of these analogues with MGS value of 0.0058 ± 0.0003 mm^2.



Graph 4.2. RI activity of C-Linked AFGP analogues **31**, **32** and **19** ((a) ice crystals from ref. 8, (b) resynthesized analogue) analyzed via DRS program.

Conformational Analysis:

The solution conformation of C-linked AFGP analogues **31** and **32** was examined by circular dichroism (CD) spectroscopy.[9] Deconvolution of the raw spectral data was accomplished

using CD Pro with Selcon III and I-Basis 4 as well as I-Basis 5 as representative databases.[10] The CD spectrum is presented in Figure 4.1.

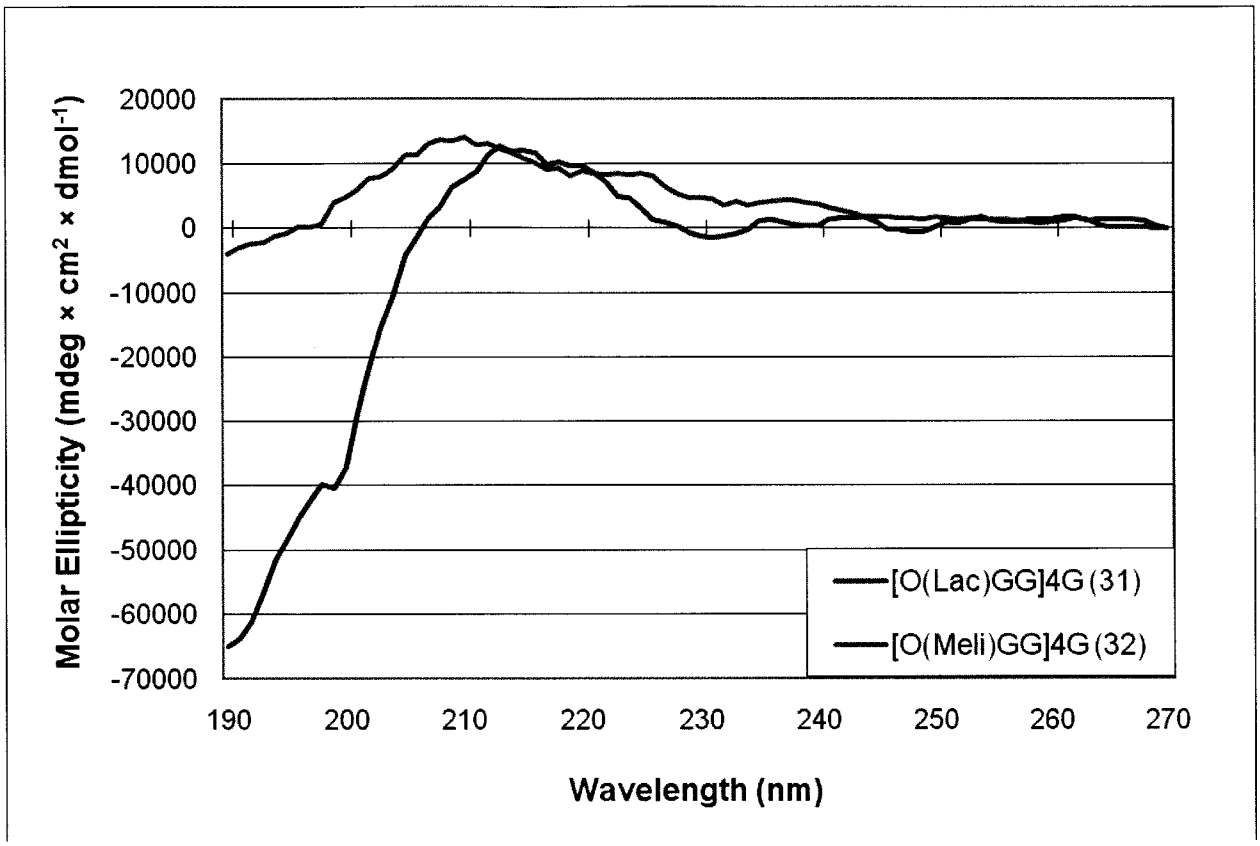


Figure 4.1. Circular Dichroism spectra of C-Linked AFGP analogues **31** and **32**.

In general, the solution conformations of analogues **31** and **32** are consistent with a large percentage of random coil structure with varying degrees of more ordered secondary structure (Table 4.1. and 4.2.). While the distribution of secondary structures varies depending on the IBASIS used, the data correlates well with earlier conformational analyses of AFGP8 using CD spectroscopy.[9, 11]

Table 4.1. CD deconvolution data of AFGP analogues **31**, **32** and AFGP8 using CD Pro with Selcon III and IBASIS 4.

Compound	Regular α helix	Distorted α helix	Regular β strand	Distorted β strand	Turn	Random coil	Sum
$[O(Lac)GG]_4G$ (**31**)	-0.017	0.039	-0.023	0.100	0.263	0.647	1.008
$[O(Meli)GG]_4G$ (**32**)	-0.014	-0.027	0.302	0.198	0.234	0.160	0.853
AFGP8	0.049	0.114	0.164	0.094	0.237	0.378	1.036

Table 4.2. CD deconvolution data of AFGP analogues **31**, **32** and AFGP8 using CD Pro with Selcon III and IBASIS 5.

Compound	α helix	β strand	β turn	PPII	Random coil	Sum
$[O(Lac)GG]_4G$ (**31**)	0.033	0.113	0.083	0.077	0.704	1.008
$[O(Meli)GG]_4G$ (**32**)	-0.089	0.353	0.102	0.109	0.373	0.853
AFGP8	0.129	0.159	0.156	0.136	0.445	1.025

4.5. Discussion: Interpretation of Antifreeze Activity and Conformational Analysis of Analogues 31 and 32

Partial molar volumes (PMVs), isentropic partial molar compressibilities and hydration numbers (HNs) of ten disaccharides were determined by Galema et al.[2] In order to synthesize $[O(disaccharide)GG]_4G$ analogues and subsequently compare their RI activity with our $[O(monosaccharide)GG]_4G$ series, the list of these hydration parameters was examined to pick out disaccharides which do not possess 1-->1 glycosidic bonds and only contain hexoses. The

lactose and melibiose were the only two that satisfied these criteria and also have the lowest isentropic partial molar compressibilites (Figure 4.2). These disaccharides can be easily modified and incorporated into our C-linked AFGP design, and their hydration and antifreeze activity can be directly compared to analogues **19**-**22** that consist of hexoses.

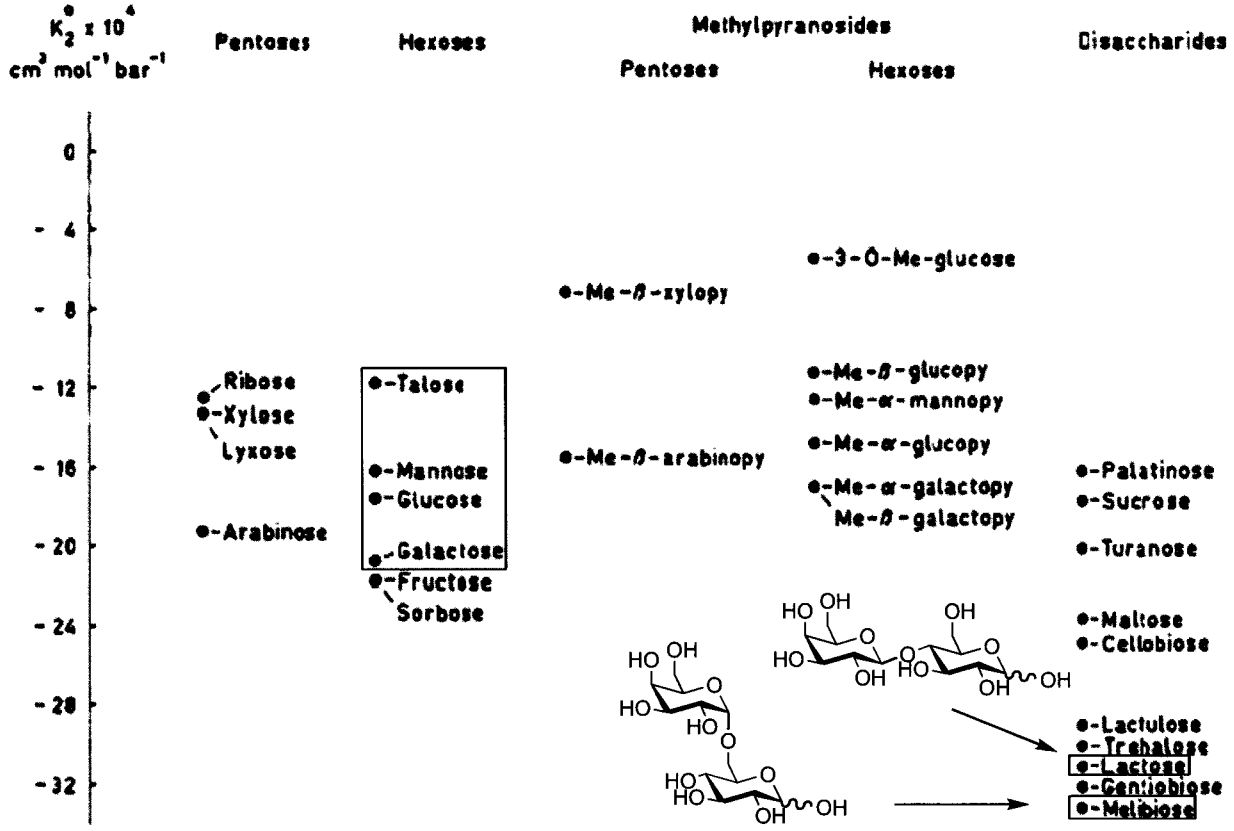


Figure 4.2. Isentropic partial molar compressibilities of carbohydrates relevant to the synthesized analogues **19**-**22**, **31** and **32**.[2]

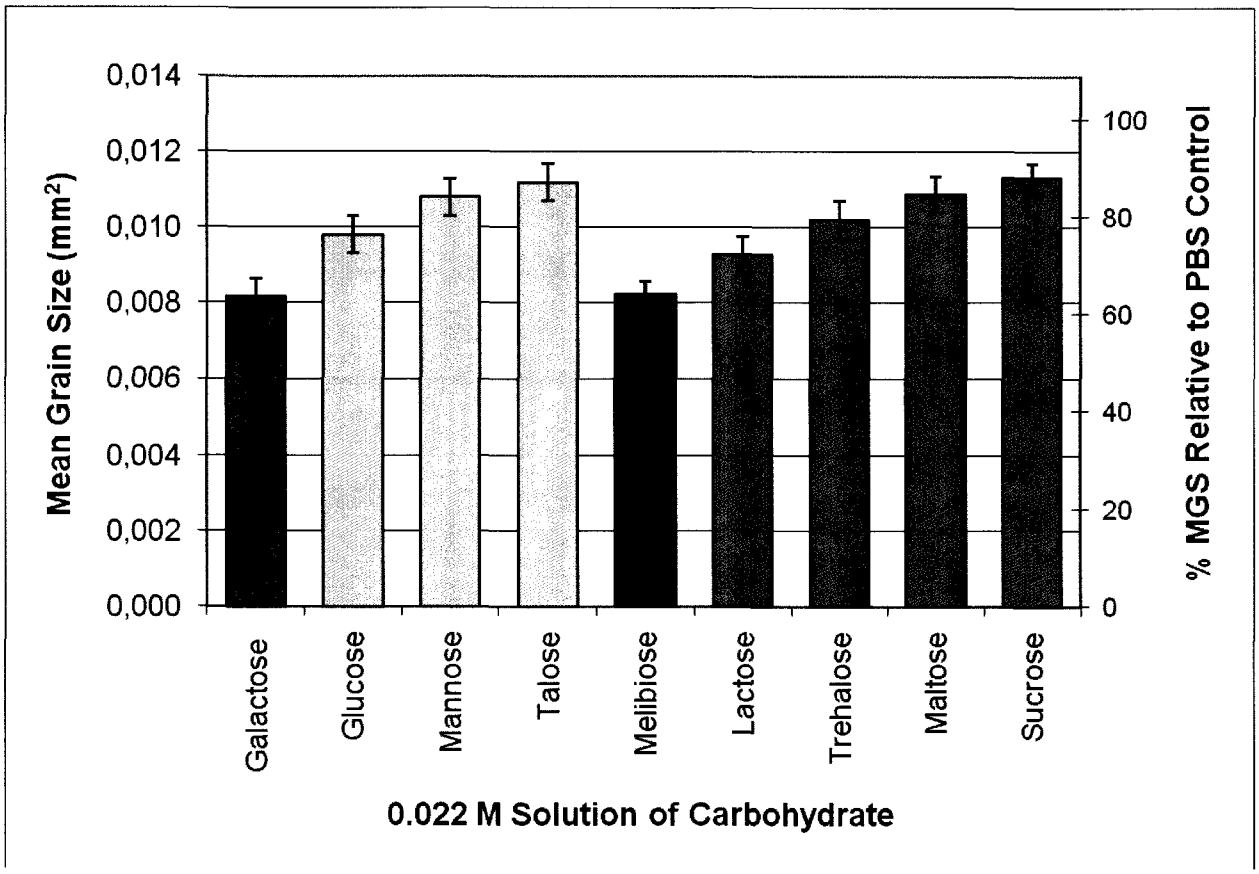
In chapter 3, $[O(Gal)GG]_4G$ **19** exibited a drastic improvement in RI activity in comparison to the $[O(Tal)GG]_4G$ analogue **22**.[1] Also, as deduced from Graph 4.2 the partial molar compressibility increase from D-talose to D-galactose is ~10 units. This increase in PMVs

is similar when comparing D-galactose to melibiose. Thus, we speculated that a significant improvement in RI activity can be achieved with $[O(Meli)GG]_4G$ analogue.

The RI activity of previously synthesized divalent analogue, $[K(Gal_2)GG]_4G$, was comparable to that of $[K(Gal)GG]_9G$ and it was more than doubled when compared to $[K(Gal)GG]_4G$ (Table 2.1.).[8, 12, 13] Hence, it was proposed that the carbohydrate concentration with respect to the peptide backbone is very important for potent recrystallization inhibition. Since the concentration of hexoses is doubled between analogue **19** and analogues **31** and **32**, we expected RI activity of the disaccharide analogues to be significantly greater than that of the monosaccharide analogues. However, the amount by which RI activity of $[O(disaccharide)GG]_4G$ analogues increases will depend on the structure (regiochemistry/stereochemistry) and thus hydration of the disaccharide.[1, 2]

Recently Ben et al. quantified RI activity of various monosaccharides and disaccharides (Graph 4.3.) and correlated these to hydration.[3] The trend in RI activity for each class of carbohydrates (monosaccharides and disaccharides, Graph 4.3.) correlated positively with isentropic molar compressibility values (Figure 4.2.).[2] However, predicted increase in RI activity for the disaccharides relative to monosaccharides was not observed. For instance, the large difference in hydration numbers between galactose (hydration number = 8.7) and melibiose (15.5) did not result in a corresponding increase in RI activity. More surprisingly, the difference in hydration numbers between galactose (8.7) and sucrose (13.9) resulted in a decrease in RI activity for the latter. Ben et al. proposed that this is due to the difference in steric volume between monosaccharides and disaccharides and have corrected for it by dividing hydration numbers by partial molar volumes. This resulted in the number of tightly bound water molecules per molar volume of carbohydrate and was termed a *hydration index*. Consequently, it has been

determined that while the absolute number of water molecules surrounding a solute is important for predicting RI activity, the number of water molecules per unit volume of solute is also important. This was the first study on RI activity of disaccharides and forms the basis for the work presented in this chapter.



Graph 4.3. RI activity of various monosaccharides and disaccharides at 0.022 M in PBS solution.[3]

As depicted in Graph 4.1., $[O(Lac)GG]_4G$ analogue **31** (MGS=0.0077±0.0004 mm^2) possesses RI activity 5% more potent as compared to $[O(Gal)GG]_4G$ analogue **19** (MGS=0.0080±0.0004mm^2) while $[O(Meli)GG]_4G$ analogue **32** shows to be the best recrystallization inhibitor with MGS value of 0.0058±0.0003 mm^2. This contradicts the results obtained when testing carbohydrates for RI activity.[3] Within error, both carbohydrates, melibiose and D-galactose had the same MGS values. The trend observed between glycopeptides **31** and **32** (Graph 4.2) conforms to the tabulated data for isentropic partial molar compressibilities (Figure 4.2). While the isentropic partial molar compressibilities values cannot be directly correlated to the values of MGS, in chapter 3 it was demonstrated that they can be used as indicators for prediction of RI activity. However, hydration index defined by Ben et al describes a more precise way for estimating hydration of a solute, as it takes into account the concentration of hydrated water molecules around a particular solute. Consequently, we set out to estimate hydration indices for glycopeptides and examine their relationship with RI activity.

In order to calculate hydration index of glycopeptides, partial molar volume and hydration number of a particular glycopeptide had to be determined. If recrystallization inhibition reported in Graph 4.1 was solely based hydration of the carbohydrate then comparable results between carbohydrates and glycopeptides should have been obtained. However, it is obvious that hydrophobic peptide backbone of AFGP plays an important role in inducing RI activity. Hence, PMVs and hydration numbers of saccharides and peptides were utilized to estimate total hydration of a glycopeptide.

The partial molar volumes of carbohydrates reported by Galema were determined experimentally as described by Hoiland.[2, 14] However, PMVs of the peptide component had to be

calculated. Two common methods were found to perform this computation: Kharakozs method[15] and McGowans method[16].

Kharakoz method employs 'group additivity' approach similar to that used by Iqbal and Verrall[17], where PMVs of peptides and proteins are calculated by sequentially adding PMVs from amino acid residue contributions (side chains, terminal groups, and ionization effects). Although, PMVs of amino acid residue contributions have been well studied,[18, 19] the reported data is inconsistent because the methods used to determine them are different. For instance, some of these methods do not account for the zwitterionic nature of the amino acids, the overlap of hydration co-spheres of the side chain and the adjacent charged terminal groups, and intramolecular folding of the peptides.[20] Thus, Kharakoz`s method was not favored.

McGowan was able to construct a table of characteristic atomic volumes that can be employed to estimate partial molar volumes of any organic solute by trivial arithmetic.[16b] To test how reliable McGowans method was PMVs of carbohydrates studied by Galema et al. were calculated. As illustrated in Table 4.3, PMVs reported by Galema for each class of carbohydrates are the same within 5% error. While partial molar volumes estimated using McGowans method are the same for each class of carbohydrates, they compare to Galemas values within 10% error. This suggested that McGowans method constituted a valid approach to estimate PMV of the peptide component of our glycoconjugates (Table 4.3).

Table 4.3. Partial molar volumes reported by Galema compared to PMVs calculated using McGowans method.

Compound	PMVs reported by Galema	PMVs calculated using McGowans method
Galactose	111.0, 110.2[a]	120
Glucose	111.7[a],	120
Mannose	111.3[a]	120
Talose	112.5	120
Melibiose	210.7, 204.0[b]	216
Lactose	208.5, 209.1[a]	216
$[OGG]_4G$	-	912

[a] Reference 21, [b] Reference 22

Subsequently, an estimated PMV of the glycopeptide was obtained by adding the PMV of the peptide and the carbohydrates. However, to obtain more accurate results, PMV of each carbohydrate was corrected to account for the loss of the anomeric hydroxyl group of the carbohydrate upon coupling to the ornithine amino acid. Consequently, PMVs of analogues **19-22**, **31** and **32** were estimated from the following equation:

$$PMV_{(Glycopeptide)} = 4 \times (PMV_{(Carbohydrate)}^{a} - PMV_{(OH)}^{b}) + PMV_{(polypeptide\ backbone)}^{b}$$

where values with superscript 'a' were obtained from reference 2 and those with superscript 'b' were calculated according to method published in reference 16.

Hydration numbers for monosaccharides and disaccharides were also determined by Galema using previously established methods.[23, 24] While hydration of some naturally occurring amino acids in various solutions has been studied extensively,[19] little precedent was found on estimating hydration of peptides.[25, 26] The first of two options would be to sum up hydration numbers of amino acids[27] within a particular polypeptide. This method would likely overestimate total hydration because hydration numbers determined for each amino acid take into account solvation of their ionic and polar termini. Furthermore, hydration numbers of amino acids largely vary and depend on the experimental method used to determine them.[23, 24, 27] The general order is as follows: conductometry < ultrasonic interferometry < infrared spetrophotometry, where with conductometry the lowest hydration numbers are obtained. The maximum hydration was obtained using the IR method which shows fair agreement with recent NMR and light scattering results on similar compounds.[24] Hence, in order to add the hydration values of the carbohydrates to that of the polypeptide, the technique used to determine both sets of hydration numbers should be the same. Since ultrasonic interferometry was used to determine hydration of carbohydrates,[2] hydration values of amino acids determined via this technique were sought however, to date such data was not reported. Thus, we turned our attention to another option.

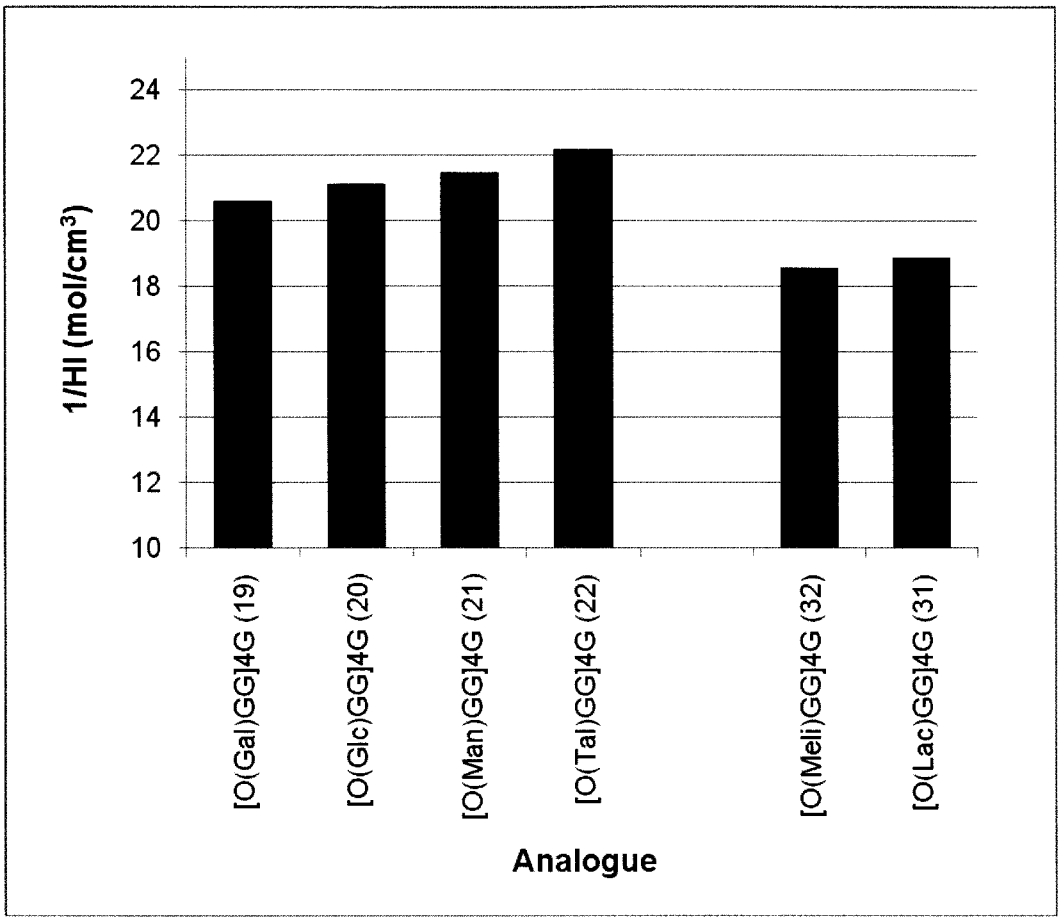
The second alternative was to use a prediction method developed by Kuntz and recently improved by Peirotti, Piaggio and Deiber. [25, 26] Kuntz was able to assign hydration number to each of the naturally occurring amino acids via NMR method.[25] The tabulated set of numbers was used to determine hydration of polypeptides via group additivity approach. Peirotti, Piaggio and Deiber refined this method by accounting for the ionic terminal groups and side chains of the amino acids.[26] Consequently, hydration number for $[OGG]_4G$ was calculated using glutamine values in place of the non-natural ornithine.

Total hydration number (HN) of the glycopeptide was estimated according to the following arithmetic:

$$HN_{(Glycopeptide)} = 4 \times HN_{(Carbohydrate)}{}^{a} + HN_{(polypeptide\ backbone)}{}^{b}$$

where values with superscript 'a' were obtained from reference 2 and those with superscript 'b' were estimated according to method published in reference 26. The calculated values are only an estimate and their error depends on the variance of their individual components. Moreover, it was discovered that small changes in the magnitude of the hydration numbers of the carbohydrates and/or the polypeptide have a large effect on the hydration index of the glycopeptide. Unfortunately, no other method exists to quantify hydration in such a trivial manner.

Consequently, having successfully estimated PMVs and hydration numbers of the glycopeptides, their hydration indices were calculated. Since it was already established that RI activity is directly proportional to hydration,[1, 3] and the potency of RI is inversely correlated to MGS values, then 1/HI was plotted for each analogue (Graph 4.4). The trend in Graph 4.4 correlates well with those of MGS values in Graph 4.2 and Graph 3.1. This confirms that hydration index, in comparison with PMC values or hydration numbers, is indeed a more proper and precise indicator for predicting RI activity.



Graph 4.4. The hydration index inverse of analogues **19-22**, **31** and **32**.

Collectively, the work described in this chapter is an extremely significant contribution for the future design of AFGP analogues because it implies that RI activity of AFGP analogues may be predicted prior to their synthesis by estimating their hydration indices. To the best of our knowledge, this is a first example of engineering RI activity. Further studies are required to ascertain if this protocol is truly applicable and can be routinely used.

Similarly to [O(monosaccharide)GG]$_4$G series, analogues **31** and **32** did not possess TH activity or DIS. This indicates there is no binding between [O(carbohydrate)GG]$_4$G analogues and the surface of the nucleated ice crystal. Hence, all the synthesized analogues possess custom tailored antifreeze properties; potent RI and no TH activity.

The conformation of our analogues **31** and **32** was examined by CD to determine if substitution of the carbohydrate did not alter the secondary structure of the glycopeptide (Figure 4.1., Table 4.1. and 4.2.). [O(Lac)GG]$_4$G analogue possesses a large percentage of random coil structure (>50%) and β-turn (>25%). [O(Meli)GG]$_4$G analogue also exhibits a large percentage of random coil (>30%) but possesses predominant proportion of β-strand (>30%). Nonetheless, both analogues show varying degrees of all secondary structures. These conformational characteristics are consistent with previous analyses of AFGP8[9, 11] and verify that the observed RI activity for **31** and **32** is a function of their hydration.

4.6. Conclusion

In summary, this is the first synthesis of C-linked AFGP analogues containing a more complex carbohydrate then a monosaccharide. Recrystallization inhibition of analogue **32** was superior to that of analogue **19**. The MGS trend determined for glycopeptides did not conform to the one observed for carbohydrates. The glycopeptides were treated as a two component system, carbohydrates and polypeptide, for which PMVs and hydration numbers were estimated. Subsequently, PMVs and hydration numbers for analogues **19-22**, **31** and **32** were calculated and their hydration indices were determined by trivial arithmetic. It was demonstrated that 1/HI clearly correlates to the RI trend which signifies that the relative hydration of the C-linked AFGP is directly proportional to the degree of recrystallization-inhibition activity. This methodology

provides a tool necessary to engineer AFGP analogues with predicted RI activity based on their hydration.

While many factors may affect hydration, conformation is one of the contributors. CD analysis of the C-linked glycoproteins **31** and **32** in solution indicates that the analogues contain a large amount of random coil structure with varying degrees of other secondary structures. Therefore, the improved RI activity associated with analogues **31** and **32** is not due to a change in their secondary structures. Collectively, the results described in this chapter have important implications in the rational design of biological antifreezes with custom-tailored antifreeze activity for cryomedical applications.

4.7. References

1. Czechura, P., Tam, R. Y., Dimitrijevic, E. Murphy A. V., Ben, R. N., *J. Am. Chem. Soc.* **2008**, 130, 2918.
2. Galema, S. A., Hoiland, H., *J. Phys. Chem.* **1991**, 95, 5321.
3. Tam, R. Y, Ferreira, S. S., Czechura, P., Chaytor, J. L., Ben, R. N., *J. Am. Chem. Soc.* **2008**, 130, 17494.
4. Pontén, F., Magnusson, G., *J. Org. Chem.* **1996**, 61, 7463.
5. (a) Carlsen, H. J., Katsuki, T., Martin, V. S., Sharpless, B. K., *J. Org. Chem.* **1981**, 46, 3936. (b) Plietker, B., *Synthesis* **2005**, 15, 2453.
6. Fields, G. B., Fields, C. G., *J. Am. Chem. Soc.* **1991**, 113, 4202.
7. Jackmann, J., Noestheden, M., Moffat, D., Pezacki, J., Findlay, S., Ben, R. N., *Biochem. Biophys. Res. Commun.* **2007**, 354, 340.
8. Murphy, A. V., Ph.D. Thesis, SUNY Binghamton, Binghamton **2004**.
9. Bouvet, V. R., Lorello, G. R., Ben, R. N., *Biomacromolecules* **2006**, 7, 565.
10. (a) Greenfield, N. J., *Anal. Biochem.* **1996**, 35, 1. (b) Sreerama, N., Venyaminov, S. Y., Woody, R. W., *Anal. Biochem.* **2000**, 287, 243. (c) Sreerama, N., Woody, R.W., *Anal. Biochem.* **2000**, 287, 252.
11. (a) Raymond, J. A., Radding, W., DeVries, A. L., *Biopolymers* **1977**, 16, 2575. (b) Wen, D. Y., Laursen, R. A., *FEBS Lett.* **1993**, 317, 31.
12. Eniade, A., Murphy, A. V., Landreau, G., Ben, R. N., *Bioconj. Chem.* **2001**, 12, 817.
13. Eniade, A., Purushotham, M., Ben, R. N., Wang, J. B., Horwath, K., *Cell Biochem. Biophys.* **2003**, 38, 115.
14. Heiland, H. J. *Solution Chem.* **1980**, 9, 857.

15. (a) Kharakoz, D. P., *J. Phys. Chem.* **1991**, 95, 5634. (b) Kharakoz, D. P., *Biochem.* **1997**, 36, 10276.

16. (a) Mellors, A., McGowan, J. C., *Biochem. Pharmacol.* **1985**, 34, 2413. (b) Abraham, M. H., McGowan, J. C., *Chromatographia*, **1987**, 23, 243.

17. Iqbal, M., Verrall, R. E., *J. Biol. Chem.* **1988**, 263, 4159.

18. (a) Shahidi, F., Farrell, P, G., *J. Chem. Soc. Faraday Trans. 1* **1978**, 74, 858. (b) Shahidi, F., *J. Chem. Soc. Faraday Trans. 1,* **1980**, 76, 101. (c) Shahidi, F., Farrell, P, G., *J. Chem. Soc. Faraday Trans. 1* **1981**, 77, 963. (d) Sinha, B., Dakua, V. K., Roy, M. N., *J. Chem. Eng. Data* **2007**, 52, 1768. (d) Millero, F. J., Lo Surdo, A., Shin, C., *J. Phys. Chem.* **1978**, 82, 784.

19. Zhao, H., *Biophys. Chem.* **2006**, 122, 157.

20. (a) Leslie, T. E., Lilley, T. H., *Biopolymers* **1985**, 24, 695.(b) Hedwig, G. R., Hoiland, H., *Phys. Chem. Chem. Phys.*, **2004**, 6, 2440.

21. Hoiland, H., Holvik, H. J. *Solution Chem.* **1978**, 7, 587.

22. Shahidi, F., Farrell, P. G., Edward, J. T. J. *Solution Chem.* **1976**, 5, 807.

23. (a) Yoshihashi, H., Ogawa, T., Shiio, H., *J. Am. Chem. Soc.* **1955**, 77, 4980. (b) Shiio, H., *J. Am. Chem. Soc.* **1958**, 80, 70.

24. Gupta, S., Moulik, S. P., *Can. J. Chem.* **1989**, 67, 356.

25. (a) Kuntz, I. D., *J. Am. Chem. Soc.* **1971**, 93, 514. (b) Kuntz, I. D., *J. Am. Chem. Soc.* **1971**, 93, 516.

26. (a) Deiber, J. A., Piaggio, M. V., Peirotti, M. B., *J. Sep. Sci.* **2008**, 31, 548. (b) Deiber, J. A., Piaggio, M. V., Peirotti, M. B., *Electrophoresis* **2007**, 3658.

27. Ifft, J. B., Hollenberg, J. L., *J. Phys. Chem.* **1982**, 86, 1938.

CHAPTER 5

Investigating Other Structural Features Necessary for Potent RI Activity of a C-linked AFGP Analogue

5.1. Introduction

Previously reported structure activity relationship (SAR) studies of antifreeze proteins and glycoproteins have been based exclusively on thermal hysteresis activity.[1, 2, 3, 4] Early SAR studies of AFGP were carried out by Feeney and DeVries,[1] and the results of these studies were re-visited by Nishimura and colleagues[4] in 2004. Based upon these results, it was concluded that the observed TH activity is dependent upon: 1) conformation, 2) N-acetyl group of the disaccharide, 3) configuration of the O-glycosidic linkages between sugars and peptide chain, and 4) the β-methyl group of the threonyl residue. Unfortunately, RI activity was not assessed in either of these studies and consequently the key structural features necessary for RI activity are unknown.

Additional SAR studies have been conducted in our laboratory where a library of C-linked AFGP analogues was synthesized and tested for TH, DIS and RI activity.[5, 6] Some of these analogues possessed custom tailored antifreeze activity (no TH and potent RI activity). From these studies structural moieties responsible for RI activity were identified to be: 1) a galactose residue (monosaccharide), 2) a glycosylated C-linked ornithine or C-linked serine side chain, and 3) glycine residues in the polypeptide backbone (i.e. a highly flexible peptide backbone).

In the past few decades many efforts have been dedicated to elucidate the secondary structure of AFGP8.[1, 7, 8, 9] Nuclear magnetic resonance (NMR) and circular dichroism (CD)

spectroscopy have been employed in the conformation study of AFGP8. The NMR studies imply that random unordered conformations predominate in AFGP8.[8] In the supercooled state, the amount of ordered structure increases especially that of β-turns. In the presence of ice, percentage of the β-turns further increases indicating a high degree of molecular order. This dynamic flexibility of AFGP8 in various states of water is thought to be essential for maximizing ice binding.[10] More recent NMR and CD studies suggest that AFGPs adopt a left handed polyproline type II (PP-II) helix.[4] It is believed that the amphiphilic nature of AFGP in PP-II conformation is important for antifreeze activity. Moreover, CD spectrum of a random coil conformation is very similar to that of PP-II and the two solution structures of AFGP cannot be easily distinguished from one another.[11] As a consequence it is apparent that the short glycopeptide adopts a variety of secondary structures in solution.

In order to further probe the structure activity relationship between RI and our C-linked AFGP analogues, several other analogues were prepared. These include: [Orn(Gal)-Ala-Ala]$_4$Gly, [Orn(GalNAc)Gly-Gly]$_4$Gly, [Orn(Gal-GalNAc)Gly-Gly]$_4$Gly, and [Orn(Glc-GalNAc)-Gly-Gly]$_4$Gly. The SAR studies of these glycopeptides will assess the importance of L-alanine residues, C-2 N-acetyl moiety of the carbohydrate as well as β-D-Gal-(1-->3)-D-GalNAc and β-D-Glc-(1-->3)-D-GalNAc disaccharide motifs on RI activity, hydration, and conformation.

5.2. Retrosynthethic Analysis of C-linked [Orn(Gal)-Ala-Ala]$_4$Gly and [Orn(GalNAc)-Gly-Gly]$_4$Gly Analogues

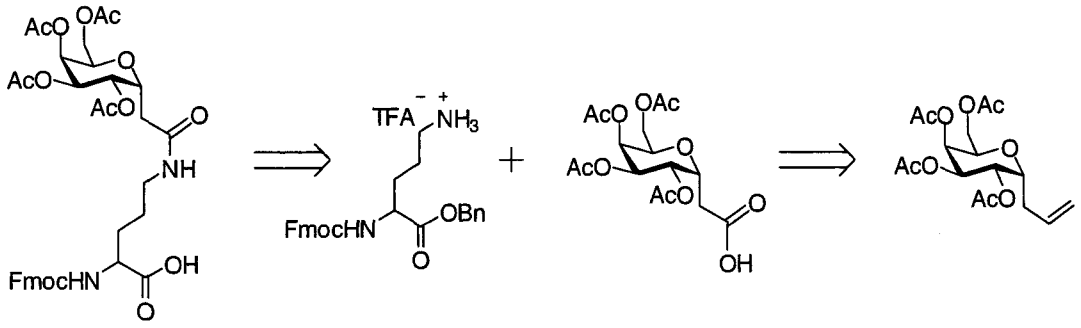
The [Orn(Gal)-Ala-Ala]$_4$Gly analogue was designed to examine the effect of L-alanine residues on antifreeze activity. The antifreeze activity results of this analogue will be compared

to that of the previously synthesized [Orn(Gal)-Gly-Gly]$_4$Gly, the most potent of the monosaccharide analogues. It has been speculated that the hydrophobic L-alanine residues are important for potent antifreeze activity.[4] Moreover, changing the peptidic backbone from predominantly glycines to alanines modifies our C-linked analogues towards the native AFGP composition.

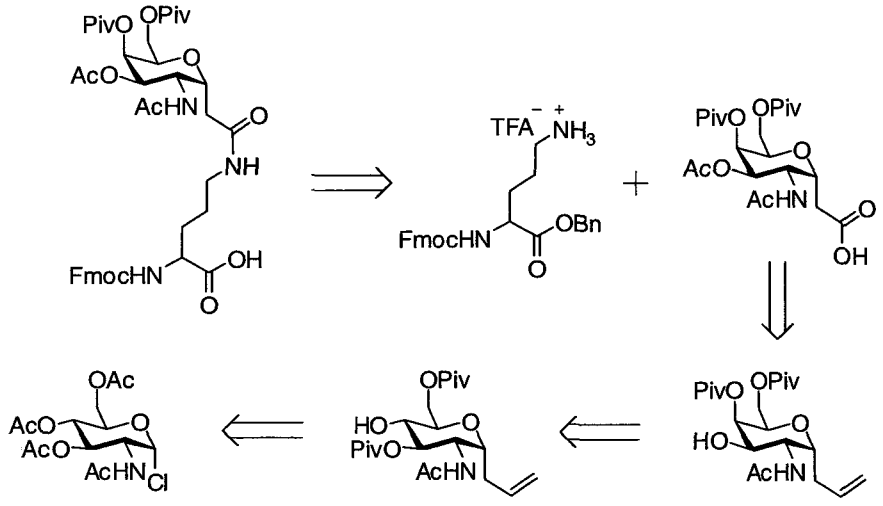
To investigate the effect of C-2 N-acetyl moiety of the carbohydrate on TH and RI activity, [Orn(GalNAc)-Gly-Gly]$_4$Gly analogue was designed, and its antifreeze activity results will also be compared to that of [Orn(Gal)-Gly-Gly]$_4$Gly. It is known that the conformation and hydration of the glycopeptide can be altered by hydrogen bonding between the nitrogen of the amide bond at C-2 and the polypeptide backbone.[12]

Consequently, to verify the effect of L-alanine residues on antifreeze activity, the activity of [Orn(GalNAc)-Gly-Gly]$_4$Gly analogue will be compared to that of the previously synthesized [Orn(GalNAc)-Ala-Ala]$_4$Gly. The antifreeze activity results of [Orn(GalNAc)-Ala-Ala]$_4$Gly will also be compared to that of [Orn(Gal)-Ala-Ala]$_4$Gly to re-evaluate the effect of C-2 N-acetyl moiety of the carbohydrate on antifreeze activity of our C-linked AFGP analogues.

The Gal-ornithine and GalNAc-ornithine building blocks will be assembled by coupling of the ornithine derivative with the corresponding carboxylic acid of galactose or N-acetyl galactose (Scheme 5.1. and 5.2.).[5c, 6] The N-acetyl galactose carboxylic acid will be prepared from the allylated N-acetyl galactosamine derivative following C-3 hydroxyl acetylation and consequent oxidation. The allylated N-acetyl galactosamine derivative will be obtained from 2-acetamido-2-deoxy-3,4,6-tri-O-acetyl-α-D-glucopyranosyl chloride by heat induced radical allylation, following protecting group manipulation and intramolecular inversion of the C-4 hydroxyl.[5a, 13]



Scheme 5.1. Retrosynthetic analysis of the C-linked Gal-ornithine building block.



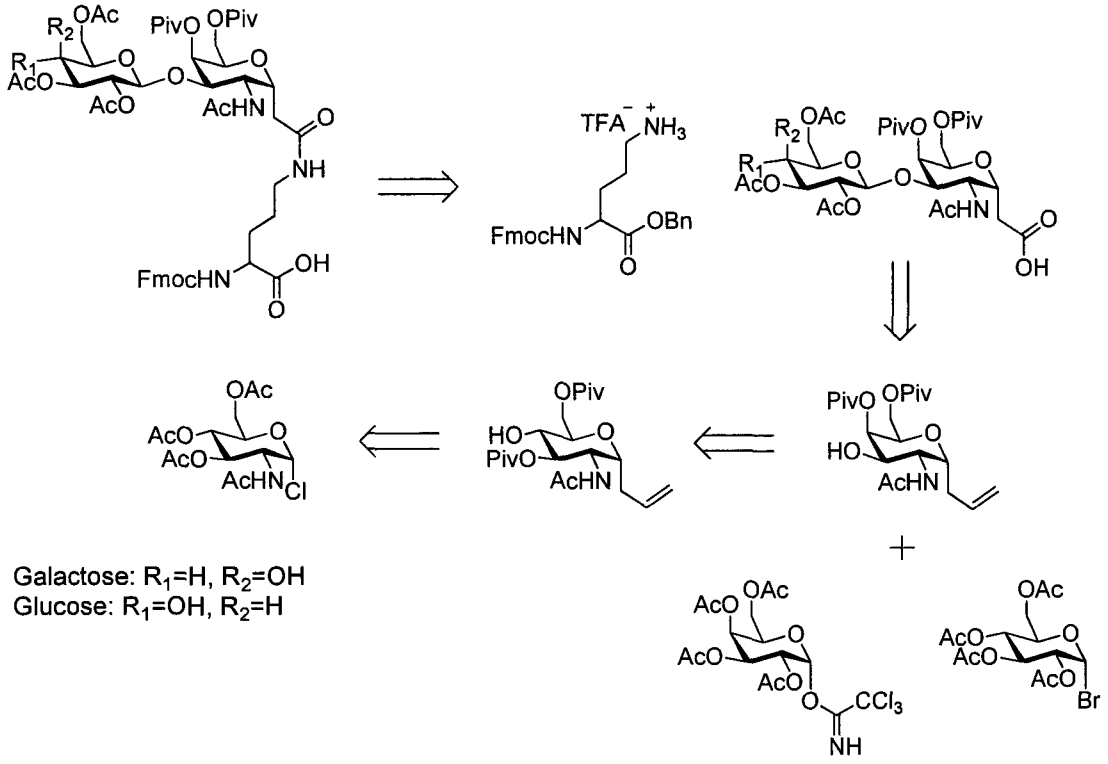
Scheme 5.2. Retrosynthetic analysis of the C-linked GalNAc-ornithine building block.

5.3. Retrosynthetic Analysis of C-linked [Orn(Gal-GalNAc)-Gly-Gly]$_4$Gly, [Orn(Glc-GalNAc)-Gly-Gly]$_4$Gly and [Orn(Gal-Gal)-Gly-Gly]$_4$Gly Analogues

The Gal-GalNAc disaccharide, which is key structural feature of the native AFGP8,[4, 14] will be incorporated into the [Orn(carbohydrate)-Gly-Gly]$_4$Gly scaffold to compare its antifreeze activity to that of the native AFGP8. These results will assess the vitality of the disaccharide

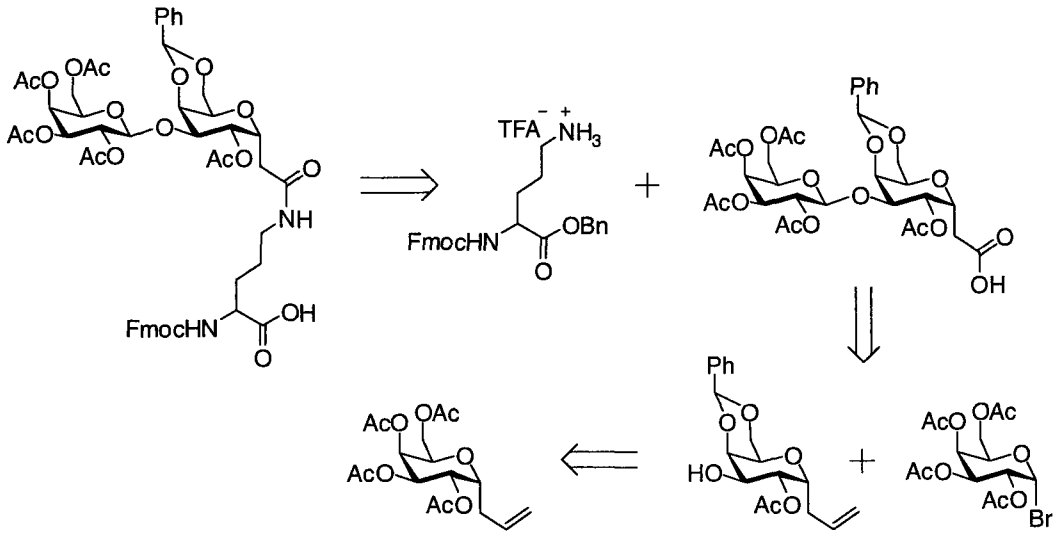
motif in: (a) binding with the ice lattice hence inducing TH, and (b) RI activity of the synthetic analogues. The Glc-GalNAc disaccharide was included to confirm the importance of the terminal carbohydrate residue in the native disaccharide, on RI activity. Finally, the [Orn(Gal-Gal)-Gly-Gly]$_4$ analogue is of interest because it can be directly compared to [Orn(Gal)-Gly-Gly]$_4$ to determine the effect of the disaccharide versus monosaccharide, and to further verify if an increase in overall carbohydrate concentration is proportional to the increase in RI. This concept can be further strengthened by comparing data between the [Orn(Gal-Gal)-Gly-Gly]$_4$ analogue containing the disaccharide and the divalent analogue [Lys(Gal$_2$)-Gly-Gly]$_4$ (Figure 2.6.).[5c] Lastly, the function of the N-acetyl group at C-2 will be again investigated by comparing the antifreeze activities of [Orn(Gal-Gal)-Gly-Gly]$_4$ and [Orn(Gal-GalNAc)-Gly-Gly]$_4$.

The preparation of Gal-GalNAc-ornithine and Glc-GalNAc-ornithine building blocks is outlined in (Scheme 5.3.). The amide bond can be constructed as described previously[5a, c, d] and several approaches toward glycosidic bond formation can be considered. Of the simpler synthetic methods is Koenigs-Knorr which requires a glycosyl acceptor and glycosyl bromide. In the presence of the acceptor, activation of the donor affords the new glycosidic linkage.[15] Another synthetic method often used involves trichloroacetimidates (TCAs) as developed by Schmidt.[16] In this approach, activation of a TCA donor with a reactive Lewis acid results in formation of the oxocarbenium ion and subsequent glycosidic bond formation. The requisite glycosyl acceptor (allylated GalNAc) can be prepared from 2-acetamido-2-deoxy-3,4,6-tri-O-acetyl-α-D-glucopyranosyl chloride as outlined in the retrosynthetic analysis of the GalNAc-ornithine building block.



Scheme 5.3. Retrosynthetic analysis of the C-linked Gal-GalNAc-ornithine and Glc-GalNAc-ornithine ornithine building blocks.

Synthesis of the Gal-Gal building block requires preparation of the Gal-Gal carboxylic acid (Scheme 5.4.). The building block synthesis and glycosidic bond formation can be accomplished though the abovementioned methods.[15, 16] The acceptor in this case can be prepared via protecting group manipulation from the corresponding allylated galactose tetraacetate.



Scheme 5.4. Retrosynthetic analysis of the C-linked Gal-Gal-ornithine building block.

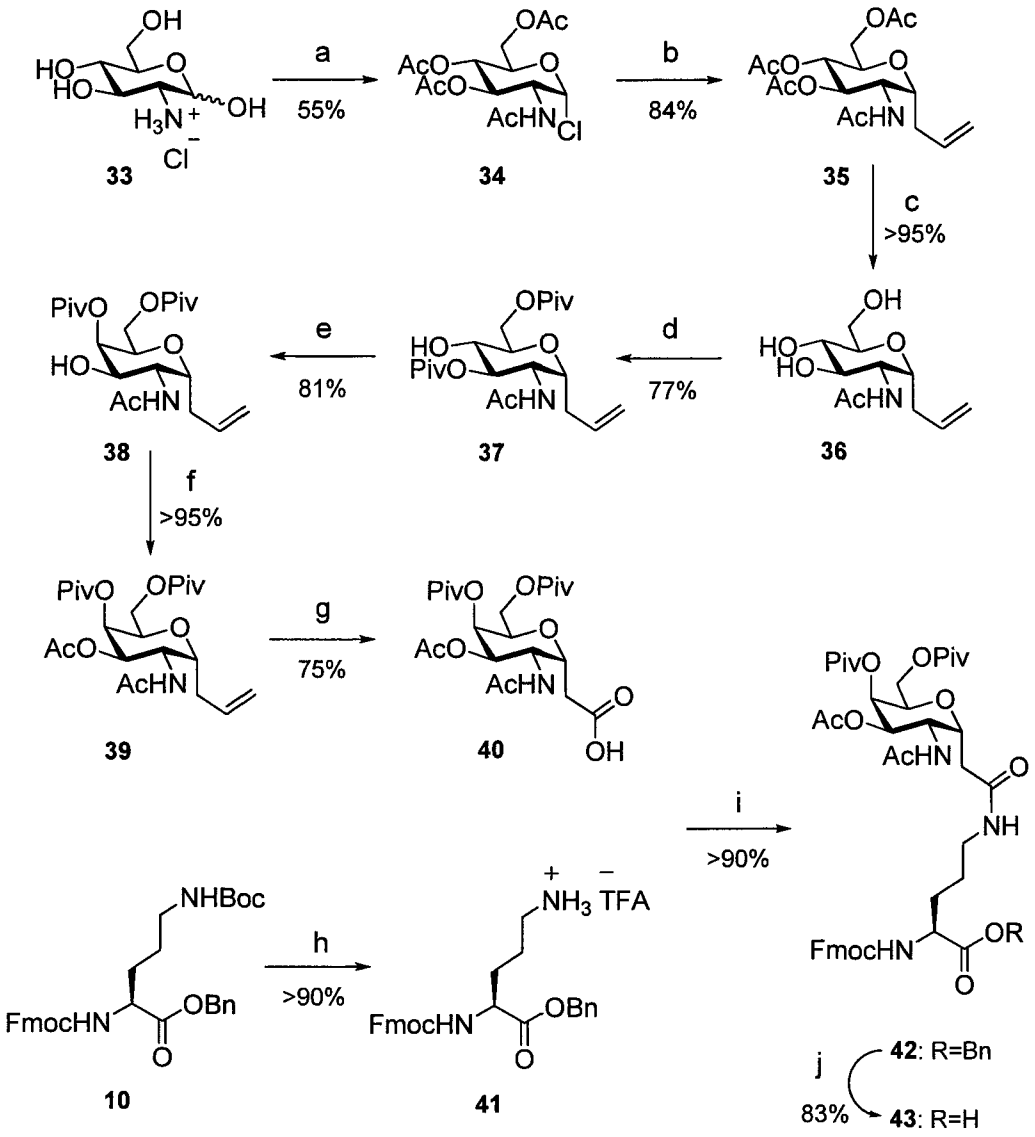
5.4. Synthesis of the C-linked AFGP Analogues 63-66

The *C*-linked AFGP analogues **63-66** were prepared in a linear manner whereby a particular building block was assembled into the desired glycopolymer using automated solid phase synthesis (Section 3.2.). In order to obtain these analogues, their key building blocks had to be prepared by coupling of the C-linked carbohydrate with an orthogonally protected ornithine. Since the general focus of this thesis is the modification of the carbohydrate component of C-linked AFGP analogues, the synthesis of the necessary C-linked carbohydrates was the major challenge in obtaining the designed glycoconjugates.

5.4.1. Synthesis of the C-linked Building Blocks 16 and 43

Synthesis of the C-linked galactose-ornithine building block **16** was already described in Chapter 3 (Section 3.5). Thus, we turn our attention towards the next target, C-linked N-acetyl galactosamine-ornithine building block. Many methods have been developed to synthesize C-

linked N-acetyl galactosamine derivatives however most of them are still lengthy and inefficient.[17] The azido nitration chemistry developed by Canadian chemist Raymond Urgel Lemieux,[18] and its variations such as azido cholorination[19a] or azido selenation[19b, c] are often utilized in a synthetic approach. The resulting derivatives include acetylenic,[20a] allylic,[20b] allenic,[20b] cyano[20c] and other groups C-lined to the anomeric carbon of the N-acetyl galactosamine. Nevertheless, the desired C-linked N-acetyl galactosamine can also be prepared via direct Keck allylation[21] of the bromo- or chloro-N-acetyl-galactosamine. However, poor yields accompanied by side products shift our focus towards clever synthetic approach developed by Ben et al.[13c] The novel synthetic route consists of Keck allylation of N-acetyl glucosamine derivative following inversion of the stereochemistry at C-4 of the pyranose (Scheme 5.5.).



Conditions: (a) $HCl_{(g)}$/AcCl; (b) Allyltributyltin, AIBN, THF, reflux; (c) NaOMe, MeOH; (d) PivCl, py; (e) Tf_2O, py, CH_2Cl_2 then H_2O; (f) Ac_2O, py, CH_2Cl_2; (g) $RuCl_3 \cdot H_2O$, $NaIO_4$, $Cl_4/CH_3CN/H_2O$ (2:2:3); (h) TFA/CH_2Cl_2; (i) HBTU, DIPEA, CH_2Cl_2; (j) H_2/Pd-C, EtOAc/EtOH (4:1).

Scheme 5.5. Synthesis of the C-linked building block **43**.

Commercially available D-glucosamine was converted to 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-α-D-glucopyranosyl chloride (55% yield) via treatment with acetyl chloride saturated

with HCl gas.[22] Keck allylation of compound **34** produced **35** in 84% yield as 12:1 α/β mixture of diastereomers.[21] No major side products were detected and the yield of this reaction was excellent when compared to the same reaction with 2-acetamido-3,4,6-tri-O-acetyl-2-deoxy-α-D-galactopyranosyl chloride as starting material.[5a] The acetate protecting groups were removed (>95% yield) using sodium methoxide solution.[23] Using 6 equivalents of pivaloyl chloride, C-6 and C-3 hydroxyls of **36** were selectively protected to produce **37** in 77% yield.[5a, 13b] Triflic anhydride was used to convert the hydroxyl at C-4 to a leaving group which was easily displaced by the carbonyl of the pivaloyl group at C-3 to afford dioxolenium ion intermediate. The intermediate was then hydrolyzed to afford **38** in overall 81% yield. The C-3 hydroxyl group was protected using pyridine and acetic anhydride to furnish **39** in quantitative yield. This derivative was then oxidized to produce the desired carboxylic acid **40** (75% yield). Fully protected ornithine was activated for the next step by cleaving its acid labile tert-butoxycarbonyl (t-Boc) group (>90% yield). The generated trifluoroacetate-ornithine salt **41**[5a, c, e, f] was coupled with **40** using HBTU and DIPEA. Following hydrogenolysis using hydrogen gas and palladium on carbon, C-linked N-acetyl galactosamine-ornithine building block **43** was obtained in 75% yield.

5.4.2. Glycosidic Bond Formation

The glycosidic bond formation dates back to over a century ago when carbohydrates were reacted together in a presence of acid.[24] General reaction occurs between glycosyl donor and glycosyl acceptor in the presence of a promoter. While to this date it is very difficult to predict the reactivity of the glycosyl species, the reactivity is known to depend on the stereochemistry and substituents, particularly protecting groups, of the carbohydrate.[25]

In 1980s, Paulsen was the first to report that by changing protecting groups from benzyl ethers to benzyl esters reactivity of glycosyl donors increased several folds.[26] In fact, he has shown that carbohydrates with electronically passive protecting groups favor formation of a positive charge at the anomeric position and are therefore "armed". The "disarmed" sugars possess electron withdrawing protecting groups that disfavor formation of the positive charge at the anomeric position and consequently exhibit lower reactivity as glycosyl donors. This concept was utilized in synthesis of many polysaccharides however reactivity of carbohydrates is yet to be fully clarified.

Stereoselectivity of the glycosylation can be influenced by many factors (Figure 5.1.).[27] The intermediate that is generated after activation of a glycosyl donor is known as oxocarbenium ion and it is common to all coupling methods. The anion of the oxocarbenium ion-pair may shield one face of the complex that may then react in a stereoselective manner with a glycosyl acceptor. In addition, similar effect can be attained in a presence of a solvent with high dielectric constant or basic lone pair of electrons that can efficiently stabilize a positive charge. Another method of controlling stereoselectivity of glycosylation reaction is though neighbouring group participation, also termed anchimeric effect. The effect arises when the anomeric center of the oxocarbenium ion is stabilized by a neighbouring group at C-2. For instance, carbonyl of an ester can easily stabilize the oxocarbenium ion by forming a cyclic dioxolenium ion intermediate. This intermediate can either undergo stereoselective reaction to give the β anomer or an orthoester. Since the reaction is in equilibrium and it is reversible, steric and electronic effects will dictate its outcome. Without the anchimeric effect, anomeric effect dominates and the glycosylation reaction gives the α anomer as a major product. The stereoselectivity outcome also depends on the type of coupling reaction.

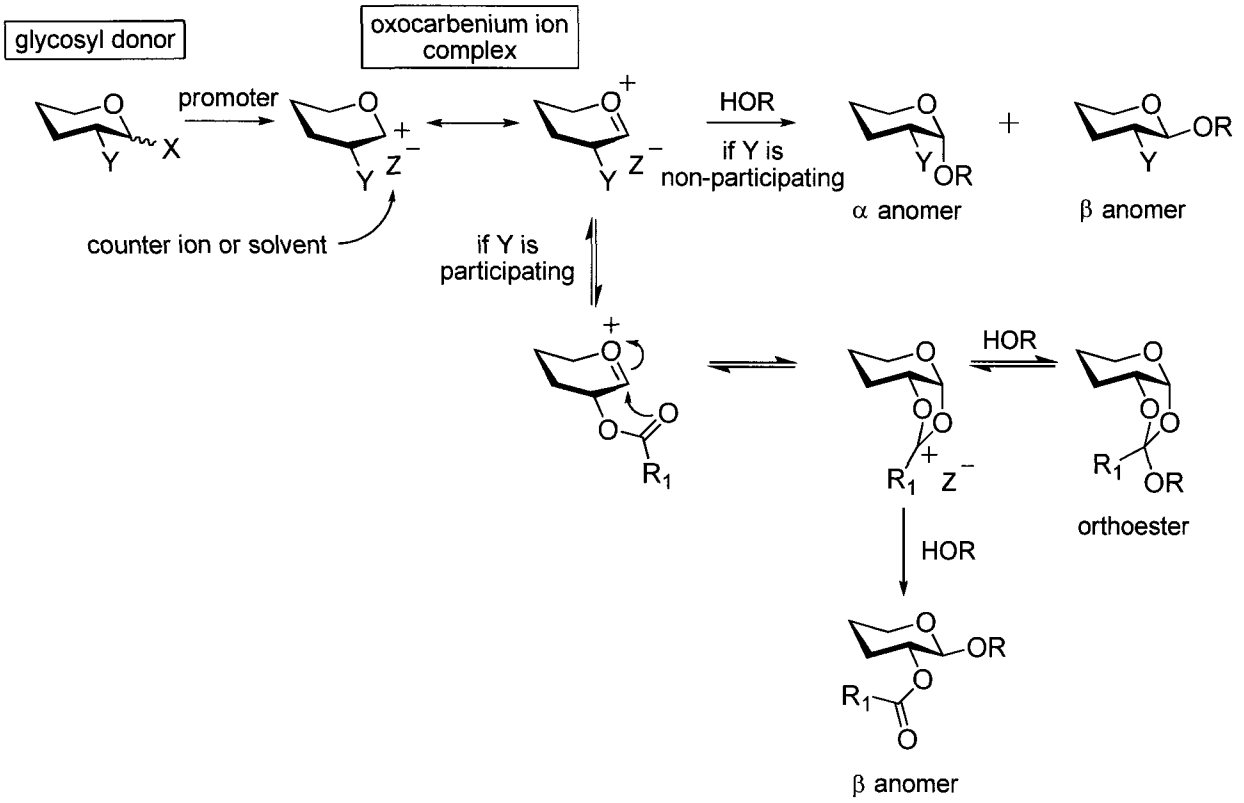


Figure 5.1. Various features of a glycosylation reaction.

To date, many glycosylation methods have been developed[28] and their preference is generally based on the overall synthetic plan of a target molecule. Frequently, the desired stereoselectivity, reaction conditions, compatibility of protecting groups with those conditions and reaction efficiency determine the choice of the glycosylation method. In addition, one of the most important factors affecting its preference is the experience of the researcher with a particular technique. Some of the common glycosyl donors and their activators are presented in Table 5.1

Table 5.1. Common glycosyl donors and their activators.

Glycosyl Donor 	**Activator**	**Developed by**
L=OH	Acid	Fisher[24]
L=Br	Ag_2CO_3, Et_4NBr, $Hg(CN)_2$, $HgBr_2$, AgOTf, Ag-silicate	Koenigs-Knorr[15a]
L=Cl	$Hg(CN)_2$, $HgBr_2$, AgOTf	
L=F	$SnCl_2$-$AgClO_4$, AgOTf-$HfCp_2Cl_2$, AgOTf-$SnCl_2$, Tf_2O-$HfCp_2Cl_2$, Tf_2O	Mukaiyama[29a, b], Nicolaou[29c]
L=OAc	BF_3•OEt_2, $SnCl_4$, TMSOTf	-
L=TCA	BF_3•OEt_2, TMSOTf, TESOTf	Sinay[30], Schmidt[16]
L=SR	I_2, DMTST, TfOH or AgOTf-NIS, IDCP, MeOTf	Garegg[31a], van Boom[31b], Fraser-Reid [31c]
L=SOR	Tf_2O	Kahne[32]
L=SeR, TeR	I_2, DMTST, TfOH or AgOTf-NIS, IDCP, MeOTf	Pinto[33]
L=$O(CH_2)_3CH=CH_2$	IDCP, TfOH-NIS, AgOTf-NIS, TESOTf-NIS	Fraser Reid[34]
glycal epoxides	$ZnCl_2$, other Lewis acids	Danishefsky[35]
L=$OP(OR)_2$	TMSOTf, other Lewis acids	Schmidt[36a, b], Wong[36c]

One of the oldest coupling procedures was discovered and reported in 1901 by Wilhelm Koenigs and Edward Knorr.[15] In the original work, 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl bromide was reacted with alcohols in the presence of silver carbonate to generate a glycosidic linkage. Over the century, a number of improvements have been reported. For instance, alternative promoters such as mercuric bromide/mercuric oxide, mercuric cyanide, silver triflate are now used in this reaction. Silver impregnated silica is also utilized as a promoter to impair the reaction of silver with bromide that drives the reaction forward.[37] As a consequence, the heterogeneous catalyst only weakens the carbon-bromide bond creating a partial positive charge at the anomeric position. Nucleophilic addition of a glycosyl acceptor to the anomeric center expels silver-bromide complex. Since glycosyl halides in the Koenigs-Knorr reaction possess α stereochemistry, this SN2-like reaction results in a β glycosidic bond. Nevertheless, other factors such as anchimeric effect and choice of solvent affect this reaction.

Another widely used glycosidation method is the trichloroacetimidate (TCA) method which was first used by Sinay[30] and further developed by Schmidt[16]. The technique involves synthesis of glycosyl TCAs that are used as glycosyl donors. Anomerically pure glycosyl TCAs can be obtained by treatment of 1-hydroxy-glycoside with a base in a presence of trichloroacetonitrile. Stronger bases such as NaH or DBU and longer reaction times favour formation of the thermodynamic product, the α anomer. The stronger base is able to deprotonate the imine resulting in a reversible reaction that 'funnels' towards the most stable product (Figure 5.2). The β anomer can be obtained with weaker bases such as Na_2CO_3 via kinetic process.

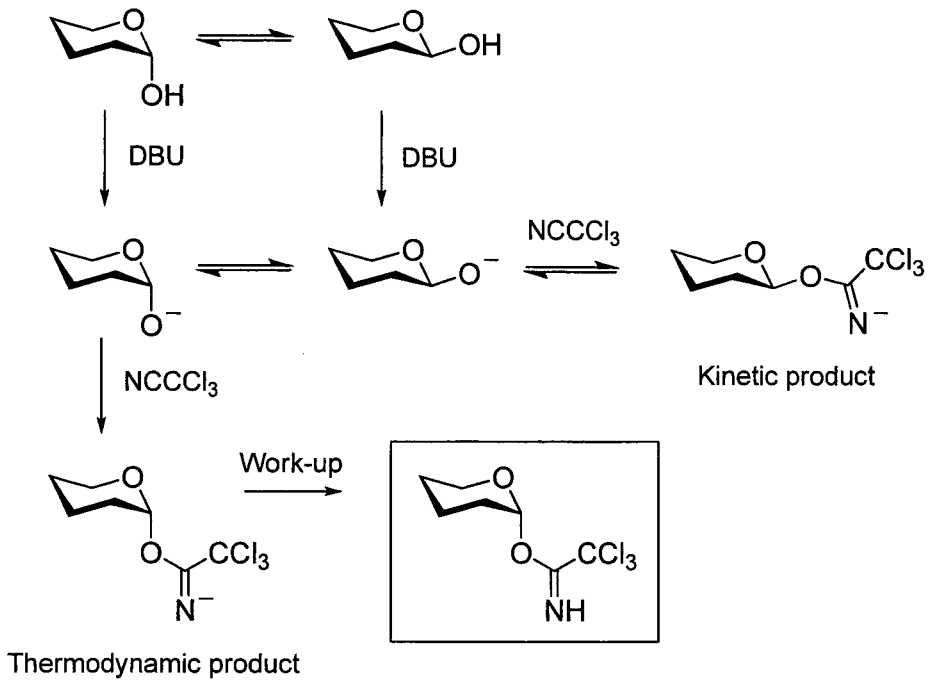


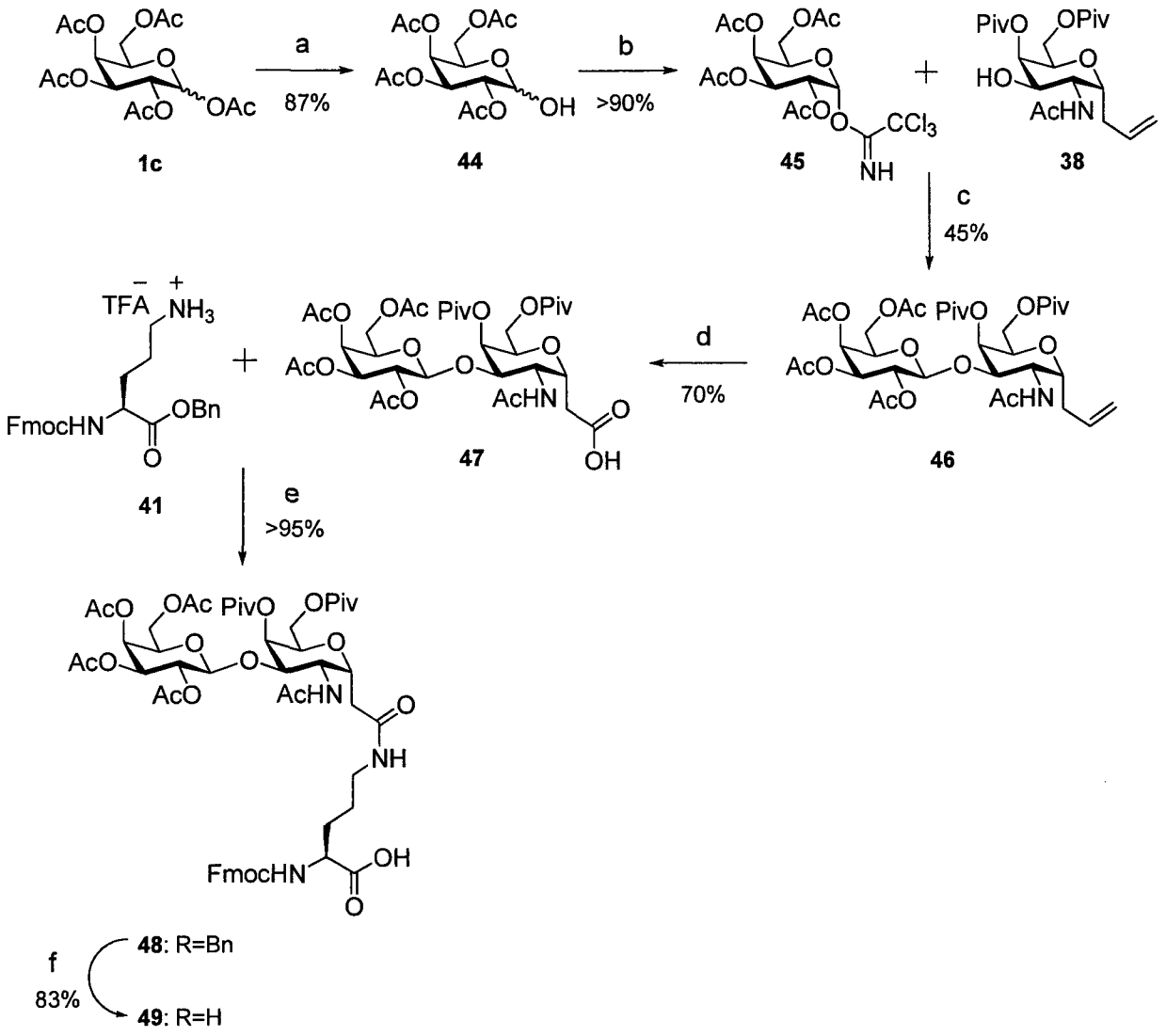
Figure 5.2. Formation of the thermodynamic TCA product with DBU.[38]

In trichloroacetimidate glycosylation reactions, glycosyl TCA donors which contain non-participating group at C-2 when treated with a mild Lewis acids such as $BF_3 \cdot OEt_2$ react in SN2-like fashion with glycosyl acceptors. When stronger Lewis acids, like TMS-OTf, are used the reaction proceeds via oxocarbenium ion and its selectivity can then be controlled via solvent choice. In CH_2Cl_2 or ether α-glycosidic linkage is generated. In CH_3CN at low temperature β-glycosidic linkage is obtained.[39] When glycosyl TCA donors with a participating group at C-2 are used, anchimeric effect determines stereochemistry of the product.

5.4.3. Synthesis of the C-linked Building Blocks 49 and 53

From among many coupling methods, well established TCA chemistry was chosen to be explored in the synthesis of Gal-GalNAc building block **49** (Scheme 5.6.). Peracetylated

galactose was selectively deprotected (87% yield) at the anomeric center using 1.1 equivalent of hydrazine.[40] Next, compound **44** was reacted with tricholoroacetonitrile in a presence of a strong base DBU to generate α-TCA-glycoside in >90% yield.[41] To a solution of glycosyl donor **45** and glycosyl acceptor **38** in CH_2Cl_2, mild Lewis acid (BF_3•OEt_2) was added dropwise to generate the desired disaccharide.[41, 42] However, due to the poor reaction yields even at low temperatures this coupling method was abandoned. Before changing glycosylation reactions, sequence of reactants addition was revised. In 1991 Schmidt published that the TCA donors in a prolonged presence of Lewis acids readily decompose. Thus, glycosyl acceptor should not be added to a solution of glycosyl donor and an activator. Inverse addition was discovered in which glycosyl acceptor was premixed with a promoter and then glycosyl donor was added.[16d] This addition resulted in improved yields and minimized side reactions. Thus, we hypothesized that if **38** is a weak nucleophile and does not promptly react with **45**, accumulation of BF_3•OEt_2 will lead to destruction of **45** and no desired product will be detected. Consequently, inverse addition was implemented into our synthesis and disaccharide **46** was obtained in 45% yield after intensive purification procedures. Stereoselectivity was well controlled due to the SN2 like reaction and the anchimeric effect of the C-2 acetate. The 1,2-trans glysoside (β anomer) was obtained as a major diasteromer whereas the α anomer was not isolated. Ruthenium catalyzed oxidation yielded **47** in 70%.[43] This derivative was then coupled to activated ornithine **41** and subsequent hydrogenolysis gave our desired Gal-GalNAc building block **49** in 80% yield.

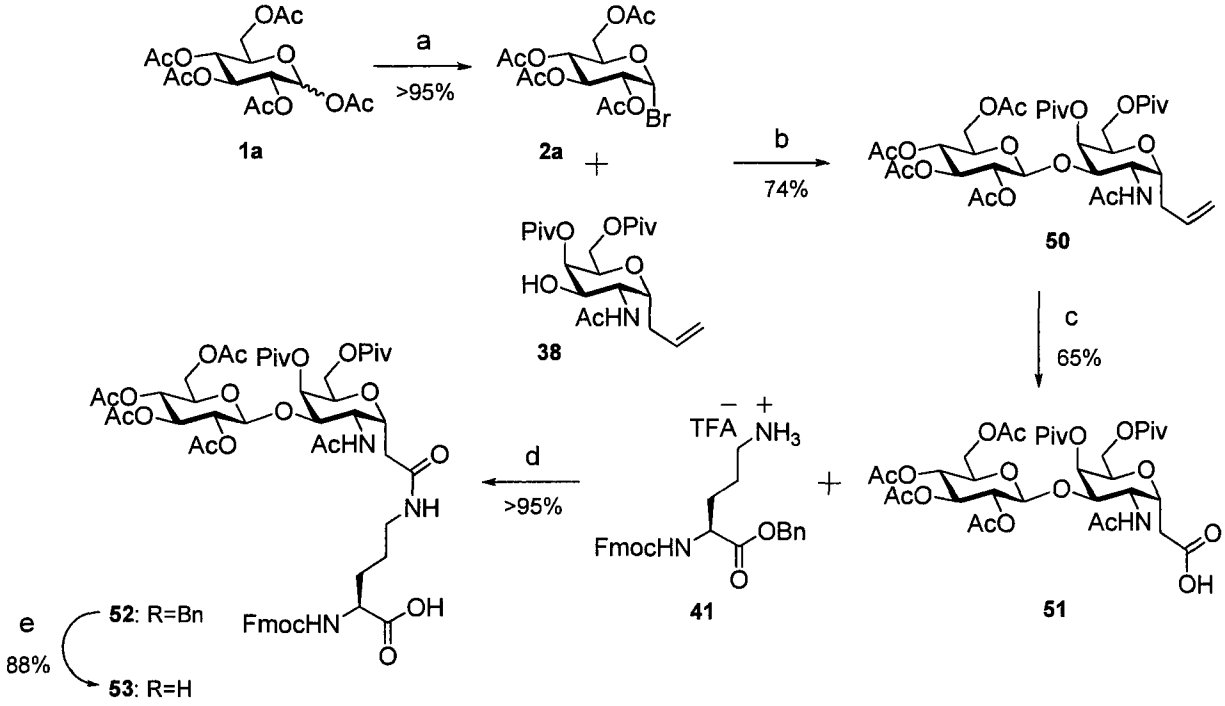


Conditions: (a) H_2NNH_2, DMF; (b) $NCCCl_3$, DBU, CH_2Cl_2; (c) $BF_3 \cdot OEt_2$, CH_2Cl_2/hexanes, 0°C to RT; (d) $RuCl_3 \cdot H_2O$, $NaIO_4$, $CCl_4/CH_3CN/H_2O$ (2:2:3); (e) HBTU, DIPEA, CH_2Cl_2; (f) H_2/Pd-C, EtOAc/EtOH (4:1).

Scheme 5.6. Synthesis of building block **49**.

The low yields in the TCA glycosidation reaction caused us to re-think the approach for the Glc-GalNAc moiety and a Koeings Knorr reaction was employed (Scheme 5.7.).[15] Glycosyl

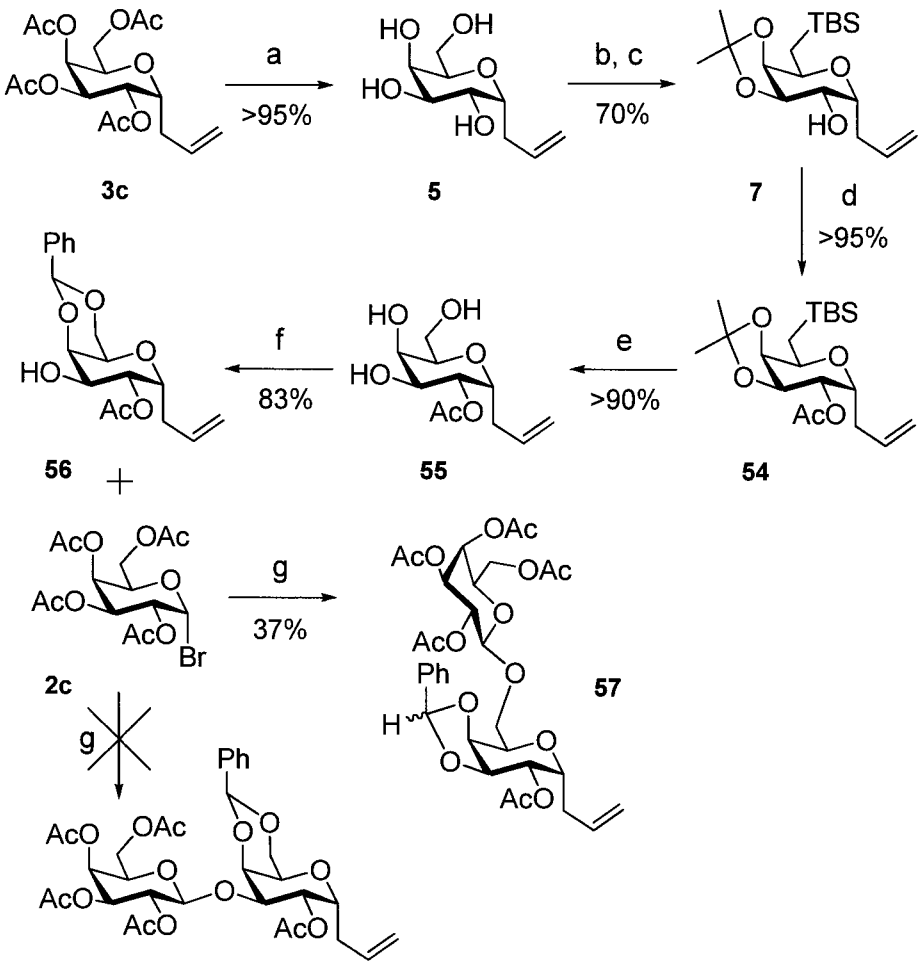
acceptor **38** was reacted with **2a** using silver triflate as the promoter that was added to the reaction mixture portionwise as a solid. The 2,6-lutidine was used as a proton sponge and the reaction condition were kept anhydrous with 4Å MS. The desired disaccharide **50** with β glycosidic linkage was obtained in 74% yield after rigorous purifications. The α anomer was not isolated. Compound **50** was oxidized with RuO_4, generated in situ with $RuCl_3$ and $NaIO_4$, to carboxylic acid derivative **51** (65% yield).[43] Following **51** was coupled with ornithine **41** using our standard reaction conditions. Hydrogenolysis of the benzyl ester furnished Glc-GalNAc building block **53** in 84% yield over two reactions.



Conditions: (a) HBr/HOAc; (b) AgOTf, 2,6-Lutidine, 4Å MS, CH_2Cl_2; (c) $RuCl_3 \cdot H_2O$, $NaIO_4$, $CCl_4/CH_3CN/H_2O$ (2:2:3); (d) HBTU, DIPEA, CH_2Cl_2; (e) H_2/Pd-C, EtOAc/EtOH (4:1).

Scheme 5.7. Synthesis of building block **53**.

In order to synthesize the Gal-Gal building block, galactosyl acceptor had to be prepared (Scheme 5.8.). Protecting group manipulation on C-linked-allyl-galactose **3c** was performed to arrive with the desired compound. The C-2 hydroxyl group of derivative **7**, synthesized as described for C-linked talose AFGP analogue **22** (Chapter 3, Scheme 3.4), was acetylated in quantitative yield. Acid labile protecting groups were cleaved off with 80% aqueous acetic acid[44] and benzylidene acetal was installed[45] between C-4 and C-6 hydroxyls in overall 75% yield. Efficiently synthesized galactosyl acceptor **56** and galactosyl bromide **2c** were subjected to Koenigs-Knorr coupling conditions.[15] However, even after meticulous scan of the reaction conditions no product formation was detected.

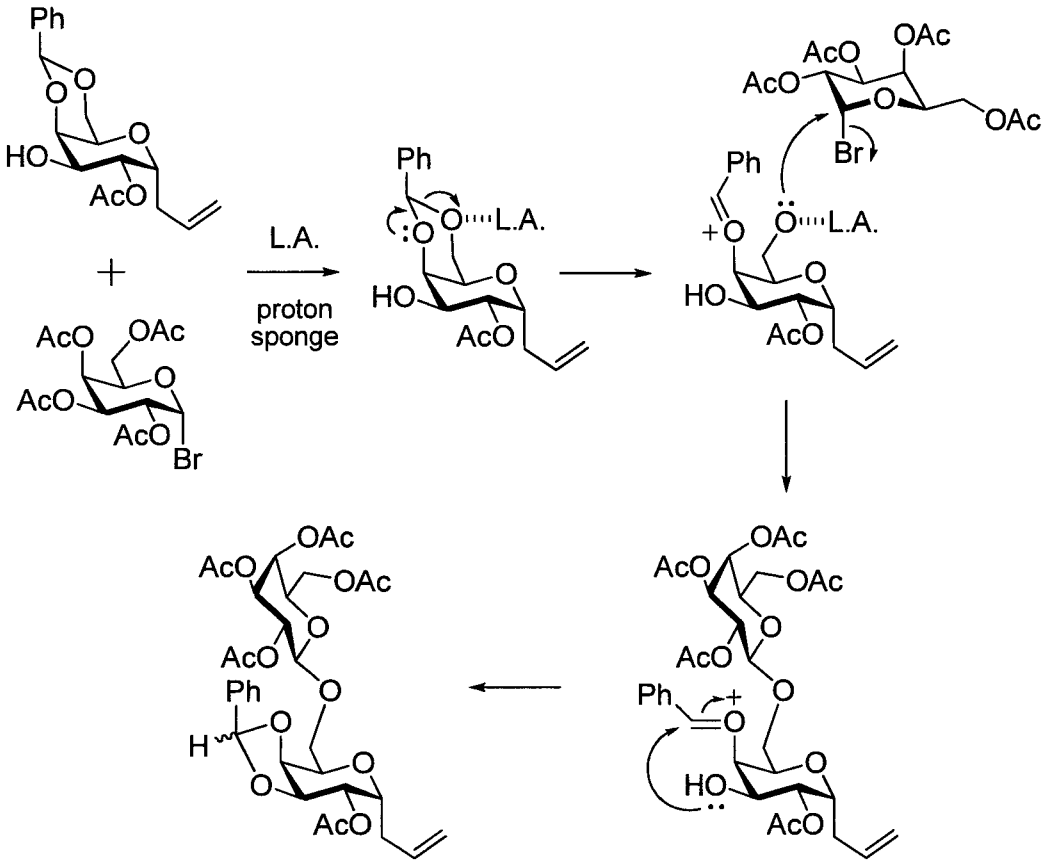


Conditions: (a) NaOMe, MeOH; (b) $Me_2C(OMe)_2$, cat. p-TSA, CH_3CN; (c) TBSCl, cat. DMAP, CH_2Cl_2/py; (d) Ac_2O, py, cat. DMAP, CH_2Cl_2; (e) 80% aque. AcOH, 80°C; (f) p-TSA, p-benzaldehyde dimethyl acetal, CH_3CN/DMF; (g) AgOTf, 2,6- Lutidine, 4Å MS, CH_2Cl_2.

Scheme 5.8. Attempted synthesis of C-linked Gal-Gal disaccharide.

The new product formed was a diasteromeric mixture of disaccharide **57** isolated in total 37% yield. Its structure was proposed by rationalizing the reaction mechanism (Scheme 5.9.). It was suggested that in presence of a Lewis acid benzylidine acetal opens exposing C-6 hydroxyl group.[46] The more reactive, primary hydroxyl group at C-6 of intermediate formed will now

react with the galactosyl donor. The secondary hydroxyl at C-3 will then close on to the activated benzaldehyde without any stereoselective preference resulting in the observed products.



Scheme 5.9. Proposed reaction mechanism for formation of disaccharide **57**.

For both diastereomers, stereochemistry of the 3,4-benzylidene acetal was verified via NOESY NMR experiments (Figure 5.3. and 5.4). As apparent from the NOESY spectrum in Figure 5.3, phenyl group protons clearly couple to hydrogens at C-3 and C-4. The methine proton of the benzylidine acetal detects C-1 and C-2 hydrogens of the pyranose. Thus, deduced from this data is a structure in which the methyne proton is situated above the C-linked pyranose

while the phenyl group is oriented away from the pyranose ring. Consequently, stereochemistry of the benzylidine acetal is assigned 'S' and designated structure is depicted in Figure 5.3. The NOSEY spectrum in Figure 5.4 shows benzylidine acetal of the other diastereoisomer with its 'R' stereochemistry. The protons of the aromatic group couple to hydrogens at C-1 and C-2, and although signal of the methyne proton overlaps with hydrogen of $-CH_2-CH=CH_2$, it detects hydrogens of C-3 and C-4.

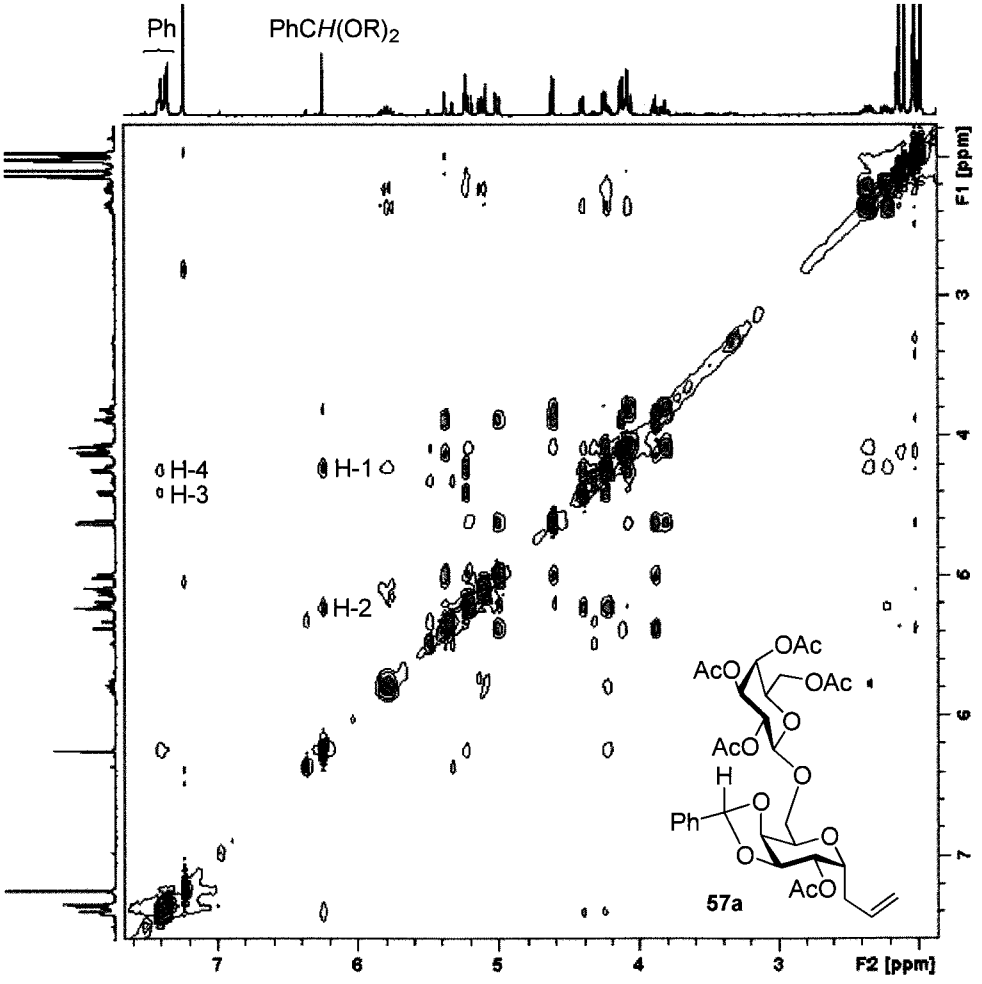


Figure 5.3. NOSEY spectrum of disaccharide **57a**.

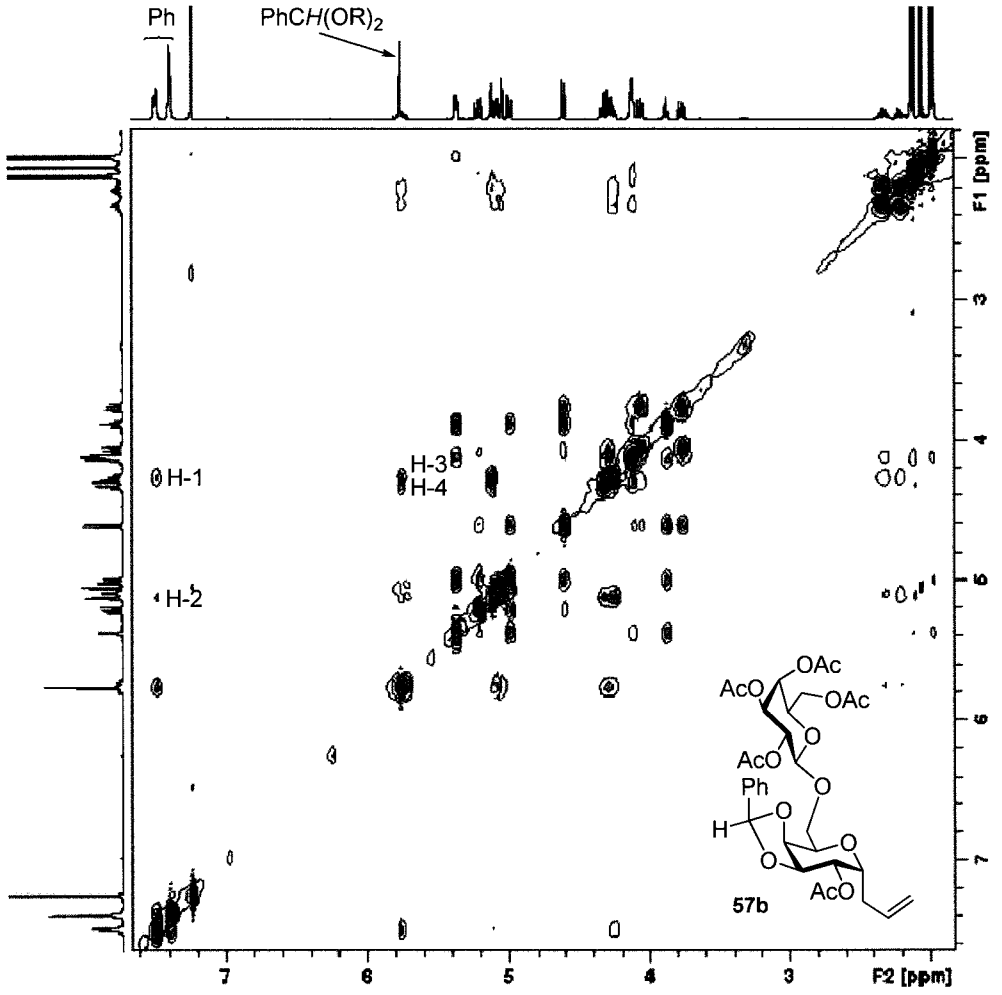
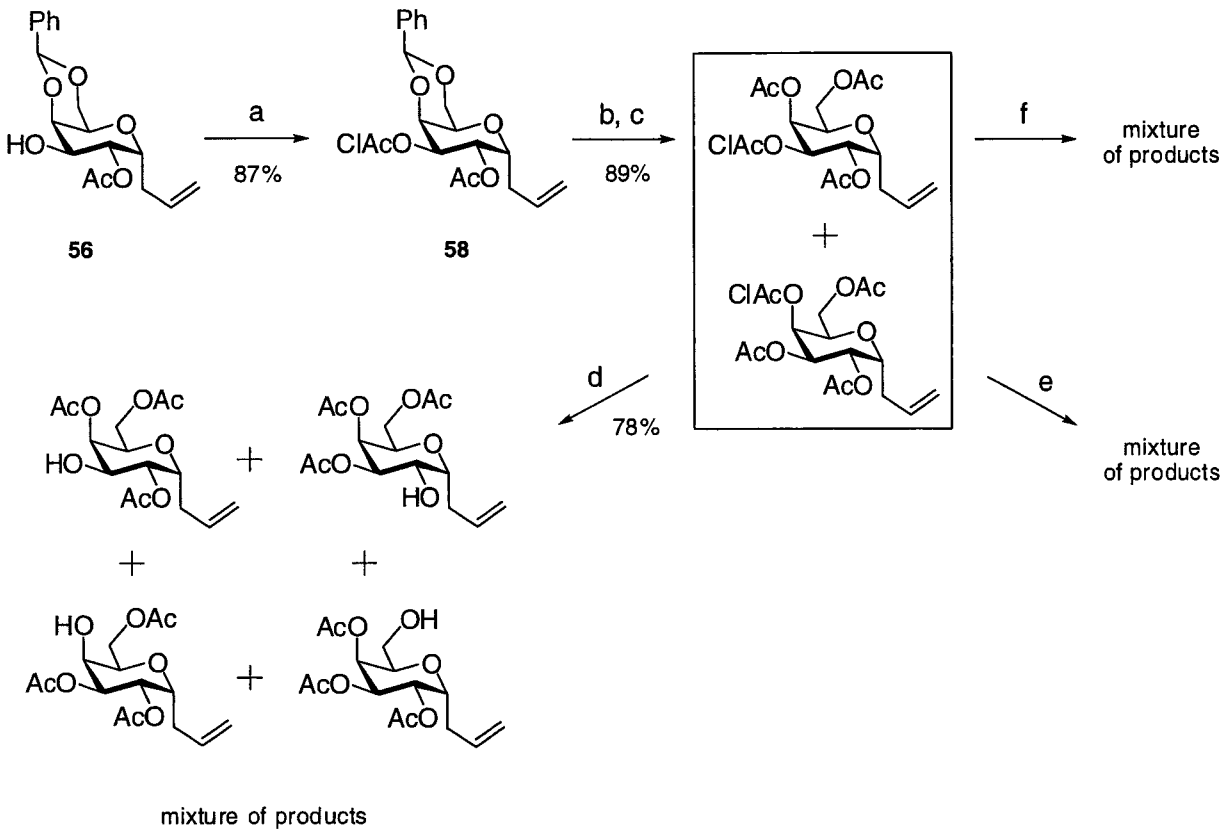


Figure 5.4. NOSEY spectrum of disaccharide **57b**.

In order to successfully synthesize the C-linked Gal-Gal building block and deviate from the acquired problems associated with opening of the acetal, galactosyl acceptor was further modified (Scheme 5.10). The C-3 hydroxyl group was protected with chloroacetate in 87% yield.[47] In subsequent step benzylidene acetal was cleaved under acidic condition[44] and using acetic anhydride, pyridine and catalytic amount of DMAP hydroxyls at C-4 and C-6 were successfully protected in 89% yield. Nevertheless all that could be isolated was an inseparable mixture of acetylated products. TLC analysis showed only one spot while in ^{1}H NMR spectrum

all signals overlapped. In the ^{13}C NMR spectrum each carbon signal was accompanied by another one next to it indicating presence of two compounds with very similar structures. Since the compounds could not be separated, resolution of the mixture was attempted after next step of the synthesis. Thus, chloroacetate was selectively cleaved using thiourea and 2,6-lutidine.[47] Following rigorous purification via flash column chromatography, again TLC showed one spot however this time ^{1}H NMR analysis revealed more than two compounds. Other methods of deprotecting chloroacetate also resulted in an inseparable mixture of more than two compounds.[48] The obtained mixture of products was rationalized by migration of the protecting groups. Chloroacetates, acetates and other ester-like protecting groups are known for their migration under basic conditions.[49] We observed that as the exposure of the compound to basic conditions increased, the number of regioisomers also increased. Consequently, this approach was not suitable for synthesis of the C-linked galactosyl acceptor.

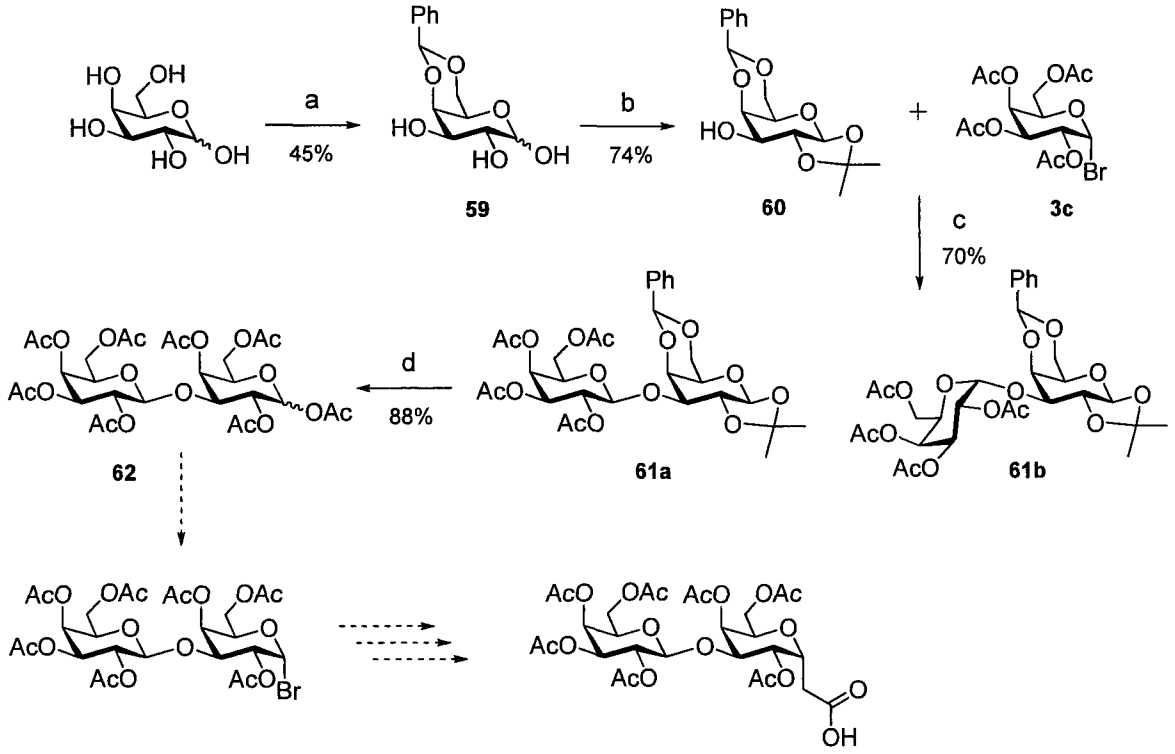


Conditions: (a) $ClAc_2O$, py, cat. DMAP, CH_2Cl_2; (b) 80% aque. AcOH, 80°C; (c) Ac_2O, py, cat. DMAP, CH_2Cl_2; (d) Thiourea, 2,6-Lutidine, $MeOH/CH_2Cl_2$ (6:4); (e) Thiourea, 2,6-Lutidine, $EtOH/CH_2Cl_2$ (8:2); (f) DABCO, EtOH/py (5:1).

Scheme 5.10. Attempted synthesis of C-linked galactosyl acceptor.

In another attempt to synthesize the desired C-linked Gal-Gal-ornithine building block (Scheme 5.11), the Gal-Gal disaccharide was to be synthesized first, then functionalized with the C-linkage using chemistry developed in chapter 4, and subsequently coupled to orthogonally protected ornithine. The synthesis started with protecting C-4 and C-6 hydroxyls of D-galactose with benzylidine acetal to furnish **59** in 45% yield.[45, 50] Further, dimethyl acetal was formed between C-1 and C-2 hydroxyls using dimethoxy propane and catalytic acid, and only the major

β anomer of **60** was isolated in 74% yield and characterized.[51] The efficiently obtained galactosyl acceptor **60** was then coupled with galactosyl bromide using silver triflate and 2,6-lutidine. Under these reaction conditions 1:1 α/β mixture of the desired product was obtained in total 70% yield. Due to the time constrains, only a small amount of **61a** was converted to peracetylated Gal-Gal disaccharide in 88% yield. The ease of preparation of the galactosyl acceptor makes this synthetic route the most promising one. Nonetheless, the glycosylation method has to be altered to obtain better coupling selectivity.

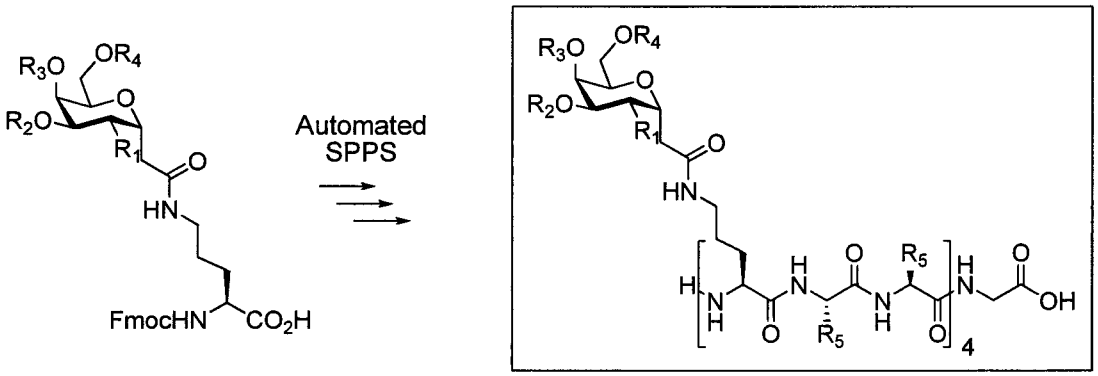


Conditions: (a) p-TSA, p-benzaldehyde dimethyl acetal, DMF; (b) $Me_2C(OMe)_2$, cat. p-TSA, (anhydrous) Acetone; (c) AgOTf, 2,6-Lutidine, CH_2Cl_2; (d) 80% aque. AcOH, 85°C, 24h; (e) Ac_2O, py, cat. DMAP, CH_2Cl_2.

Scheme 5.11. Towards the synthesis of C-linked Gal-Gal disaccharide.

5.4.4. Synthesis of the Glycopeptides

Building blocks **16**, **43**, **49**, and **53** were assembled to the desired glycopeptides in an automated matter using APEX 396 and standard Fmoc-based chemistry (Scheme 5.12.). The glycopolymers were cleaved from the Wang resing by treatment with 50% TFA solution in dichloromethane following ether precipitation. Ester protecting groups on the carbohydrates were removed using wet NaOMe/MeOH solution.[23] Optimal basicity used to remove acetates and pivaloyls without affecting glycosidic linkage was pH=10. Purification by C-18 reversed-phase HPLC furnished the desired *C*-linked AFGP analogues **63-66** in yields ranging from 68-75%.



16: Gal
43: GalNAc
49: GalGalNAc
53: GlcGalNAc

63: [O(Gal)AA]$_4$G: R_1=OH, R_2=H, R_3=H, R_4=H, R_5=CH_3: 75%
64: [O(GalNAc)GG]$_4$G: R_1=NHAc, R_2=H, R_3=H, R_4=H, R_5=H: 68%
65: [O(GalGalNAc)GG]$_4$G: R_1=NHAc, R_2= β-D-galactopyranosyl-(1-3), R_3=H, R_4=H, R_5=H: 72%
66: [O(GlcGalNAc)GG]$_4$G: R_1=NHAc, R_2=β-D-glucopyranosyl-(1-3), R_3=H, R_4=H:R_5=H: 70%

Scheme 5.12. Synthesis of C-linked AFGP Analogues **63-66**.

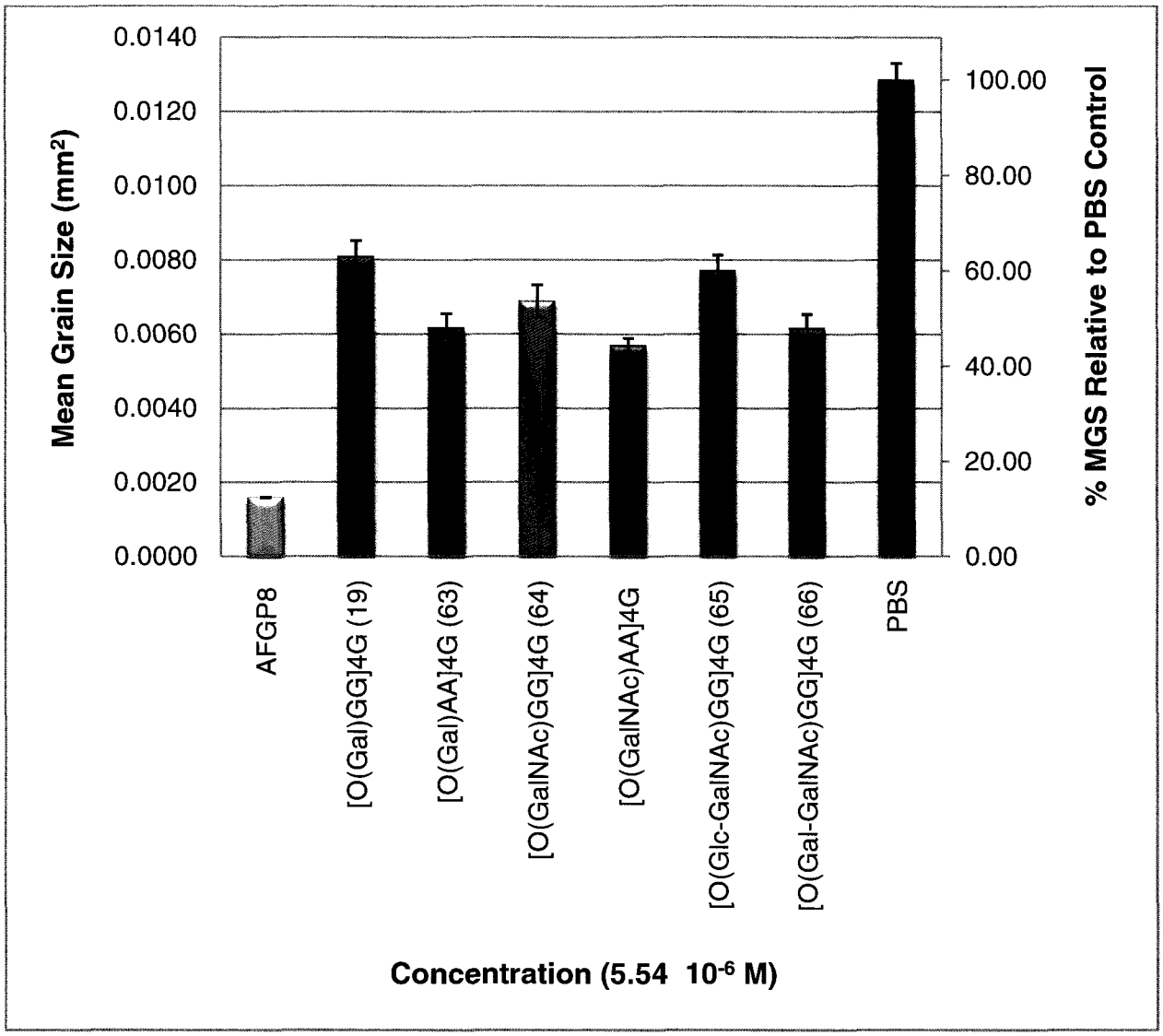
5.5. Results: Assessment of Antifreeze Activity and Conformational Analysis of C-linked AFGP analogues 63-66

Thermal Hysteresis (TH):

The C-linked analogues **63-66** did not show any TH or DIS. This was surprising since Gal-GalNAc of the native AFGP was a key structural motif necessary for TH activity. Nevertheless, these results are consistent with those of the previously synthesized C-linked AFGP analogues.

Recrystallization-Inhibition (RI):

The C-linked AFGP analogues **63-66** were assessed for recrystallization inhibition activity (Graph 5.1.). All samples were compared to a solution of phosphate buffered saline (PBS) and a solution of AFGP 8 isolated from *Gagus ogac* which was used as a positive control. Analogues **63-66** were also compared to analogue **19** $[O(Gal)GG]_4G$ which was used as a standard (a second positive control) for all the previously synthesized AFGP mimetics. Ice crystals size was analyzed using domain recognition software (DRS) program.[52] The Y-axis in Graph 5.1 represents the mean grain size (MGS) that was normalized to MLGS where small bars are representative of potent recrystallization-inhibition activity. The X-axis represents concentration of 5.54×10^{-6} M at which significant RI activity was previously observed.



Graph 5.1. RI activity of the C-linked AFGP analogues **63-66**.

Analogue **63** varies from **19** only in its polypeptide composition (where glycines were exchanged for alanines) and shows a notable increase in RI activity. Similarly, analogue **64** also shows an increase in activity when compared to **19** and the only difference between them is their carbohydrate component. The RI activity of analogue $[O(GalNAc)AA]_4G$ is better than that of **64**, thus these results indicate that an increase in RI can be achieved by incorporating L-alanine

residues into the glycopepetide backbone. Since RI activity of $[O(GalNAc)AA]_4G$ also exceeds that of analogue **63**, it is evident that C-2 N-acetyl group of the carbohydrate also increases RI activity of the C-linked AFGP analogues.

The disaccharide analogue **66**, which contains an additional glucose residue when compared to **64**, showed no improvement in RI activity. However, disaccharide analogue **65** bearing terminal galactose instead of glucose exhibited improved RI activity and exceeded that of monosaccharide analogues **19** and **64** but it did not match that of the native AFGP8. Nevertheless, these results indicate that a proper stereochemistry of the terminal carbohydrate is crucial in obtaining a potent recrystallization inhibitor.

Conformational Analysis:

The solution conformation of C-linked AFGP analogues **63-66** were examined by circular dichroism (CD) spectroscopy (Figure 5.5). Deconvolution of the raw spectral data was accomplished using CD Pro with Selcon III and I-Basis 4 as well as I-Basis 5 as representative databases (Table 5.2 and 5.3.). The molar ellipticity of analogue **63** (Figure 5.5.) is the only one that closely resembles that of the native AFGP8 (Figure 3.3.). Indeed, deconvoluted data using IBASIS 4 indicates that secondary structure of analogue **63** is comparable to that of the AFGP8, which possesses a large percentage of random coil. Surprisingly, solution conformation of analogue **64** is consistent with large percentage of α helix, and analogues **66** and **65** exhibit a large percentage of β strand. IBASIS 5 gave similar results, except the percentage of random coil significantly increased for all the analogues, and analogue **64** showed a larger percentage of β strand instead of α helix.

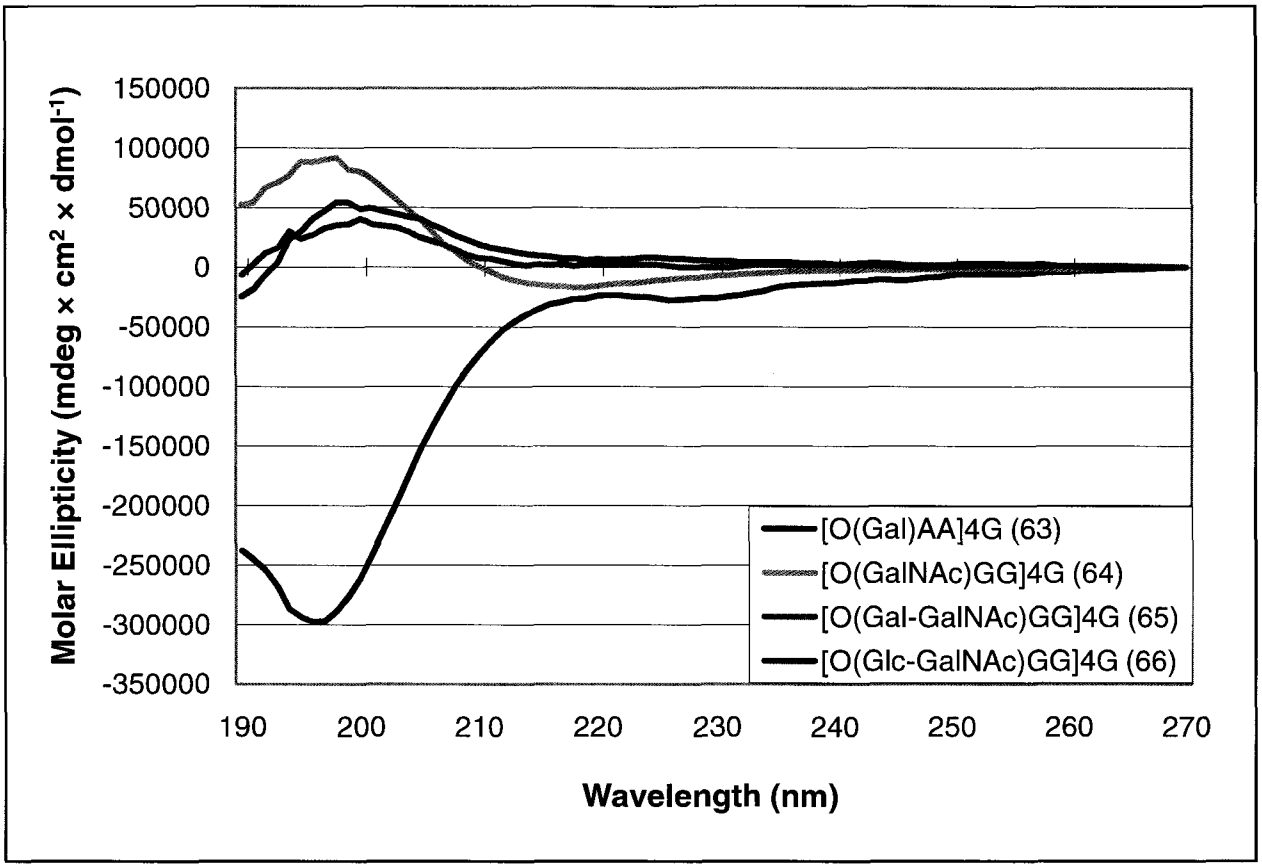


Figure 5.5. CD specta of the C-linked AFGP analogues **63-66**.

Table 5.2. CD deconvolution data of AFGP analogues **63-66** and AFGP8 using CD Pro with Selcon III and IBASIS 4.

Compound	Regular α helix	Distorted α helix	Regular β strand	Distorted β strand	Turn	Random coil	Sum
$[O(Gal)AA]_4G$ (**63**)	-0.025	0.057	0.195	0.112	0.237	0.470	1.047
$[O(GalNAc)GG]_4G$ (**64**)	0.524	0.148	0.173	0.014	-0.001	0.175	1.033
$[O(Gal\text{-}GalNAc)GG]_4G$ (**65**)	-0.010	-0.008	0.433	0.164	0.145	0.260	0.984
$[O(Glc\text{-}GalNAc)GG]_4G$ (**66**)	-0.024	0.013	0.511	0.164	0.093	0.255	1.013
AFGP8	0.049	0.114	0.164	0.094	0.237	0.378	1.036

Table 5.3. CD deconvolution data of AFGP analogues **63-66** and AFGP8 using CD Pro with Selcon III and IBASIS 5.

Compound	α helix	β strand	β turn	PPII	Random coil	Sum
$[O(Gal)AA]_4G$ (**63**)	0.034	0.070	0.162	0.212	0.557	1.036
$[O(GalNAc)GG]_4G$ (**64**)	0.104	0.402	0.087	0.055	0.348	0.996
$[O(Gal\text{-}GalNAc)GG]_4G$ (**65**)	-0.022	0.379	0.077	0.109	0.452	0.996
$[O(Glc\text{-}GalNAc)GG]_4G$ (**66**)	-0.016	0.367	0.046	0.107	0.508	1.012
AFGP8	0.129	0.159	0.156	0.136	0.445	1.025

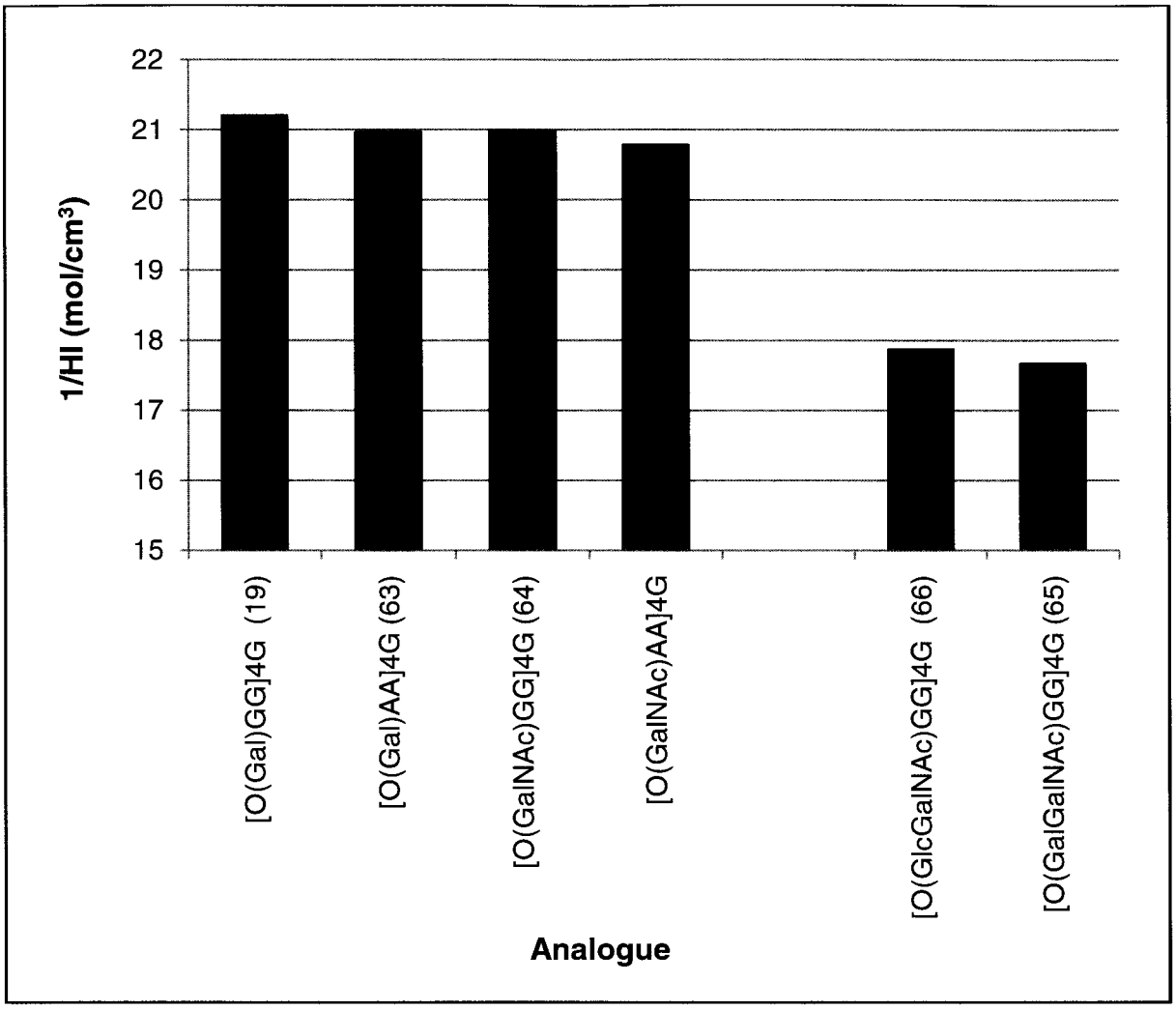
5.6. Discussion: Interpretation of Antifreeze Activity and Conformational Analysis of Analogues 19, 63, 64, 65, 66 and $[O(GalNAc)AA]_4G$

By changing the peptide backbone from glycines (analogue **19**) to alanines (analogue **63**), RI activity improved ~25%, suggesting that L-alanines are important structural features for the activity (graph 5.1). This also supports the idea that the hydrophobic methyl groups of the L-alanines are necessary for the antifreeze activity, yet to date there had been no SAR studies performed to verify this for RI activity. These SAR results show that this simple modification to the polypeptide backbone of the C-linked AFGP analogues drastically improves RI activity and should be implemented into future analogues.

Another important feature studied was the N-acetyl moiety at the C-2 position of the carbohydrate. This was done by comparing the RI activity of analogue **64** $[O(GalNAc)GG]_4G$ to analogue **19** $[O(Gal)GG]_4G$. Our data shows that the presence of the N-acetyl moiety in this series improves the RI activity by ~10%. This trend was also observed for the O-linked

[T(galNAc/gal)AA]$_4$G].[5a] Yet, in the previously synthesized (Table 2.1) C-linked [S(galNAc/gal)GG]$_4$G] series the opposite was true, whereby the absence of the C-2 N-acetyl moiety improved RI activity.[5b, h] This demonstrates that the presence of the C-2 N-acetyl moiety can have a positive or negative effect on RI activity which depends on the individual series.

Further, by incorporating both the N-acetyl moiety and the alanine residues into the C-linked AFGP mimetics, the [O(GalNAc)AA]$_4$G analogue[5a] showed MGS of ice crystals ~45% relative to that of a PBS control, making it a better recrystallization inhibitor then the previously discussed analogues **19**, **63** and **64**. A rational explanation for the aforementioned results is anticipated to be via the hydration model developed in chapter 4, where the hydration index (HI) of the AFGP analogues was proven to be directly proportional to RI activity. Unfortunately, this model predicts that RI activity of **63** [O(Gal)AA]$_4$G will be equal to that of [O(Man)GG]$_4$G (Graph 5.2 and 4.3), but this was not observed experimentally. Hence, predicted data for hydration does not correctly fit that of the RI activity. This was the first discrepancy encountered which suggests that our method for estimating hydration indices of glycopeptides needs to be further refined. Nonetheless, this is the first and only method which correlates hydration of AFGP analogues to their RI activity.



Graph 5.2. The hydration index inverse of analogues **63-66**.

In order to estimate hydration indices for **64** and $[O(GalNAc)AA]_4G$, we had to search for hydration numbers of the N-acetylated galactosamine throughout the literature. The previously reported data by Galema[53] did not include hydration numbers for any N-acetylated carbohydrates. The only data reported on hydration of these derivatives was by Uedaira et al, who determined hydration numbers of monosaccharides from accessible surface area (ASA) of the solute molecules.[54] Uedaira et al found that the hydration number of D-galactose was

approximately three times the value reported by Galema.. Therefore, in order to determine hydration index of analogue **64** [O(GalNAc)GG]$_4$G, Uedairas hydration number for GalNAc had to be converted to that of Galemas. Thus, hydration number of galactose determined by Galema was multiplied by the ratio of hydration numbers of N-acetyl galactosamine/D-galactose reported by Uedaira. The calculated hydration value of N-acetyl galactosamine (10.5) was then applied to the method for computing hydration index formulated in chapter 4, and hydration indices for **64** and [O(GalNAc)AA]$_4$G were calculated (Graph 5.2). Notably, the precision of Uedaira's method is limited in comparison to that of Galema's since hydration numbers for galactose and glucose are the same. It is also apparent that this method only provides a rough estimation of the hydration value of N-acetyl galactosamine.

Nonetheless, the inverse of the hydration indicies were calculated and it was found that the value of [O(GalNAc)GG]$_4$G was lower than that of [O(Gal)GG]$_4$G, which matches the pattern of experimental MGS values (Graph 5.1). The same trend was observed for analogues **63** (gal oaa) and [O(GalNAc)AA]$_4$G, whereby the C-2 N-acetyl group improved the RI activity. However the calculated 1/HI values of the [O(Gal/GalNAc)GG]$_4$G series were not lower than those of the [O(Gal/GalNAc)AA]$_4$G series as was observed experimentally. This suggests that the model of predicting hydration is still preliminary and other factors affecting hydration were not taken into account.

Consequently, we speculate that other effects such as intramolecular hydrogen bonding within a peptide, water pockets and hydrophobic interactions[12, 55] are likely to influence hydration of the AFGP analogues and subsequently alter their antifreeze activity. The N-acetyl group is known to form hydrogen bonds with amide residues of the polypeptide backbone, and depending on its location and strength the degree of hydration is tempered. Moreover, the C-2

acetamide group may also hydrogen bond with the residues of a polypeptide through water molecule bridges, which occurs in a similar fashion as with molecules that contain both cationic and anionic groups and form intramolecular ionic bonds or salt bridges.[55, 56] Thus, water pockets, including one or more H_2O molecules, have been commonly found in proteins as well as glycoproteins.[56] Consequently, it is plausible that this effect of trapping water molecules will increase hydration of the glycopeptide and subsequently perturb its antifreeze activity. Furthermore, it is possible that hydrophobic methyl group of the N-acetyl moiety functions by interacting with the hydrophobic layers between the ice crystals, as was proposed for alanine residues. These abovementioned factors are likely to affect both glycopeptides hydration and its overall conformation.

In light of further enhancement of RI activity N-acetyl carbohydrate series was expanded towards disaccharides. Two analogues were designed: $[O(Gal\text{-}GalNAc)GG]_4G$ (**65**) and $[O(Glc\text{-}GalNAc)GG]_4G$ (**66**). The choice of the carbohydrate was based on the native AFGPs which possess Gal-GalNAc disaccharide, and work done by Fletcher and Nishimura who determined that any chemical modifications to the Gal-GalNAc disaccharide would either diminish or eliminate AFGP function.[4] In addition, Ben et al. showed that by doubling the carbohydrate concentration with respect to polypeptide backbone (Figure 2.6.), the RI activity would be more than doubled (Table 2.1).[5c] Furthermore, as described in chapter 4, the analogues which contained dissacharides, $[O(Lac)GG]_4G$ and $[O(Meli)GG]_4G$, exhibited improved RI activity in comparison to analogue **19**, with $[O(Meli)GG]_4G$ showing the best RI of the $[O(carbohydrate)GG]_4G$ series.

As expected, the RI activity of **65** $[O(Gal\text{-}GalNAc)GG]_4G$ was improved over that of analogue **19** $[O(Gal)GG]_4G$ and **64** $[O(GalNAc)GG]_4G$ (Graph 5.1). But the activity of **65** only

increased ~5%, when compared to analogue **64**, even though carbohydrate concentration was doubled with respect to the glycopeptides backbone. Moreover, analogue **65** did not exceed the superior antifreeze activity of the native AFGP8, indicating that Gal-GalNAc does not play a vital role in RI activity of our analogues. However, in order to fully validate these results more Gal-GalNAc containing analogues would have to be synthesized and compared to other analogues of its series. For example, an obvious target would be the C-linked [S(Gal-GalNAc)GG]$_4$G.

To explore the significance of the terminal carbohydrate of the disaccharide, analogue **66** was tested, and was found to possess RI activity inferior to that of analogue **65**. This was not surprising as it was demonstrated previously that glucose is less hydrated then galactose, and RI activity is directly proportional to the hydration of the carbohydrate.[6, 53, 57] Furthermore, [O(Glc-GalNAc)GG]$_4$G exhibits a lower RI activity then [O(GalNAc)GG]$_4$G, even though it contains a double concentration of carbohydrates. This result further supports our suggestion that a disaccharide does not play an essential role in inducing potent RI activity of our C-linked AFGP mimetics. However, this data signifies that a proper choice of terminal carbohydrate is crucial in obtaining an enhanced recrystallization inhibition. As discussed in chapter 3, this choice should consider ideal stereochemistry, and therefore the most hydrated carbohydrate, which is D-galactose.[6, 53, 57] This is demonstrated in the experimental data whereby analogues **19**, **64** and **65** are better recrystallization inhibitors then analogue **66**.

As per previous results the hydration indicies of **65** and **66** should demonstrate that the latter should contain better RI activity. However, it is difficult to calculate HI for these analogues because no hydration numbers of either Gal-GalNAc or Glc-GalNAc have been reported. A vague estimate of these values could be based on sum of the hydration numbers of the individual

monosaccharides. Thus, estimated hydration number for Gal-GalNAc is 19.2 (8.7_{Gal}+10.5_{GalNAc}) and for Glc-GalNAc it is 18.9 (8.4_{Glc}+10.5_{GalNAc}). These values are overestimated because hydration spheres of the monosaccharides comprising the disaccharides are overlapping. Nevertheless, hydration indices were calculated using McGowans method[58] for estimating molar volumes and Peirottis equation[59] for determining hydration numbers of polypeptides. The 1/HI value of analogue **65** [O(Gal-GalNAc)GG]$_4$G was calculated to be lower than that of analogue **66** [O(Glc-GalNAc)GG]$_4$G (Graph 5.2). This data clearly matches the RI pattern of the two analogues (Graph 5.1.).

To determine whether specific change of conformation affected RI activity, analogues **63**, **64**, **65** and **66** were analyzed by circular dichroism spectroscopy (Figure 5.5, Table 5.2 and 5.3). It has been demonstrated that solution conformation of analogues **19** and **63**, which consists mainly of random coil with various degrees of other secondary structures, is alike that of AFGP8. Hence, the increased RI activity of **63** over **19** is due to the inherent property of alanine residues and not its conformation. Conversely, analogue **64** [O(GalNAc)GG]$_4$G exhibits solution conformation significantly different than that of AFGP8, **19** and **63**, where a predominant percent of its conformation consists of α-helix and/or β-sheet. Moreover, analogues **65** and **66** display a large percentage of random coil but they also possess notably greater amount of β-strand similar to analogue **64**. The formation of these more rigid structures, α-helix and β-sheet, could be caused by interaction of the N-acetyl moiety with the polypeptide backbone. However, this is difficult to conclude as the conformation of previously prepared [O(GalNAc)AA]$_4$G resembles more that of AFGP8 and not **64**. Nevertheless, we postulate that the observed conformational change of [O(N-acetyl-carbohydrate)GG]$_4$G series is due to the C-2 N-acetyl moiety of the saccharide which also affects hydration and RI activity. Consequently, synthesis

and conformational analysis of other C-linked AFGP analogues, such as $[O(GalGal)GG]_4G$, would definitely contribute to validation of this hypothesis. Computational modeling studies of the analogues prepared in this chapter would also provide more insight to the function of the carbohydrates C-2 acetamide group.

5.7. Conclusion

Collectively, it has been presented that C-2 N-acetyl moiety of the carbohydrate, and the L-alanine residues of the polypeptide backbone and carbohydrates stereochemistry influence RI activity of the C-linked AFGP analogues (Graph 5.1). The RI activity of $[O(GalNAc)AA]_4G$ analogue was superior to that of analogues **19**, **63** and **64**, which suggests that C-2 N-acetyl moiety and L-alanine residues have a positive effect on recrysallization inhibition. Analogue **65** showed RI activity only equivalent to that of $[O(GalNAc)AA]_4G$, implying that custom synthesis of labour intensive discharide is not necessary for potent recrystallization inhibition. The effect of the terminal carbohydrate of the disaccharide was explored by synthesizing analogue **66**, which showed RI activity lower than analogue **64** and **65**. This proved that stereochemistry of the terminal carbohydrate is very important for RI activity and is optimal when it matches that of galactose.

In order to explain RI activity of our analogues, in chapter 4 we were able to develop a method for estimating hydration numbers, partial molar volumes of glycopeptides, and subsequently hydration indices. While hydration index was determined to be proportional to RI activity, predictions in Graph 5.2 do not accurately match the results in Graph 5.1, because of error associated with calculating hydration indices. In chapter 4, it has been demonstrated that small changes in the magnitude of hydration numbers have a significant effect on the value of

hydration index. Thus, if we change the total calculated hydration number of the peptide component of [O(carbohydrate)AA]$_4$G series only by +5%, then 1/HI trend for analogues **19**, **63**, **64** and [O(GalNAc)AA]$_4$G exactly matches their trend for MGS values. Also, if the overestimated hydration numbers for Gal-GalNAc and Glc-GalNAc were corrected by accounting for overlapping hydration spheres of their individual monosaccharides, then the newly determined values for 1/HI would closer resemble the RI trend in Graph 5.1. Nevertheless, it is clear that we need better thermodynamic hydration values in order to determine the extent to which hydration allows us to predict RI activity. Therefore, it is necessary to experimentally determine hydration numbers and partial molar volumes of our C-linked AFGP analogues and their components.

5.8. References

1. Komatsu, S. K., DeVries, A. L., Feeney, R. E., *J. Biol. Chem.* **1970**, 245, 2909.

2. (a) Ahmed, A. I., Yeh, Y., Osuga, D. T., Feeney, R. E., *J. Biol. Chem.* **1976**, 251, 3033. (b) DeVries, A. L., *Science* **1971**, 171, 1152. (c) Osuga, D. T., Feather, M. S., Shah, M. J., Feeney, R. E., *J. Prot. Chem.* **1989**, 8, 519.

3. Haymet, A. D. J., Ward, L. G., Harding, M. M., *J. Am. Chem. Soc.* **1999**, 121, 941.

4. Tachibana, Y., Fletcher, G. L., Fujitani, N., Tsuda, S., Monde, K., Nishmura, S., *Angew. Chem. Int. Ed.*, **2004**, 43, 856.

5. (a) Bouvet, V., *Ph.D. Thesis*, University of Ottawa, Ottawa, **2006**. (b) Liu, S., *Ph.D. Thesis*, University of Ottawa, Ottawa, **2006**. (c) Murphy, A., *Ph.D. Thesis*, State University of Binghamtom, Binghamton, **2004**. (d) Ben, R. N., Eniade, A., Hauer, L., *Org. Lett.* **1999**, 1, 1759. (e) Eniade, A., Ben, R. N., *Biomacromolecules* **2001**, 2, 557. (f) Eniade, A., Murphy A. V., Landreau, G., Ben, R. N., *Bioconjgate Chem.* **2001**, 12, 817. (g) Eniade, A., Purushotham, M., Ben, R. N., Wang, J. B., Horwath, K., *Cell Biochem. Biophys.* **2003**, 38, 115. (h) Liu, S., Ben, R. N., *Org. Lett.* **2005**, *7,* 2385.

6. Czechura, P., Tam, R. Y., Dimitrijevic, E. Murphy A. V., Ben, R. N., *J. Am. Chem. Soc.* **2008**, 130, 2918.

7. Yeh, Y., Feeney, R. E., *Chem. Rev.* **1996**, 96, 601.

8. (a) Franks, F., Morris, E. R., *Biochim. Biophys. Acta* **1978**, 540, 246. (b) Lane, A. N., Hays, L. M., Feeney, R. E., Crowe, L. M., Crowe, J. H., *Protein Sci.* **1998**, 7, 1555. (c) Lane, A. N., Hays, L. M., Tsvetkova, N., Feeney, R. E., Crowe, L. M., Crowe, J. H., *Biophys. J.* **2000**, 78, 3195.

9. (a) Raymond, J. A., Radding, W., DeVries, A. L., *Biopolymers* **1977**, 16, 2575. (b) Filira, F., Scolaro, B. B., Fofani, M. T., Mammi, S., Peggion, R., Rocchi, R., *Int. J. Biol. Macromol.* **1990**, 12, 41. (c) Tomimatsu, Y., Scherer, J., Yeh. Y., Feeney, R. E., *J. Biol. Chem.* **1976**, 251, 2290. (d) Berman, E., Allerhand, A., DeVries, A. L., *J. Biol. Chem.* **1980**, 255, 4407.

10. (a) Bouvet, V., Ben, R. N., *Cell Biochem. Biophys.* **2003**, 39, 133. (b) Tsvetkova, N. M., Phillips, B. L., Krishnan, V. V., Feeney, R. E., Fink, W. H., Crowe, J. H., Risbud, S. H., Tablin, F., and Yeh, Y. *Biophys. J.* **2002**, 82, 464–473.

11. (a) Steim, J. M., Fleischer, S., *Proc. Natl. Acad. Sci.* **1967**, 58, 1292. (b) Davies, P. L., Hew, C. L., *FASEB J.* **1990**, 4, 2460.

12. (a) Mimura, Y., Yamamoto, Y., Inoue, Y., Chujo, R., Int. J. Biol. Macromol. 1992, 14, 242-248. (b) Tam, R. Y., Ben, R. N., unpublished results.

13. Roe, B. A., Boojamra, C. G., Griggs, J. L., Bertozzi, C. R., *J. Org. Chem.* **1996**, 61, 6442. (b) Cipolla, L., La Ferla, B., Lay, L., Peri, F., Nicotra, F., *Tetrahedron: Asymmetry* **2000**, 11, 295. (c) Bouvet, V. R., Ben, R. N., *J. Org. Chem.* **2006**, 71, 3619.

14. Komatsu, S. K., DeVries, A. L., Feeney, R. E., *J. Biol. Chem.* **1970**, 245, 2909.

15. (a) Koenigs, W., Knorr, E., *Ber. Dtsch. Chem. Ges.* **1901**, 34, 957. (b) Garegg, P. J., Norberg, T., *Acta. Chem. Scand. B* **1979**, 33, 116. (c) Banoub, J., Bundle, D. R., *Can. J. Chem.* **1979**, 57, 2091.

16. (a) Michel, J., Schmidt, R. R., *Angew. Chem. Int. Ed.* **1980**, 19, 731. (b) Zimmermann, P., Schmidt, R. R., *Angew. Chem. Int. Ed.* **1986**, 25, 725. (c) Schmidt, R. R., *Pure. Appl. Chem.* **1989**, 61, 1257. (d) Schmidt, R. R., Toepfer, A., *Tet. Lett.* **1991**, 32, 3353. (e) Zhu, X. M., Schmidt, R. R., *Angew. Chem. Int. Ed.* **2009**, 48, 1900.

17. (a) Giannis, A., Sandhoff, K., *Carbohydr. Res.* **1987**, 171, 201. (b) Carcano, M., Nicotra, F., Panza, L., Russo, G., *J. Chem. Soc., Chem. Commun.* **1989**, 297. (c) Hoffman, M., Kessler, H., *Tet. Lett.* **1994**, 32, 6067.

18. Lemieux, R. U., Ratcliffe, R M., *Can. J. Chem.* **1979**, 57, 1244.

19. (a) Bovin, N. V., Zurabyan, S. E., Khorlin, A. Y., *Carbohydr. Res.* **1981**, 98, 25. (b) Czernecki, S., Ayadi, E., Randriamandimby, D., *J. Org. Chem.* **1994**, 59, 8256. (c) Giuliano, R. M., Davis, R. S., Boyko, W. J., *Carb. Chem.* **1994**, 13, 1135.

20. (a) Dondoni, A., Mariotti, G., Marra, A., *J. Org. Chem.* **2002**, 67, 4475. (b) Bertozzi, C. R., Bednarski, M. D., *Tet. Lett.* **1992**, 33, 3109. (c) Hoffmann, M. G., Schmidt, R. R., *Liebigs. Ann. Chem.* **1985**, 2403.

21. Keck, G. E., Yates, J. B. *J. Am. Chem. Soc.* **1982**, 104, 5829.

22. Pravdic, N., Franjic-Mihalic, I., Danilov, B., *Carbohydr. Res.* **1975**, 45, 302.

23. Hadacsy, I., Gerecs, A., Zemplen, G., *Ber. Dtsch. Chem. Ges.* **1936**, 69B, 1827.

24. (a) Fischer, E., *Ber. Dtsch. Chem. Ges.* **1893**, 26, 2400. (b) Beensch, L., Fischer, E., *Ber. Dtsch. Chem. Ges.* **1894**, 27, 2478. (c) Fischer, E., *Ber. Dtsch. Chem. Ges.* **1895**, 28, 1145.

25. (a) Zhang, Z., Ollmann, I. R., Ye, X.-S., Wischnat, R., Baasov, T., Wong, C.-H., J. Am. Chem. Soc. **1999**, 121, 734-753. (b) Xin-Shan, Y., Wong, C.-H., *J. Org. Chem.* **2000**, 65, 2410. (c) Sears, P., Wong, C.-H., *Science* **2001**, 291, 2344.

26. (a) Paulsen, H., *Angew. Chem. Int. Ed.* **1982**, 21, 155. (b) Paulsen, H., *Angew. Chem. Int. Ed.* **1982**, 34, 957.

27. Stick, R. V., *Carbohydrates: The sweet moleculaes of life*, Academic Press **2001**.

28. (a) Schmidt, R. R., Kinzy W., *Adv. Carb. Chem. Biochem.* **1994**, 50, 21. (b) Davis, B. G., *J. Chem. Soc. Perkin Trans. 1* **2000**, 2137. (c) Jensen, K. J., *J. Chem. Soc. Perkin Trans. 1* **2002**, 2219. (d) Demchenko, A. V., *Synlett* **2003**, 1225. (e) Toshima, K., *Carbohydr. Res.* **2006**, 341, 1282. (f) Galonic, D. P., Gin, D. Y., *Nature* **2007**, 446, 1000.

29. (a) Mukaiyama, T., Murai, Y., Shoda, S., *Chem. Lett.* **1981**, 431. (b) Mukaiyama, T., Hashimoto, Y., Shoda, S., *Chem. Lett.* **1983**, 935. (c) Nicolaou, K. C., Chucholowski, A., Dolle, R. E., Randall, J. L., *J. Chem. Soc., Chem. Commun.* **1984**, 1155. (d) Nicolaou, K. C., Caulfield, T. J., Groeneberg, R. D., *Pure Appl. Chem.* **1991**, 63, 555. (e) Toshima, K., *Carbohydr. Res.* **2000**, 327, 15.

30. Pougny, J. R., Jacquinet, J. C., Nassr, M., Duchet, D., Milat, M. L., Sinay, P. *J. Am. Chem. Soc.* **1977**, 99, 6762.

31. (a) Foged, P., Garegg, P. J., Lonn, H., Norberg, T., *Glycoconj. J.* **1987**, 4, 97. (b) Veeneman, G. H., van Leeuwen, S. H., van Boom, J. H., *Tet. Lett.* **1990**, 31, 1331. (c) P. Konradsson, P.,Udodong, U. E., Fraser-Reid, B., *Tet. Lett.* **1990**, 31, 4313.

32. (a) Khane, D., Walker, S., Cheng, Y., Van Engen, D., *J. Am. Chem. Soc.* **1989**, 111, 6881. (b) Raghavan, S., Kahne, D., *J. Am. Chem. Soc.* **1993**, 115, 1580. (c) Liang, R., Yan. L., Loebach, J., Ge, Min., Uozumi, Y., Sekanina, K., Horan, N., Gildersleeve, J., Thompson, C., Smith, A., Biswas, K., Still, C., Khane, D., *Science* **1996**, 274, 1520.

33. (a) Mehta, S., Pinto, B. M., *Tet. Lett.* **1991**, 32, 4435. (b) Mehta, S., Pinto, B. M., *J. Org. Chem.* **1993**, 58, 3269.

34. (a) Fraser-Reid, B., Udodong, U. E., Wu, Z., Ottosson, H., Merritt, J. R., Rao, C. S., Roberts, C., Madsen, R., *Synlett.* **1992**, 927. (b) Fraser-Reid, B., Mootoo, D. R.,

Konradsson, P., Udodong, U. E., Andrews, W. C., Ratcliffe, A. J., Wu, Z., Yu, K. L., *Pure Appl. Chem.* **1989**, 61, 1243

35. (a) Randolph, J. T., Danishefsky, S. J., *J. Am. Chem. Soc.* **1993**, 115, 8473. (b) Danishefsky, S. J., Bilodeau, M., *Angew. Chem. Int. Ed.* **1996**, 1381.

36. (a) Schmidt, R. R., Stumpp, M., Michel, J., *Tet. Lett.* **1982**, 23, 405, (b) Schmidt, R. R., Stumpp, M., *Liebigs Ann. Chem.* **1984**, 680. (c) Kondo, H., Aoki, S., Ichikawa, Y., Halcomb, R. L., Ritzen, H., Wong, C.-H., *J. Org. Chem.* **1994**, 59, 864. (d) Wong, C.-H., Halcomb, R. L., Ichikawa, Y., Kajimoto, T., *Angew. Chem. Int. Ed.* **1995**, 34, 412.

37. (a) Sugarwara, T., Susaki, H., Nogusa, H., Gonsho, A., Iwasawa, H., Irie, K., Ito, K., Shibukawa, M., *Carbohydr. Res.* **1993**, 238, 163. (b) Li, T.-S., Li, J.-T., Li, H.-Z., *J. Chomatgr. A* **1995**, 715, 372. (c) Holm, B., Broddefalk, J., Flodell, S., Wellner, E., Kihlberg, J., *Tetrahedron* **2000**, 56, 1579.

38. Schmidt, R. R., Michel, J., *Tet. Lett.* **1984**, 25, 821.

39. Schmidt, R. R., Behrendt, M., Toepfer, A., *Synlett* **1990**, 694.

40. de la Funente, J. M., Penades, S., *Tet. Asym.* **2002**, 13, 1879.

41. Liao, W., Locke R. D., Matta, K. L., *Chem. Commun.* **2000**, 369.

42. Kinzy, W., Schmidt, R. R., *Carbohydr. Res.* **1987**, 164, 265.

43. (a) Carlsen, H. J., Katsuki, T., Martin, V. S., Sharpless, B. K., *J. Org. Chem.* **1981**, 46, 3936. (b) Plietker, B., *Synthesis* **2005**, 15, 2453.

44. Bundle, D. R., Eichler, E., Lowary, T. L., *Can. J. Chem.* **2002**, 80, 1112.

45. Lonn, H., *Carbohydr. Res.* **1985**, 139, 105.

46. Johansson, R., Samuelsson, B., *Chem. Commun.* **1984**, 201.

47. Belot, F., Jacquinet, J.-C., *Carbohydr. Res.* **2000**, 325, 93.

48. (a) Yang, G., Kong, F., *Carbohydr. Res,* **2005**, 340, 39. (b) Kovac, P., *Carbohydr. Res,* **1986**, 153, 237. (c) Lefeber, D. J., Kamerling, J. P., Vliegenthart, J. F. G., *Org. Lett.* **2000**, 2, 701.

49. (a) Gros, E.G., Deulofeu, V., *J. Org. Chem.* **1964**, 29, 3647. (b) Roslund, M. U., Aitio, O., Warna, J., Maaheimo, H., Murzin, D. Y., Leino, R., *J. Am. Chem. Soc.* **2008**, 130, 8769.

50. Duclos Jr., *Chem. Phys. Lipids* **2001**, 111, 111.

51. Busse, H., Hakoda, M., Stanley, M., Streicher, H., *J. Carb. Chem.* **2007**, 26, 159.

52. Jackmann, J., Noestheden, M., Moffat, D., Pezacki, J., Findlay, S., Ben, R. N., *Biochem. Biophys. Res. Commun.* **2007**, 354, 340.

53. Galema, S. A., Hoiland, H., *J. Phys. Chem.* **1991**, 95, 5321.

54. Uedaira, H., Okouchi, S., Tsuda, S., Uedaira, H., *Bull. Chem. Soc. Jpn.* **2001**, 74, 1857.

55. Franks, F., *Protein Biotechnology*, The Humana Press **1993**.

56. Wyttenbach, T., Liu, D., Bowers, M. T., *Int. J. Mass Spec.* **2005**, 240, 221.

57. Tam, R. Y, Ferreira, S. S., Czechura, P., Chaytor, J. L., Ben, R. N., *J. Am. Chem. Soc.* **2008**, 130, 17494.

58. (a) Mellors, A., McGowan, J. C., *Biochem. Pharmacol.* **1985**, 34, 2413. (b) Abraham, M. H., McGowan, J. C., *Chromatographia*, **1987**, 23, 243.

59. (a) Deiber, J. A., Piaggio, M. V., Peirotti, M. B., *J. Sep. Sci.* **2008**, 31, 548. (b) Deiber, J. A., Piaggio, M. V., Peirotti, M. B., *Electrophoresis* **2007**, 3658.

CHAPTER 6

Experimental Procedures

6.1. General Procedures

Unless otherwise stated all anhydrous reactions were performed in flame-dried or oven-dried glassware under a positive pressure of dry argon or nitrogen. Air or moisture-sensitive reagents and anhydrous solvents were transferred with oven-dried syringes or cannulae. All flash chromatography was performed with E. Merck silica gel 60 (230-400 mesh). All solution phase reactions were monitored using analytical thin layer chromatography (TLC) with 0.2 mm pre-coated silica gel aluminum plates 60 F254 (E. Merck). Components were visualized by illumination with a short-wavelength (254 nm) ultra-violet light and/or staining (ceric ammonium molybdate, potassium permanganate, or phosphomolybdate stain solution).

6.2. Solvents

All solvents used for anhydrous reactions were distilled. Tetrahydrofuran (THF) and diethyl ether were distilled from sodium/benzophenone under nitrogen atmosphere. Dichloromethane, acetonitrile, triethylamine, benzene and diisopropylethylamine (DIPEA) were distilled from calcium hydride. Methanol was distilled from calcium sulfate. *N,N*-dimethylformamide (DMF) and acetone were stored over activated 4Å molecular sieves under argon atmosphere.

All deuterated solvents ($CDCl_3$, CD_3OD and D_2O) were obtained from Cambridge Isotope Laboratories Ltd. and $CDCl_3$ was stored over Linde 4Å molecular sieves.

6.3. Instrumentation

^{1}H (300, 360, 400, or 500 MHz) and ^{13}C NMR (75, 90, 100 or 125 MHz) spectra were recorded at ambient temperature on a Bruker Avance 300, Bruker AM 360, Bruker Avance 400 or Bruker Avance 500 spectrometer. Chemical shifts are reported in ppm downfield from TMS and corrected using the solvent residual peak or TMS as an internal standard. Splitting patterns are designated as follows: s, singlet; d, doublet; t, triplet; q, quartet; m, multiplet and br, broad. Low resolution mass spectrometry (LRMS) was performed on a Micromass Quatro-LC Electrospray spectrometer with a pump rate of 20μL/min using electrospray ionization (ESI) or a Voyager DE-Pro matrix-assisted desorption ionization-time of flight (MALDI-TOF), (Applied Biosystem, Foster City, CA) mass spectrometer operated in the reflectron/positive-ion mode with DHB in 20% EtOH/H_2O as the MALDI matrix. High resolution mass spectrometry (HRMS) data was acquired on Applied Biosystems/Sciex QStar (Concord, ON). Samples in CH_2Cl_2/MeOH 1:1 were mixed with Agilent ES tuning mix for internal calibration, and infused into the mass spectrometer at 5 μL/min. Infrared absorption spectra (IR) were recorded on a Shimadzu FTIR-8400S spectrometer as a neat film from 4000 cm^{-1} to 650 cm^{-1}. Analytical and preparatory scale RP-HPLC were carried out with C-18 columns on a Varian Dynomax HPLC system equipped with a variable wavelength detector (ProStar 330 PDA) or a Waters Delta 600E HPLC system equipped with a variable wavelength detector. Automated solid phase peptide synthesis (SPPS) was performed on APEX 396 (Advanced ChemTech) equipped with a 40 well reaction vessel. Circular Dichroism was performed on JASCO Model J-810 automatic spectrophotometer interfaced with Acer computer. Spectral data were obtained with the following parameters: 1nm bandwidth; four scans at a scan speed of 50nm/min; 190-350nm wavelength range. Samples were measured at 22 °C in a quartz cell (Hellma QS) with a 1 cm path length. A molar concentration of 43 μM was used for each sample, diluted in doubly distilled water. Data

obtained from CD spectroscopy was converted into molar ellipticity (deg cm^2 $dmol^{-1}$). Glycopeptide secondary structures were estimated using the deconvolution software CD Pro (SELCON 3 program and IBASIS 4 protein dataset).

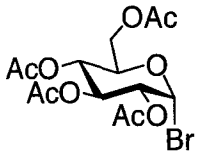
6.4. Synthetic Procedures A

6.4.1. General protocol for bromination of a D-pyranose pentaacetate[1]

To commercially available D-pyranose pentaacetate (5 g, 12.81 mmol), 30 mL of HBr in AcOH (33% solution) was added at room temperature. The reaction was stirred for 40 mins and then diluted with CH_2Cl_2 (50 mL). The solution was transferred to a separatory funnel containing ice water and the organic layer was washed with water until neutral pH. The organic extract was dried over $MgSO_4$, filtered and concentrated *in vacuo* to afford product in ~90% yield. The crude product was crystallized from diethyl ether to form a white powder.

6.4.1.1. 2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl bromide (2a)

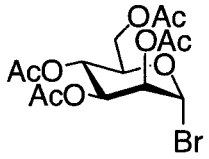
Reaction yield: 94%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 6.16 (1H, d, *J*=1.3 Hz), 5.53 (1H, dd, *J*=3.6, 7.8 Hz), 5.27 (2H, m), 4.15 (1H, dd, *J*=9.6, 18 Hz), 4.07 (1H, m), 3.98 (1H, dd, *J*=2.1, 12 Hz), 2.01, (3H, s), 1.98 (3H, s), 1.91 (3H, s), 1.84 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.28, 169.51, 169.39, 169.35, 83.20, 72.71, 71.90, 67.78, 65.05, 61.27, 20.60, 20.49, 20.41; LRMS (ESI): Calcd for $C_{14}H_{23}BrNO_9$ $[M+NH_4]^+$ 429.2, found 429.9.

6.4.1.2. 2,3,4,6-tetra-O-acetyl-α-D-mannopyranosyl bromide (2b)

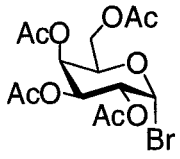
Reaction yield: 92%.



^{1}H NMR (360 MHz, $CDCl_3$) δ 6.45 (1H, d, *J*=3.9 Hz), 5.39 (1H, dd, *J*=9.6, 10 Hz), 5.01 (1H, dd, *J*=9.6, 9.9Hz), 4.67, (1H, ddd, *J*=1.5, 3.9, 5.4Hz), 4.16 (2H, m), 3.95 (1H, d, *J*=12.0 Hz), 1.92 (3H, s), 1.91 (3H, s), 1.91 (3H, s), 1.90 (3H, s); ^{13}C NMR (90 MHz, $CDCl_3$) δ 169.98, 169.70, 169.64, 169.42, 88.19, 70.94, 76.73, 67.57, 66.78, 60.67, 20.47, 20.37, 20.31; LRMS (ESI): Calcd for $C_{14}H_{23}BrNO_9$ $[M+NH_4]^+$ 429.2, found 429.9.

6.4.1.3. 2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl bromide (2c)

Reaction yield: 89%.

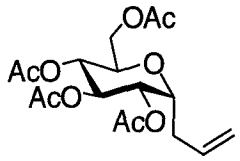


^{1}H NMR (360 MHz, $CDCl_3$) δ 6.56 (1H, d, *J*=3.9 Hz), 5.39 (1H, m), 5.22 (1H, dd, *J*=4.2, 8.2 Hz), 4.85 (1H, d, *J*=4.3 Hz), 4.82 (1H, d, *J*=4.2 Hz), 4.25 (1H, m), 3.97 (1H, m), 1.98 (3H, s), 1.98 (3H, s), 1.87 (3H, s), 1.82 (3H, s); ^{13}C NMR (90 MHz, $CDCl_3$) δ 169.98, 169.70, 169.64, 169.42, 88.19, 70.94, 67.73, 67.51, 66.78, 60.67, 20.47, 20.37, 20.31; LRMS (ESI): Calcd for $C_{14}H_{23}BrNO_9$ $[M+NH_4]^+$ 429.2, found 429.9.

6.4.2. General protocol for photochemical allylation of a pyranosyl bromide[2]

To a solution of 2,3,4,6-tetra-O-acetyl-α-D-pyranosyl bromide (1.24 g, 3.01 mmol) in benzene (15 mL), allyl phenyl sulfone (1.418 mL, 7.52 mmol) and bis-tributyl tin (2.19 mL, 4.21 mmol) were added. The solution was degassed and sonicated for 30 mins under argon atmosphere. The sealed flask was then irradiated for 9 hrs under a 450W mercury arc lamp and the reaction was monitored using TLC at time intervals of 2, 6, 8, 10 and 12 hrs. The reaction mixture was loaded directly onto a silica gel column packed with hexanes. The organostannanes were flushed with 3 void volumes of hexanes, after which the solvent system was changed to 5:1 hexanes/EtOAc to remove the unreacted allyl phenyl sulfone. The product was eluted with 3:1 hexanes/EtOAc and concentrated *in vacuo* to afford the allylated derivative as colorless oil (~80% yield).

6.4.2.1. 2,3,4,6-tetra-O-acetyl-1-deoxy-C-allyl-α-D-glucopyranoside (3a)

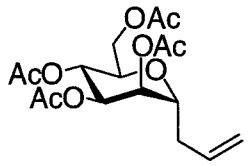


The product was co-eluted with starting material using 3:1 hexanes/EtOAc and concentrated under reduced pressure. The residue was re-dissolved in a 1:1 mixture of acetone and water, to which silver carbonate was added. The reaction mixture was allowed to stir for 30 mins, after which it was filtered. The product was extracted with CH_2Cl_2, washed with water, brine, dried over $MgSO_4$ and concentrated *in vacuo*. The colorless oil was purified by flash column chromatography using 3:1 hexanes/EtOAc, and afforded **7b** as a colorless oil in ~70% yield over two steps.

^{1}H NMR (360 MHz, $CDCl_3$) δ 5.76 (1H, dddd, *J*=6.6, 6.6, 8.0, 16.0 Hz), 5.41 (1H, dd, *J*=6.8, 13.2 Hz), 5.18 (2H, m), 5.01 (1H, dd, *J*=6.5, 12.1 Hz), 4.35 (1H, m), 4.23 (1H, dd, *J*=5.9, 12.8 Hz), 4.10 (1H, dd, *J*=3.0, 13.2 Hz), 3.89 (1H, m), 2.59 (1H, m), 2.38 (1H, m), 2.19 (3H, s), 2.01 (3H, s), 1.95 (3H, s), 1.91 (3H, s); ^{13}C NMR (90 MHz, $CDCl_3$) δ 177.3, 174.5, 170.6, 169.6, 169.4, 168.7, 163.7, 140.1, 115.3, 77.0, 76.6, 69.3, 68.8, 67.7, 67.47, 32.8, 26.6, 20.5, 20.5, 20.4, 20.2, 20.1, 20.1, 19.9; HRMS (ESI): Calcd for $C_{17}H_{25}O_9$ $[M+H]^+$ 373.1498, found 373.1461.

6.4.2.2. 2,3,4,6-tetra-O-acetyl-1-deoxy-C-allyl-α-D-mannopyranoside (3b)

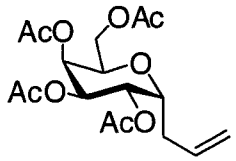
Reaction yield: 82%.



^{1}H NMR (360 MHz, $CDCl_3$) δ 5.77 (1H, dddd, J=6.8, 6.8, 10.0, 10.4 Hz), 5.23 (1H, dd, J=3.2, 6.8 Hz), 5.15 (1H, m), 5.07 (1H, m), 4.29 (1H, dd, J=6.4, 12.2 Hz), 4.06 (1H, dd, J=2.9, 8.2 Hz), 4.03 (1H, m), 3.86 (1H, ddd, J=2.9, 6.8, 8.6 Hz), 2.52 (1H, m), 2.29 (1H, m), 2.07 (3H, s), 2.02 (3H, s), 1.99 (3H, s), 1.94 (3H, s); ^{13}C NMR (90 MHz, $CDCl_3$) δ 170.6, 170.1, 169.9, 169.6, 132.4, 118.2, 74.0, 70.5, 69.9, 68.7, 66.8, 62.3, 33.4, 20.8, 20.7, 20.6; LRMS (ESI): Calcd for $C_{17}H_{25}O_9$ $[M+H]^+$ 373.4, found 373.1.

6.4.2.3. 2,3,4,6-tetra-O-acetyl-1-deoxy-C-allyl-α-D-galactopyranoside (3c)

Reaction yield: 88%.



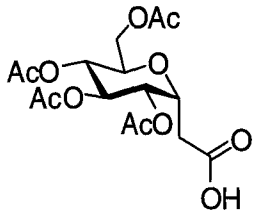
^{1}H NMR (360 MHz, $CDCl_3$) δ 5.76 (1H, dddd, J=6.5, 7.1, 10.0, 16.6 Hz), 5.42 (1H, dd, J=2.4, 3.1 Hz), 5.28 (1H, dd, J=4.8, 9.3 Hz), 5.22 (1H, dd, J=3.1, 9.3 Hz), 5.13 (1H, ddd, J=1.6, 3.0, 16.6 Hz), 5.12 (1H, ddd, J=1.4, 3.0, 10.0 Hz), 4.30 (1H, ddd, J=4.7, 5.2, 10.3 Hz), 4.21 (1H, dd, J=8.9, 12.5 Hz), 4.09 (2H, m), 2.46 (1H, m), 2.29 (1H, m), 2.12 (3H, s), 2.07 (3H, s), 2.04 (3H, s), 2.03 (3H, s); ^{13}C NMR (90 MHz, $CDCl_3$) δ 170.5, 170.1, 169.9, 169.8, 133.3, 117.6, 71.4, 68.2, 67.9, 61.4, 30.9, 20.8, 20.7, 20.6; LRMS (ESI): Calcd for $C_{17}H_{25}O_9$ $[M+H]^+$ 373.4, found 373.1.

6.4.3. General protocol for oxidation of allylated pyranose[3]

Allylated pyranose tetraacetate (0.83 g, 2.23 mmol) was dissolved in 14 mL of 2:2:3 acetonitrile/carbon tetrachloride/water mixture, followed by the addition of sodium periodate (1.90 g. 8.92 mmol). A catalytic amount of ruthenium trichloride trihydrate was then added, and the reaction was allowed to stir at room temperature for 2-3 hrs. The solution was filtered through celite and transferred to a separatory funnel with dichlroromethane. The organic layer was washed successively with saturated ammonium chloride solution, saturated brine solution, dried over $MgSO_4$, filtered and concentrated *in vacuo*. Flash column chromatography (0-10% methanol/CH_2Cl_2) afforded the carboxylic acid derivative in >75 % yield.

6.4.3.1. 2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-glucopyranosyl)-ethanoic acid (4a)

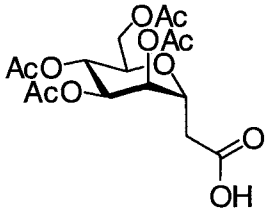
Reaction yield: 76%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 9.50-7.20 (1H, brs), 5.24 (1H, dd, *J*=8.5, 8.9 Hz), 5.14 (1H, dd, *J*=5.6, 9.1 Hz), 4.98 (1H, dd, *J*=8.5, 9.1 Hz), 4.69-4.62 (1H, m), 4.22 (1H, dd, *J*=5.1, 12.3 Hz), 4.09 (1H, dd, *J*=2.8, 12.3 Hz), 3.96-3.89 (1H, m), 2.78 (1H. dd, *J*=9.3, 15.5 Hz), 2.67 (1H, dd, *J*=5.4, 15.5 Hz), 2.05 (3H, s), 2.02 (6H, s), 2.01 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 175.1, 170.7, 170.0, 169.5, 169.4, 70.0, 69.9, 69.4, 69.1, 68.2, 61.9, 32.8, 20.6, 20.6, 20.5; HRMS (ESI): Calcd for $C_{16}H_{23}O_{11}$ $[M+H]^+$ 391.1240, found 391.1216.

6.4.3.2. 2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-mannopyranosyl)-ethanoic acid (4b)

Reaction yield: 82%.

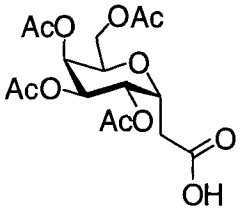


^{1}H NMR (300 MHz, $CDCl_3$) δ 9.50-7.50 (1H, brs), 5.23 (1H, dd, *J*=3.3, 7.2 Hz), 5.13, (1H, dd, *J*=3.3, 5.3 Hz), 5.08 (1H, dd, *J*=6.4, 6.8 Hz), 4.45-4.37 (2H, m), 4.13 (1H, dd, *J*=3.8, 12.1 Hz), 4.00-3.94 (1H, m), 2.71 (1H, dd, *J*=9.4, 15.5 Hz), 2.64 (1H, dd, *J*=4.5, 15.5 Hz), 2.07 (3H, s), 2.06 (3H, s), 2.05 (3H, s), 2.04 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 174.7, 170.8, 169.9,

169.7, 169.6, 71.9, 69.3, 69.1, 67.9, 67.2, 61.4, 35.3, 20.7, 20.6, 20.6; IR (thin film): 3300-2350, 1750 cm^{-1}; HRMS (ESI): Calcd for $C_{16}H_{23}O_{11}$ $[M+H]^+$ 391.1240, found 391.1228.

6.4.3.3. 2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl)-ethanoic acid (4c)

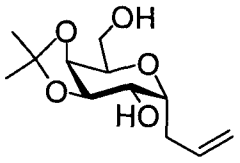
Reaction yield: 91%.



^1H NMR (400 MHz, $CDCl_3$) δ 5.43 (1H, t, *J*=2.8 Hz), 5.33 (1H, dd, *J*=8.9, 5.0 Hz,), 5.17 (1H, dd, *J* =8.9, 3.3 Hz), 4.70 (1H, ddd, *J*=9.3, 8.3, 5.3 Hz), 4.30-4.08 (3H, m), 2.73 (1H, dd, *J*=15.6, 8.7 Hz), 2.63 (1H, dd, *J*=15.6, 5.7 Hz), 2.13 (3H, s), 2.07 (3H, s), 2.04 (6H, s); ^{13}C NMR (90 MHz, $CDCl_3$) δ: 175.5, 170.8, 170.1, 170.0, 169.7, 69.4, 68.9, 67.8, 67.6, 67.1, 61.2, 33.15, 20.7; IR (thin film): 3706-2355, 1748 cm^{-1}.

6.4.4. Talose Synthesis[1a, 3, 4]

6.4.4.1. Deoxy-C-allyl-3,4-O-isopropylidene-α-D-galactopyranoside (5)

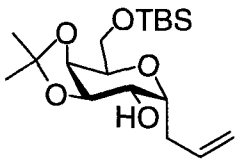


To a solution of allylgalactoside (0.81 g, 3.97 mmol) and dimethoxypropane (2.44 mL, 19.85 mmol) in dry acetonitrile (15 mL), p-TsOH (40 mg) was added and the reaction mixture was stirred for 2 hrs. Water (3 mL) was then added and after 30 mins the reaction was

neutralized with triethylamine and evaporated. Flash column chromatography (10% CH_2Cl_2/EtOAc) afforded **5** (0.69 g, 70%) as white crystalline solid.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.86 (1H, dddd, *J*=7.1, 7.1, 10.2, 17.2 Hz), 5.15 (1H, dd, *J*=1.8, 17.2 Hz), 5.09 (1H, dd, *J*=1.5, 17.2 Hz), 4.33-4.25 (2H, m), 4.08-4.02 (2H, m), 3.87-3.77 (2H, m), 2.45-2.30 (2H, m), 2.20 (1H, dd, *J*=3.2, 9.3 Hz), 2.00 (1H, d, *J*=4.5 Hz), 1.47 (3H, s), 1.32 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.3, 117.7, 109.7, 74.8, 73.1, 70.6, 69.6, 69.1, 63.6, 34.9, 26.8, 24.6; HRMS (ESI): Calcd for $C_{12}H_{21}O_5$ $[M+H]^+$ 245.1388, found 245.1377.

6.4.4.2. Deoxy-C-allyl-3,4-O-isopropylidene-6-O-tert-butyldimethylsilyl-α-D-galactopyranoside (6)

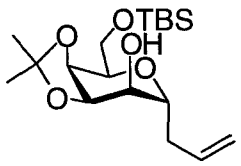


To a solution of diol **9** (0.30 g, 1.21 mmol) in dry pyridine/CH_2Cl_2 (5:5 mL), tert-butyldimethylsilyl chloride (0.22 g, 1.45 mmol) and DMAP (7 mg) were added sequentially. The reaction mixture was stirred overnight, diluted with CH_2Cl_2 (10 mL) and washed successively with 0.5 N HCl, water and brine. The organic extract was dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (25% EtOAc/hexanes) gave the product in quantitative yield (0.43 g, >95%).

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.82 (1H, dddd, *J*=7.0, 7.0, 10.2, 17.1 Hz), 5.16-5.10 (1H, m), 5.08-5.04 (1H, m), 4.34 (1H, dd, *J*=1.8, 7.1 Hz), 4.17 (1H, dd, *J*=3.4, 7.1 Hz), 3.99-3.93 (2H, m), 3.76-3.64 (3H, m), 2.45 (1H, br s), 2.35-2.28 (2H, m), 1.43 (3H, s), 1.28 (3H, s), 0.84 (9H, s),

0.02 (3H, s), 0.02 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.4, 117.5, 109.0, 74.6, 71.7, 70.5, 70.0, 69.7, 62.6, 35.0, 27.0, 25.9, 24.5, -5.3 ,-5.4; HRMS (ESI): Calcd for $C_{18}H_{35}O_5Si$ $[M+H]^+$ 359.2253, found 359.2257.

6.4.4.3. Deoxy-C-allyl-3,4-O-isopropylidene-6-O-tert-butyldimethylsilyl-α-D-talopyranoside (7)

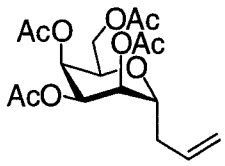


To a solution of oxalyl chloride (0.13 mL, 1.51 mmol) in CH_2Cl_2 (10 mL) at -78 °C, DMSO (0.23 mL, 3.31 mmol) was added dropwise and the reaction mixture was stirred for 30 mins. Then, **6** (0.27 g, 0.75 mmol) in CH_2Cl_2 (10 mL) was added over 10 mins. After an additional 30 mins of stirring, triethylamine (0.73 mL, 5.27 mmol) was added at -60 °C and the reaction mixture was brought to room temperature. After 1 hr, the reaction mixture was diluted with CH_2Cl_2 (10 mL) and water was added. The layers were separated and the organic extract was washed with water, brine, dried over $MgSO_4$, and concentrated under reduced pressure. The crude product was immediately dissolved in MeOH (20 mL) and $NaBH_4$ (0.06 g, 1.51 mmol) was added. After stirring for 30 mins, the reaction mixture was quenched with acetic acid and the solvent was evaporated. The residue was taken up in CH_2Cl_2 (30 mL), washed with water, brine and then dried over $MgSO_4$. The organic phase was concentrated *in vacuo* and after flash column chromatography (25% EtOAc/hexanes), compound **7** was obtained as an oil (0.41 g, 75%).

^1H NMR (400 MHz, $CDCl_3$) δ 5.93 (1H, dddd, *J*=6.7, 7.4, 10.2, 17.1 Hz), 5.19-5.07 (2H, m), 4.50 (1H, dd, *J*=3.6, 8.0 Hz), 4.42 (1H, dd, *J*=1.5, 8.0 Hz), 3.77 (1H, ddd, *J*=4.24, 6.50, 9.65 Hz),

3.74-3.57 (4H, m), 2.52-2.27 (2H, m), 1.99 (1H, br s), 1.51 (3H, s), 1.36 (3H, s), 0.89 (9H, s), 0.06 (3H, s), 0.06 (3H,s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 134.4, 117.6, 109.8, 73.4, 73.1, 72.5, 70.5, 68.2, 62.3, 36.9, 26.2, 25.9, 24.7, -5.3, -5.5; HRMS (ESI): Calcd for $C_{18}H_{38}NO_5Si$ $[M+NH_4]^+$ 376.2519, found 376.2508.

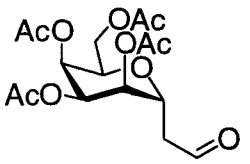
6.4.4.4. 2,3,4,6-tetra-O-acetyl-1-deoxy-C-allyl-α-D-talopyranoside (8a)



A solution of acetal **7** (0.22 g, 0.61 mmol) in 80% acetic acid (20 mL) was heated for 1 hr at 80 °C. The reaction mixture was cooled and the solvent was evaporated. The residue was taken up in pyridine (20 mL), followed by addition of acetic anhydride (10 mL) and DMAP (10 mg). The reaction mixture was stirred overnight, then diluted with CH_2Cl_2 (30 mL) and washed successively with saturated $CuSO_4$ solution, water and brine. The organic extract was dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography afforded compound **8a** as a white crystalline solid (0.17 g, 75%).

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.77 (1H, dddd, *J*=6.9, 6.9, 10.7, 16.6 Hz), 5.44 (1H, t, *J*=3.2 Hz), 5.18 (1H, dd, *J*=3.2, 5.1 Hz), 5.11-5.08 (1H, m), 5.08-5.06 (1H, m), 4.83 (1H, dd, *J*=3.2, 7.3 Hz), 4.62 (1H, dd, *J*=8.9, 12.1 Hz), 4.18 (1H, ddd, *J*=3.8, 4.8, 8.8 Hz), 4.09 (1H, dd, *J*=3.6, 12.2 Hz), 4.01 (1H, ddd, *J*=4.8, 7.5, 7.5 Hz), 2.39-2.23 (2H, m), 2.08 (3H, s), 2.05 (3H, s), 2.04 (3H, s), 2.02 (3H, s); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.6, 169.6, 169.4, 169.3, 132.8, 117.7, 70.5, 69.2, 68.7, 67.1, 66.4, 60.3, 34.6, 20.7, 20.6, 20.5; HRMS (ESI): Calcd for $C_{17}H_{28}NO_9$ $[M+NH_4]^+$ 390.1764, found 390.1777.

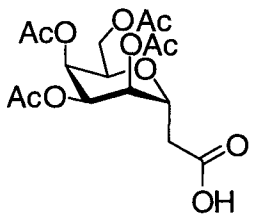
6.4.4.5. 2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-talopyranosyl) ethanal (8b)



Ozone gas was bubbled through a solution of **8a** (0.86 g, 2.33 mmol) in 20 mL of dry CH_2Cl_2 at -78 °C until the solution turned blue in color. Nitrogen was then bubbled through the solution until it turned colorless. Triphenylphosphine (1.5 g, 5.7 mmol) was then added, the solution was allowed to warm up to room temperature and it was stirred overnight. Dichloromethane was removed under reduced pressure and the residue was purified by flash column chromatography (toluene/acetone 4:1) to afford **8b** in 85% yield.

^{1}H NMR (500 MHz, $CDCl_3$) δ 9.70 (1H, t, *J*=2.1 Hz), 5.56 (1H, t, *J*=3.0 Hz), 5.18 (1H, dd, *J*=3.0, 6.3 Hz), 4.90 (1H, dd, *J*=9.8, 12.6 Hz), 4.77 (1H, dd, *J*=3.0, 9.3 Hz), 4.53 (1H, ddd, *J*=5.2, 7.4, 9.2 Hz), 4.23 (1H, ddd, *J*=2.7, 6.3, 9.6 Hz), 4.07 (1H, dd, *J*=2.8, 12.7 Hz), 2.57-2.54 (2H, m), 2.15 (3H, s), 2.07 (3H, s), 2.05 (3H, s), 2.00 (3H, s); ^{13}C NMR (125 MHz, $CDCl_3$) δ 199.1, 171.0, 169.6, 169.3, 169.1, 71.8, 69.1, 67.4, 66.5, 63.3, 59.5, 44.9, 20.8, 20.8, 20.6, 20.5; HRMS (ESI): Calcd for $C_{16}H_{26}NO_{10}$ $[M+NH_4]^+$ 392.1556, found 392.1568.

6.4.4.6. 2-(2,3,4,6-tetra-O-acetyl-α-D-talopyranosyl) ethanoic acid (9)



To a solution of **8b** (2.0 g, 5.4 mmol) in 1:1 (v:v) mixture of *tert*-butyl alcohol and 2-methyl-2-butene (20 mL), an aqueous solution of potassium dihydrogen phosphate (3.7 g, 27.2 mmol) and sodium chlorite (2.4 g, 26.5 mmol) in 10 mL distilled water was added. The mixture was stirred vigorously at 0 °C for 6 hrs. The product was extracted with EtOAc and washed with 100 mL water and 50 mL brine solution. The organic layer was then dried over $MgSO_4$ and concentrated *in vacuo* to afford **9** in 96% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.57 (1H, t, *J*=3.0 Hz), 5.20 (1H, dd, *J*=3.1, 6.2 Hz), 4.83 (1H, dd, *J*=3.0, 9.2 Hz), 4.76 (1H, dd, *J*=9.5,12.5 Hz), 4.38 (1H, ddd, *J*=3.7, 8.9, 8.9 Hz), 4.29 (1H, ddd, *J*=3.0, 6.2, 9.3 Hz), 4.17 (1H, dd, *J*=3.0, 12.6 Hz), 2.62 (1H, dd, *J*=3.7, 15.8 Hz), 2.53 (1H, dd, *J*=8.8, 15.8 Hz), 2.16 (3H, s), 2.08 (3H, s), 2.07 (3H, s), 2.03 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 173.9, 171.0, 169.7, 169.3, 169.2, 71.6, 68.9, 67.4, 66.5, 65.0, 59.8, 36.1, 20.9, 20.7, 20.6, 20.6; LRMS (ESI): Calcd for $C_{16}H_{23}O_{11}$ $[M+H]^+$ 391.3, found 391.4; HRMS (ESI): Calcd for $C_{16}H_{26}NO_{11}$ $[M+NH_4]^+$ 408.1505, found 408.1512.

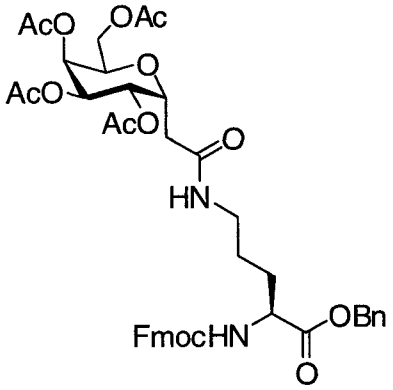
6.4.5. General protocol for amide coupling of C-linked glycoconjugate[1]

To a solution of fully protected amino acid derivative **10** (0.21 g, 0.38 mmol) in 10 mL of CH_2Cl_2, 2 mL of TFA was added. The reaction was stirred for 40 mins and then concentrated under reduced pressure. The syrup was re-dissolved in a 1:1 mixture of toluene and CH_2Cl_2, concentrated, then re-dissolved in diethyl ether and concentrated *in vacuo* to afford a white solid. To a solution of carboxylic acid carbohydrate derivative (0.15 g, 0.38 mmol) in 15 mL of CH_2Cl_2, HBTU (0.17 g, 0.46 mmol) and 0.20 mL of DIPEA (0.15 g, 1.14 mmol) were added. The reaction was allowed to stir for 20 mins and then the activated amino acid derivative was

added. After stirring overnight, the reaction mixture was transferred to a separatory funnel and washed successively with saturated ammonium chloride, water, brine, dried over $MgSO_4$ and concentrated *in vacuo*. The product was purified by flash column chromatography in 50:1 CH_2Cl_2/MeOH to afford a C-linked glycosyl amide building block in ~80% yield.

6.4.5.1. Benzyl Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-galactopyranoside)acetyl]-L-ornithine (11)

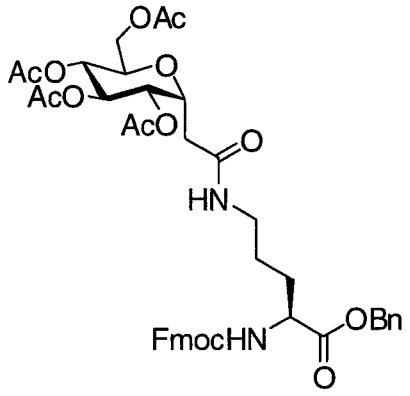
Reaction yield: >95%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (2H, d, *J*=7.5 Hz), 7.59 (2H, d, *J*=7.21 Hz), 7.42-7.28 (9H, m), 6.12 (1H, br s), 5.55 (1H, d, *J*=8.4 Hz), 5.4 (1H, t, *J*=3.3 Hz), 5.29-5.23 (1H, m), 5.17-5.12 (3H, m), 4.67-4.61 (1H, m), 4.46-4.34 (3H, m), 4.25-4.19 (2H, m), 4.15-4.09 (2H, m), 3.29-3.23 (2H, m), 2.57-2.47 (1H, m), 2.41-2.34 (1H, m), 2.10 (3H, s), 2.05 (3H, s), 2.04 (3H, s), 2.00 (3H, s), 1.93-1.82 (1H, m), 1.73-1.47 (3H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.9, 170.5, 169.9, 169.7, 169.5, 169.2, 156.0, 143.6, 141.2, 135.0, 128.6, 128.5, 127.7, 127.0, 125.0, 119.9, 69.4, 68.7, 67.8, 67.7, 67.2, 67.0, 66.8, 61.0, 53.5, 47.0, 38.6, 34.4, 30.0, 25.3, 20.7, 20.6; LRMS (ESI): Calcd for $C_{43}H_{49}N_2O_{14}$ $[M+H]^+$ 817.8, found 817.4.

6.4.5.2. Benzyl Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-glucopyranoside)acetyl]-L-ornithine (12)

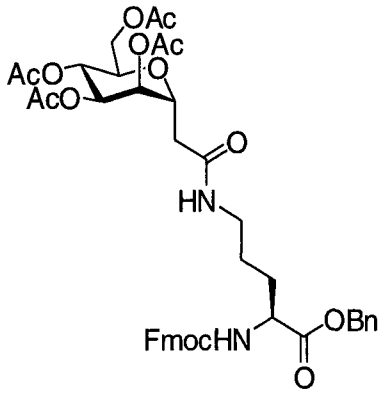
Reaction yield: >95%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (2H, d, *J*=7.4 Hz), 7.60 (2H, d, *J*=7.1 Hz), 7.45-7.25 (9H, m), 6.17 (1H, br s), 5.64 (1H, d, *J*=7.9 Hz), 5.23 (1H, t, *J*=8.4 Hz), 5.20-5.04 (3H, m), 4.97 (1H, t, *J*=8.3 Hz), 4.25-4.12 (3H, m), 3.97-3.90 (1H, m), 3.35-3.17 (2H, m), 2.60 (1H, dd, *J*=9.6, 12.4 Hz), 2.45 (1H, dd, *J*=4.4, 15.4 Hz), 2.07-1.98 (12H, m), 1.95-1.77 (1H, m), 1.76-1.60 (1H, m), 1.60-1.46 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.9, 170.6, 169.9, 169.4, 169.0, 156.0, 143.8, 143.6, 141.2, 135.1, 128.6, 128.5, 128.5, 128.3, 127.7, 127.0, 125.0, 119.9, 69.9, 69.7, 69.4, 69.3, 68.2, 67.3, 67.0, 62.0, 53.5, 47.1, 38.9, 38.6, 34.2, 30.0, 25.3, 20.7, 20.6, 20.6; HRMS (ESI): Calcd for $C_{43}H_{49}N_2O_{14}$ $[M+H]^+$ 817.3183, found 817.3192.

6.4.5.3. Benzyl Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside)acetyl]-L-ornithine (13)

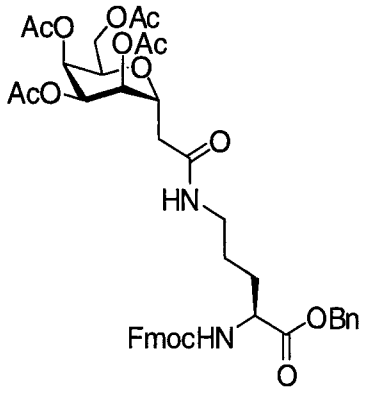
Reaction yield: >95%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (2H, d, *J*=7.4 Hz), 7.60 (2H, d, *J*=7.1 Hz), 7.44-7.25 (9H, m), 6.18 (1H, br s), 5.71 (1H, d, *J*=8.1 Hz), 5.25 (1H, dd, *J*=3.2, 7.2 Hz), 5.20-5.07 (3H, m), 4.47-4.30 (5H, m), 4.25-4.15 (2H, m), 4.02-3.94 (1H, m), 3.40-3.20 (2H, m), 2.52 (2H, d, *J*=6.87 Hz), 2.10-2.06 (12H, m), 2.00-1.82 (1H, m), 1.80-1.65 (1H, m), 1.64-1.50 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.0, 170.6, 169.9, 169.8, 169.5, 168.8, 156.1, 143.8, 143.6, 141.2, 135.1, 128.6, 128.6, 128.5, 128.5, 128.3, 127.7, 127.6, 127.0, 126.9. 125.0, 119.9, 72.0, 69.9, 69.1, 68.0, 67.3, 67.2, 67.0, 65.2, 61.7, 53.6, 47.1, 38.9, 38.6, 37.3, 29.9, 25.5, 20.7, 20.7, 20.6; HRMS (ESI): Calcd for $C_{43}H_{49}N_2O_{14}$ $[M+H]^+$ 817.3183, found 817.3198.

6.4.5.4. Benzyl Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-talopyranoside)acetyl]-L-ornithine (14)

Reaction yield: >95%.



^{1}H NMR (500 MHz, $CDCl_3$) δ 7.69 (2H, d, *J*=7.3 Hz), 7.52 (2H, d, *J*=7.8 Hz), 7.35-7.17 (9H, m), 6.18-6.11 (1H, m), 5.54 (1H, d, *J*=7.6 Hz), 5.46 (1H, t, *J*=3.3 Hz), 5.14-5.05 (3H, m), 4.70 (1H, dd, *J*=2.6, 9.0 Hz), 4.55 (1H, dd, *J*=9.8, 12.6 Hz), 4.40-4.00 (7H, m), 3.23-3.10 (2H, m), 2.37 (1H, dd, *J*=2.7, 15.1 Hz), 2.23 (1H, dd, *J*=9.0, 14.8 Hz), 2.06 (3H, s), 1.99 (3H, s), 1.96 (3H, s), 1.93 (3H, s), 1.87-1.77 (1H, m), 1.67-1.57 (1H, m), 1.53-1.42 (2H, m); ^{13}C NMR (125 MHz, $CDCl_3$) δ 172.0, 170.7, 169.6, 169.4, 169.2, 169.2, 156.0, 143.8, 143.6, 141.2, 135.1, 128.6, 128.5, 128.2, 127.6, 127.0, 125.0, 125.0, 119.9, 71.7, 69.0, 67.4, 67.2, 67.0, 66.4, 65.6, 60.1, 53.6, 47.0, 38.9, 38.5, 38.2, 29.9, 25.4, 20.8, 20.8, 20.6; HRMS (ESI): Calcd for $C_{43}H_{49}N_2O_{14}$ $[M+H]^+$ 817.3183, found 817.3149.

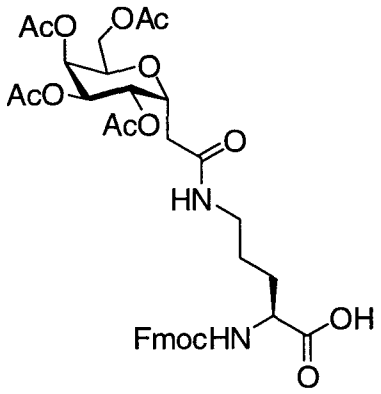
6.4.6. General protocol for debenzylation of a *C*-linked glycoconjugate[1]

To a solution of benzyl ester in 1:1 ethanol/ethyl acetate, a catalytic amount (8% w/w) of palladium on charcoal (10% w/w) was added. The reaction mixture was stirred under H_2 atmosphere and monitored via TLC. After the reaction was completed, the catalyst was gravity filtered through Whatman No.2 filter paper. The combined filtrates were concentrated *in vacuo*

and the compound was purified by column chromatography (5% MeOH/CH_2Cl_2) to afford a crystalline solid in >80% yield.

6.4.6.1. Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-galactopyranoside)acetyl]-L-ornithine (15)

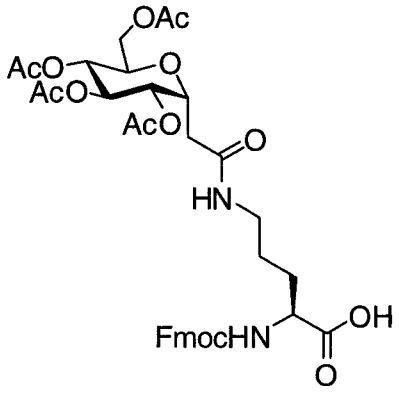
Reaction yield: 84%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.76 (2H, d, *J*=7.2 Hz), 7.60 (2H, d, *J*=6 Hz), 7.42-7.37 (2H, m), 7.33-7.28 (2H, m) 6.25 (1H, br s), 5.6 (1H, d, *J*=7.5 Hz), 5.39 (1H, t, *J*=3.3 Hz), 5.27-5.23 (1H, m), 5.16-5.12 (1H, m), 4.68-4.63 (1H, m), 4.39 (2H, d, *J*=6.6), 4.33-4.05 (5H, m), 3.37-3.25 (2H, m), 2.60 (1H, dd, *J*=15.5, 9.8 Hz), 2.44 (1H, dd, *J*=15.5, 4.2 Hz), 2.11 (3H, s), 2.05 (3H, s), 2.04 (3H, s), 2.02 (3H, s), 1.98-1.86 (1H, m), 1.82-1.69 (1H, m), 1.70-1.54 (2H, m); ^{13}C NMR (100 MHz, $CDCl_3$) δ 174.4, 172.4, 170.9, 170.4, 170.0, 169.8, 156.1, 143.6, 141.2, 127.7, 127.0, 125.0, 119.9, 69.1, 68.9, 67.8, 67.7, 66.9, 61.2, 53.4, 50.6, 47.0, 39.1, 34.1, 31.7, 28.6, 22.1, 20.7; IR (thin film): 3367, 1748 cm^{-1}; LRMS (ESI): Calcd for $C_{36}H_{43}N_2O_{14}$ $[M+H]^+$ 726.3, found 727.0.

6.4.6.2. Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-glucopyranoside)acetyl]-L-ornithine (16)

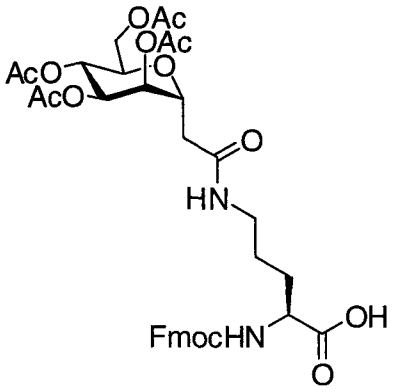
Reaction yield: 85%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75 (2H, d, *J*=7.6 Hz), 7.63-7.55 (2H, m), 7.38 (2H, t, *J*=7.4 Hz), 7.33-7.25 (2H, m), 6.64-6.53 (1H, m), 5.83 (1H, d, *J*=7.4 Hz), 5.24 (1H, t, *J*=8.3 Hz), 5.07 (1H, dd, *J*=5.2, 8.6 Hz), 4.98 (1H, t, *J*=8.3 Hz), 4.70-4.60 (1H, m), 4.45-4.30 (3H, m), 4.25-4.12 (2H, m), 4.05-3.90 (3H, m), 3.40-3.20 (2H, m), 2.66 (2H, dd, *J*=9.9, 15.2 Hz), 2.49 (2H, dd, *J*=4.3, 15.3 Hz), 2.05 (3H, s), 2.03 (3H, s), 2.02 (3H, s), 2.01 (3H, s), 1.97-1.82 (1H, m), 1.81-1.67 (1H, m), 1.67-1.51 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 171.0, 170.0, 169.6, 169.5, 156.3, 143.8, 143.7, 141.3, 127.7, 127.1, 125.1, 120.0, 77.2, 69.9, 69.7, 69.5, 69.4, 68.2, 67.1, 62.1, 53.5, 47.1, 39.1, 34.2, 29.8, 25.2, 20.7, 20.7, 20.6, 20.6; HRMS (ESI): Calcd for $C_{36}H_{43}N_2O_{14}$ $[M+H]^+$ 727.2714, found 727.2739.

6.4.6.3. Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-mannopyranoside)acetyl]-L-ornithine (17)

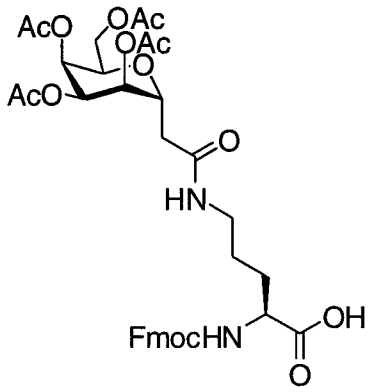
Reaction yield: 93%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.74 (2H, d, *J*=7.4 Hz), 7.58 (2H, dd, *J*=4.8, 7.1 Hz), 7.37 (2H, dd, 7.2, 7.9 Hz), 7.27 (2H, dt, *J*=1.1, 7.7, 7.8 Hz), 6.72 (1H, br m), 5.94 (1H, d, *J*=7.5 Hz), 5.24 (1H, dd, *J*=3.1, 7.4 Hz), 5.15-5.08 (2H, m), 4.45-4.30 (4H, m), 4.22-4.12 (2H, m), 4.02-3.92 (1H, m), 3.38-3.20 (2H, m), 2.65-2.45 (2H, m), 2.05 (3H, s), 2.04 (3H, s), 2.03 (6H, s), 1.97-1.82 (1H, m), 1.80-1.65 (1H, m), 1.65-1.52 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 174.5, 170.9, 170.2, 169.9, 169.8, 169.5, 156.3, 143.7, 143.6, 141.2, 127.7, 127.0, 125.0, 119.9, 77.2, 71.7, 70.1, 69.3, 68.0, 67.2, 67.0, 61.8, 53.4, 47.0, 39.1, 36.9, 29.6, 25.2, 20.7, 20.7, 20.7, 20.6; HRMS (ESI): Calcd for $C_{36}H_{43}N_2O_{14}$ $[M+H]^+$ 727.2714, found 727.2715.

6.4.6.4. Fluorenylmethoxycarbonyl-[2-(deoxy-2,3,4,6-tetra-O-acetyl-α-D-talopyranoside)acetyl]-L-ornithine (18)

Reaction yield: 93%.



^{1}H NMR (400 MHz, $CDCl_3$) δ 9.3-8.2 (1H, br s), 7.74 (2H, d, *J*=7.4 Hz), 7.59 (2H, d, *J*=6.9 Hz), 7.33 (4H, dt, *J*=7.1, 28.6 Hz), 6.73 (1H, br s), 5.99 (1H, br d, *J*=5.6 Hz), 5.52 (1H, br s), 5.16 (1H, br dd, *J*=3.0, 5.2), 4.80 (1H, br dd, *J*=2.2, 8.3 Hz), 4.62 (1H, br dd, *J*=9.8, 12.6 Hz), 4.50-4.11 (6H, m), 3.4-3.15 (2H, br m), 2.58-2.30 (2H, br m), 2.12 (3H, s), 2.06 (3H, s), 2.03 (3H, s), 2.00 (3H, s), 1.96-1.83 (1H, m), 1.80-1.66 (1H, m), 1.66-1.52 (2H, m); ^{13}C NMR (100 MHz, $CDCl_3$) δ 171.0, 170.2, 169.7, 169.5, 156.3, 143.8, 143.6, 141.2, 127.7, 127.0, 125.1, 120.0, 71.5, 68.9, 67.3, 66.9, 66.4, 66.0, 60.3, 47.0, 39.0, 37.9, 29.7, 25.3, 20.8, 20.8, 20.6; HRMS (ESI): Calcd for $C_{36}H_{43}N_2O_{14}$ $[M+H]^+$ 727.2714, found 727.2701.

6.4.7. General protocol for automated and manual solid phase synthesis of C-linked analogues[5]

Automated solid phase peptide synthesis (SPPS) was performed on APEX 396 equipped with a 40 well reaction vessel. Manual SPPS was carried out in SPS flask with a 3 way valve (drain, argon, and outlet). All polypeptides were prepared by linear solid-phase synthesis using standard Fmoc chemistry.

A typical procedure started from Fmoc-glycine Wang resin with loading capacities of ~0.066 mmol/g. The resin was swollen in DMF for 30 mins. The solvent was then drained and 20% piperidine solution in DMF was added. The solution was allowed to stir for 1 hr, then it was drained, and the resin was washed with three aliquots of DMF. Kaiser and TNBS tests for free amine were then performed. In a separate flask, building block (1.5 eq) was premixed with 1.5 eq of HBTU and 1.5 eq of DIPEA in DMF. The reaction mixture was stirred for 30 mins, transferred to the SPS flask and stirred for 24 hrs. The flask was drained and the resin was rinsed three times with DMF. Kaiser and TNBS tests were performed to verify that coupling had reached completion (negative test results). The resin was treated with 20% piperdine in DMF solution for 1 hr, the flask was drained and the resin was washed thoroughly with DMF. Kaiser and TNBS tests were performed to ensure the presence of free amine (positive test results). In a separate flask, 5 eq of the successive building block (commercially available amino acid) were premixed with 5 eq of HBTU and 5 eq of DIPEA in DMF for 20 mins. The reaction solution was then transferred to the SPS flask with the resin, and the reaction mixture was allowed to stir for 4 hrs. The Fmoc deprotection and coupling steps were repeated until twelve amino acid residues (four repeating tripeptides) were coupled to the resin. To cleave the glycopeptide from the resin, the resin was successively rinsed with DMF, MeOH, and CH_2Cl_2, and stirred for 2 hrs in 1:1 (v:v) trifluoroacetic acid/CH_2Cl_2. The resin was filtered off, the solution was concentrated under reduced pressure, and the product was crystallized from diethyl ether to produce a white powder. The glycopolymer was dissolved in a sodium methoxide (pH=10) solution (4:1 sodium methoxide/ water) and stirred overnight. The solution was neutralized with Dowex® 50WX8 ion exchange resin, filtered and concentrated *in vacuo* to remove MeOH. Purification via HPLC using reverse phase chromatography (C-18 column, liquid phase: 0.1% TFA in acetonitrile/water) was performed to obtain the desired glyopeptide. Acetonitrile, water and 0.1%

TFA were evaporated under a flow of air. The residue was then dissolved in water which was lyophilized to give the product as white, fluffy solid.

6.4.7.1. [L-Ornithine(galactose)-glycine-glycine]$_4$-glycine (19)

Reaction yield: 78%.

^{1}H NMR (300 MHz, D_2O) δ 4.42-4.32 (4H, m), 4.28-4.18 (4H, m), 4.00-3.80 (29H, m), 3.75-3.67 (4H, m), 3.67-3.47 (14H, m), 3.20-3.05 (8H, m), 2.62-2.42 (8H, m), 1.85-1.70 (5H, m), 1.70-1.57 (4H, m), 1.57-1.37 (9H, m); ^{13}C NMR (75 MHz, D_2O) δ 175.0, 174.9, 174.1, 173.9, 173.6, 172.3, 172.3, 172.2, 172.0, 171.9, 171.8, 170.7, 163.6, 163.1, 73.3, 72.8, 70.0, 69.1, 67.9, 61.2, 54.1, 53.2, 42.9, 42.7, 41.5, 41.4, 39.1, 32.7, 28.4, 25.2; LRMS (MALDI-TOF): Calcd for $C_{70}H_{117}N_{17}O_{38}$ $[M+H]^+$ 1805.8, found 1805.3.

6.4.7.2. [L-Ornithine(glucose)-glycine-glycine]$_4$-glycine (20)

Reaction yield: 76%.

^{1}H NMR (360 MHz, D_2O) δ 4.47 (32H, s), 4.08 (2H, br s), 3.73 (6H, m), 3.32 (10H, m), 3.01 (6H, m), 2.45 (2H, m), 2.19 (18H, s), 1.98 (8H, s), 1.68 (8H, m), 1.30 (4H, m), 0.87 (2H, m); ^{13}C NMR (90 MHz, D_2O) δ 187.76, 186.77, 185.19, 184.83, 184.24, 180.00, 159.30, 143.77, 142.41, 142.34, 139.011, 138.55, 97.19, 86.21, 85.68, 82.81, 81.66, 80.76, 80.63, 74.08, 71.31, 66.13, 60.96, 55.79, 55.60, 51.98, 45.61, 45.43, 38.05, 36.468, 19.55; LRMS (ESI): Calcd for $C_{70}H_{117}N_{17}O_{38}$ $[M+H]^+$ 1805.8, found 1805.2

6.4.7.3. [L-Ornithine(mannose)-glycine-glycine]$_4$-glycine (21)

Reaction yield: 79%.

^{1}H NMR (500 MHz, D_2O) δ 4.83 (32H, m), 4.45 (4H, m), 4.07 (6H, m), 3.93 (8H, m), 3.65 (4H, m), 3.46 (4H, m), 3.21 (3H, m), 2.98 (1H, m), 2.80 (4H, m), 2.68 (1H, m), 2.58 (1H, m), 2.25 (4H, m), 2.07 (16H, m), 1.79 (4H, m), 1.60 (8H, m), 1.33 (8H, m), 1.06 (4H, m); ^{13}C NMR (125 MHz, D_2O) δ 181.01, 170.59, 117.47, 117.05, 74.60, 74.33, 74.07, 70.34, 70.06, 66.70, 62.04, 60.51, 42.73, 42.07, 41.95, 38.29, 35.01, 34.79, 29.77, 27.69, 24.29, 22.81; LRMS (ESI): Calcd for $C_{70}H_{117}N_{17}O_{38}$ $[M+H]^+$ 1805.8, found 1804.1.

6.4.7.4. [L-Ornithine(2,3,4,6-peracetylated talose)-glycine-glycine]$_4$-glycine

Reaction yield: 81%.

^{1}H NMR (300 MHz, MeOD) δ 5.51 (1H, br s), 5.21 (1H, br s), 4.65-4.53 (1H, m), 4.46-3.82 (10H, m), 3.24 (3H, m), 2.60-2.45 (2H, m), 2.11 (3H, s), 2.08 (3H, s), 2.06 (6H, s), 1.97-1.83 (1H, m), 1.83-1.72 (1H, m), 1.72-1.52 (2H, m); ^{13}C NMR (75 MHz, MeOD) δ 174.3, 174.0, 171.9, 171.5, 171.4, 171.1, 170.8, 170.5, 170.5, 170.4, 71.1, 69.4, 68.1, 67.5, 67.0, 60.9, 53.9, 53.9, 53.2, 42.7, 40.9, 39.2, 38.9, 37.6, 29.8, 28.7, 25.9, 25.0, 24.9, 22.8, 20.0, 19.8, 19.7, 19.7; LRMS (MALDI-TOF): Calcd for $C_{102}H_{149}N_{17}O_{54}$ $[M+H]^+$ 2478.4, found 2476.8.

6.4.7.5. [L-Ornithine(talose)-glycine-glycine]$_4$-glycine (22)

Reaction yield: 86%.

^{1}H NMR (500 MHz, D_2O) δ 4.27-4.17 (5H, m), 3.92-3.72 (20H, m), 3.67-3.55 (7H, m), 3.45-3.40 (1H, m), 3.14-3.07 (6H, m), 2.54-2.40 (6H, m), 1.80-1.70 (2H, m), 1.70-1.55 (3H, m), 1.55-1.37 (6H, m); ^{13}C NMR (125 MHz, D_2O) δ 174.4, 172.6, 171.8, 171.5, 170.9, 99.9, 73.9, 72.9, 72.8, 70.6, 68.8, 66.8, 60.0, 53.8, 53.5, 43.0, 42.4, 42.3, 42.2, 38.8, 38.6, 36.0, 28.0, 24.7, 24.6, 24.4; LRMS (MALDI-TOF): Calcd for $C_{70}H_{117}N_{17}O_{38}$ $[M+H]^+$ 1805.8, found 1804.8.

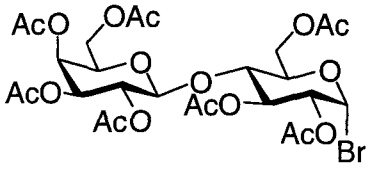
6.5. Synthetic Procedures B

6.5.1. General protocol for bromination of a disaccharide

To a solution of peracetylated disaccharide in CH_2Cl_2 at 0 °C, an equal volume of HBr in AcOH (33% solution) was added. The ice bath was removed and the reaction was allowed to stirr at room temerature. The progress of reaction was monitored via TLC. After completion, the reaction was diluted with CH_2Cl_2, poured over ice and transferred to a separatory funnel. The organic layer was washed with water until neutral pH. The organic extract was dried over $MgSO_4$, filtered and concentrated *in vacuo* to afford product in >95% yield. The crude product was crystallized from diethyl ether to form a white powder.

6.5.1.1. 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->4)-2,3,6-tri-O-acetyl-α-D-glucopyranosyl bromide (24a)

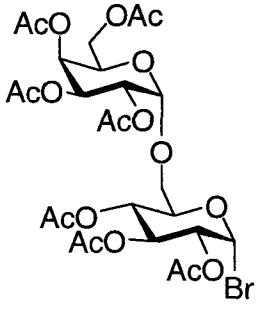
Reaction yield: >95%.



^{1}H NMR (500 MHz, $CDCl_3$) δ 6.47 (1H, d, *J*=4.0 Hz), 5.49 (1H, t, *J*=9.6 Hz), 5.29 (1H, dd, *J*=0.9, 3.5 Hz), 5.06 (1H, dd, *J*=7.9, 10.4 Hz), 4.9 (1H, d, *J*=10.4, 3.5 Hz), 4.70 (1H, dd, *J*=4.0, 10.0 Hz), 4.48-4.42 (2H, m), 4.14-4.06 (3H, m), 4.02 (1H, dd, *J*=7.2, 11.2 Hz), 3.87-3.78 (2H, m), 2.10 (3H, s), 2.07 (3H, s), 2.03 (3H, s), 2.01 (3H, s), 2.00 (3H, s), 1.99 (3H, s), 1.91 (3H, s); ^{13}C NMR (125 MHz, $CDCl_3$) δ 170.2, 170.0, 170.0, 169.9, 169.8, 169.1, 168.8, 100.6, 86.3, 74.8, 72.8, 70.8, 70.7, 70.6, 69.4, 68.9, 66.5, 60.9, 60.7, 20.7, 20.6, 20.5, 20.3; LRMS (ESI): Calcd for $C_{26}H_{39}BrNO_{17}$ $[M+NH_4]^+$ 717.4, found 1418.2, 1416.3, 718.4, 716.5.

6.5.1.2. 2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl-(1-->6)-2,3,4-tri-O-acetyl-α-D-glucopyranosyl bromide (24b)

Reaction yield: >95%.

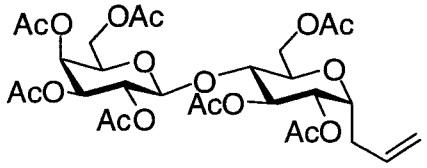


^{1}H NMR (300 MHz, $CDCl_3$) δ 5.90 (1H, d, *J*=4.0 Hz), 5.60-5.50 (1H, m), 5.48-5.44 (1H, m), 5.33 (1H, dd, *J*=3.3, 10.8), 5.22-5.12 (2H, m), 5.08 (1H, dd, *J*=3.7, 10.8 Hz), 4.79 (1H, dd, *J*=4.0, 10.0 Hz), 4.28-4.13 (2H, m), 4.10-4.03 (1H, m), 3.77 (1H, dd, *J*=4.4, 12.0 Hz), 3.63 (1H, dd, *J*=2.2, 12.0 Hz), 2.13 (3H, s), 2.12 (3H, s), 2.10 (3H, s), 2.07 (3H, s), 2.05 (3H, s), 2.04 (3H, s), 1.99 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.9, 170.8, 170.6, 170.4, 170.3, 170.2, 169.8, 96.6, 86.9, 73.3, 71.0, 70.6, 68.4, 68.3, 68.0, 67.8, 66.8, 65.8, 62.0, 21.2, 21.1, 21.1, 21.0, 21.0; LRMS (ESI): Calcd for $C_{26}H_{39}BrNO_{17}$ $[M+NH_4]^+$ 717.4, found 1418.1, 1416.2, 718.3, 716.3.

6.5.2. 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->4)-2,3,6-tri-O-acetyl-1-deoxy-C-allyl-α-D-glucopyranoside (25a)

For procedure please refer to Section **6.4.2.**

Reaction yield: 78%.

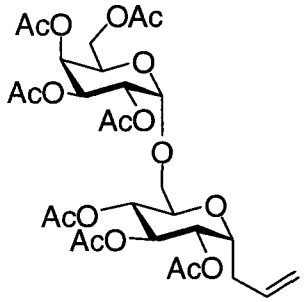


^{1}H NMR (300 MHz, $CDCl_3$) δ 5.74 (1H, dddd, *J*=6.9, 6.9, 9.9, 17.5 Hz), 5.40-5.32 (2H, m), 5.20-5.07 (3H, m), 5.02-4.94 (2H, m), 4.52 (1H, d, *J*=4.5 Hz), 4.35 (1H, dd, *J*=11.7, 2.6 Hz), 4.23-4.07 (4H, m), 3.94-3.86 (1H, m), 3.85-3.77 (1H, m), 3.70-3.62 (1H, m), 2.60-2.46 (1H, m), 2.38-2.26 (1H, m), 2.15 (3H, s), 2.10 (3H, s), 2.08-2.06 (9H, m), 1.97 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.9, 170.8, 170.6, 170.5, 170.4, 170.0, 169.6, 133.5, 118.2, 101.7, 77.7, 71.9, 71.4, 71.1, 70.4, 70.2, 69.5, 67.1, 62.7, 61.3, 31.4, 21.3, 21.3, 21.2, 21.1, 21.1, 20.9; HRMS (ESI): Calcd for $C_{29}H_{44}NO_{17}$ $[M+NH_4]^+$ 678.2594, found 678.2609.

6.5.3. 2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl-(1-->6)-2,3,4-tri-O-acetyl-1-deoxy-C-allyl-α-D-glucopyranoside (25b)

For procedure please refer to Section **6.4.2.**

Reaction yield: 84%.

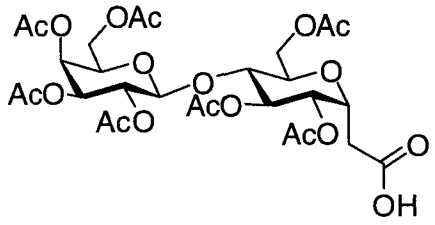


^{1}H NMR (300 MHz, $CDCl_3$) δ 5.85-5.66 (1H, m), 5.47-5.41 (1H, m), 5.38-5.27 (2H, m), 5.25-5.19 (1H, m), 5.17-5.07 (3H, m), 5.02 (1H, dd, *J*=5.8, 9.3 Hz), 4.97-4.87 (1H, m), 4.35-4.27 (1H, m), 4.27-4.18 (1H, m), 4.13 (1H, dd, *J*=6.1, 11.2 Hz), 4.00 (1H, dd, *J*=7.0, 11.2 Hz), 3.90-3.82 (1H, m), 3.68 (1H, dd, *J*=6.2, 10.8 Hz), 3.51 (1H, dd, *J*=2.6, 10.8 Hz), 2.65-2.47 (1H, m), 2.37-2.25 (1H, m), 2.13 (3H, s), 2.10 (3H, s), 2.08-2.01 (12H, m), 1.98 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.9, 170.8, 170.6, 170.6, 170.3, 170.1, 170.0, 133.3, 118.6, 96.3, 77.7, 72.2, 70.8, 70.8, 70.0, 69.8, 68.6, 68.4, 67.8, 67.0, 66.8, 62.2, 30.8, 21.2, 21.1, 21.1, 21.1; HRMS (ESI): Calcd for $C_{29}H_{44}NO_{17}$ $[M+NH_4]^+$ 678.2597, found 678.2609.

6.5.4. 2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->4)-deoxy-2,3,6-tri-O-acetyl-α-D-glucopyranosyl)-ethanoic acid (26a)

For procedure please refer to Section **6.4.3.**

Reaction yield: 80%.

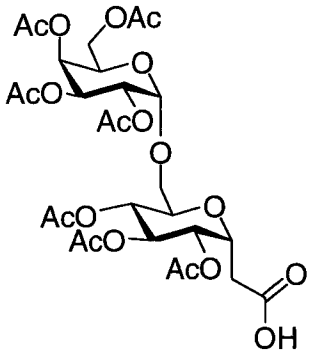


^{1}H NMR (300 MHz, $CDCl_3$) δ 5.39-5.29 (2H, m), 5.13 (1H, dd, *J*=7.9, 10.4 Hz), 5.04 (1H, dd, *J*=5.1, 7.9 Hz), 4.98 (1H, dd, *J*=3.4, 10.4 Hz), 4.64-4.55 (1H, m), 4.53 (1H, d, *J*=4.5 Hz), 4.37 (1H, dd, *J*=2.9, 11.9 Hz), 4.22-4.08 (3H, m), 3.95-3.85 (2H, m), 2.79 (1H, dd, *J*=9.0, 15.7 Hz), 2.65 (1H, dd, *J*=5.2, 15.7 Hz), 2.16 (3H, s), 2.11 (3H, s), 2.09 (3H, s), 2.08-2.05 (9H, m), 1.98 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 175.6, 171.0, 170.9, 170.6, 170.5, 170.2, 170.0, 169.6, 101.7, 77.7, 76.2, 71.7, 71.4, 71.2, 69.6, 69.4, 69.3, 69.1, 67.2, 62.3, 61.3, 33.7, 21.3, 21.2, 21.1, 21.1, 20.9; HRMS (ESI): Calcd for $C_{28}H_{42}NO_{19}$ $[M+NH_4]^+$ 696.2343, found 696.2351.

6.5.5. 2-(2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl-(1-->6)-deoxy-2,3,4-tri-O-acetyl-α-D-glucopyranosyl)-ethanoic acid (26b)

For procedure please refer to Section **6.4.3.**

Reaction yield: 82%.

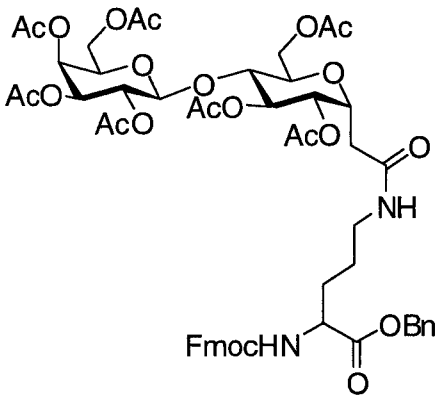


^{1}H NMR (300 MHz, $CDCl_3$) δ 5.47-5.37 (2H, m), 5.32 (1H, dd, *J*=8.9, 9.5 Hz), 5.23-5.06 (3H, m), 4.82 (1H, dd, *J*=8.8, 9.6 Hz), 4.78-4.66 (1H, m), 4.52-4.44 (1H, m), 4.17-4.00 (3H, m), 3.71 (1H, dd, *J*=8.9, 9.8 Hz), 3.45 (1H, dd, *J*=1.8, 9.9 Hz), 2.95 (1H, dd, *J*=12.4, 15.2 Hz), 2.61 (1H, dd, *J*=4.0, 15.3 Hz), 2.16 (3H, s), 2.10 (3H, s), 2.09 (3H, s), 2.08 (3H, s), 2.08 (3H, s), 2.06 (3H, s), 2.05 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.2, 171.4, 170.9, 170.7, 170.4, 170.1, 169.8, 95.2, 77.6, 70.8, 70.7, 70.1, 69.9, 69.7 ,69.0, 68.7, 68.5, 66.7, 65.5, 62.0, 32.5, 32.1, 31.9, 21.4, 21.3, 21.2, 21.1; HRMS (ESI): Calcd for $C_{28}H_{42}NO_{19}$ $[M+NH_4]^+$ 696.2316, found 696.2351.

6.5.6. Benzyl Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->4)-deoxy-2,3,6-tri-O-acetyl-α-D-glucopyranosyl)acetyl]-L-ornithine (27)

For procedure please refer to Section **6.4.5.**

Reaction yield: >90%.

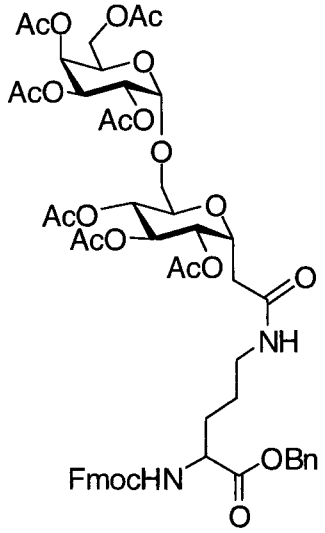


^{1}H NMR (300 MHz, $CDCl_3$) δ 7.81-7.75 (2H, m), 7.66-7.58 (2H, m), 7.45-7.30 (9H, m), 6.16-6.06 (1H, m), 5.65 (1H, d, *J*=8.1 Hz), 5.40-5.30 (2H, m), 5.25-5.10 (3H, m), 5.05-4.93 (2H, m), 4.61-4.52 (2H, m), 4.50-4.35 (3H, m), 4.27-4.18 (1H, m), 4.16-4.09 (3H, m), 3.97-3.82 (2H, m), 3.72-3.65 (1H, m), 3.35-3.20 (1H, m), 2.58 (1H, dd, *J*=9.6, 15.3 Hz), 2.39 (1H, dd, *J*=4.1, 15.3 Hz), 2.14 (3H, s), 2.11-2.03 (15H, m), 1.99 (3H, s), 1.95-1.82 (1H, m), 1.79-1.64 (1H, m), 1.63-1.49 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.5, 170.9, 170.5, 170.4, 170.1, 169.9, 169.8, 169.6, 166.1, 156.5, 144.3, 144.1, 141.7, 135.6, 129.1, 128.9, 128.7, 128.1, 127.5, 125.5, 120.4, 102.0, 77.8, 76.4, 71.9, 71.3, 71.2, 69.6, 69.5, 69.4, 69.1, 67.6, 67.4, 67.2, 62.4, 61.4, 54.1, 47.5, 39.3, 39.0, 35.6, 30.3, 30.1, 25.9, 21.3, 21.2, 21.1, 21.0, 21.0, 21.0; HRMS (ESI): Calcd for $C_{55}H_{64}N_2O_{22}$ $[M+H]^+$ 1105.3950, found 1105.4028.

6.5.7. Benzyl Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl-(1-->6)-deoxy-2,3,4-tri-O-acetyl-α-D-glucopyranosyl)acetyl]-L-ornithine (28)

For procedure please refer to Section **6.4.5.**

Reaction yield: >90%.



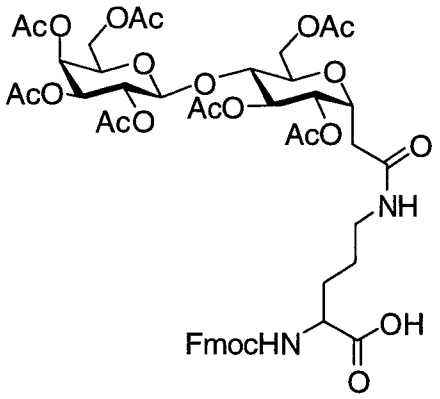
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.79-7.73 (2H, m), 7.65-7.57 (2H, m), 7.45-7.27 (9H, m), 6.28-6.19 (1H, m), 5.83 (1H, d, *J*=8.2 Hz), 5.47-5.44 (1H, m), 5.32 (1H, dd, *J*=3.3, 10.7 Hz), 5.27-5.20 (1H, m), 5.20-5.12 (2H, m), 5.10 (1H, dd, *J*=3.6, 10.7 Hz), 5.02 (1H, dd, *J*=5.0, 8.2 Hz), 4.99-4.91 (1H, m), 4.65-4.55 (1H, m), 4.47-4.31 (3H, m), 4.27-4.17 (2H, m), 4.12-4.04 (2H, m), 3.98-3.90 (1H, m), 3.76 (1H, dd, *J*=5.7, 11.0 Hz), 3.6 (1H, dd, *J*=4.0, 11.0 Hz), 3.40-3.12 (2H, m), 2.58 (1H, dd, *J*=9.0, 15.5 Hz), 2.46 (1H, dd, *J*=4.6, 15.4 Hz), 2.12 (3H, s), 2.09 (3H, s), 2.07-2.01 (12H, m), 1.97 (3H, s), 1.94-1.80 (1H, m), 1.76-1.65 (1H, m), 1.65-1.47 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 172.5, 171.0, 170.6, 170.5, 170.2, 169.9, 169.8, 169.4, 166.1, 156.6, 144.3, 144.1, 141.7, 135.7, 129.1, 128.9, 128.7, 128.1, 127.5, 125.6, 125.5, 120.4, 96.8, 77.7, 71.5, 69.8, 69.6, 69.5, 68.9, 68.5, 68.4, 67.9, 67.6, 67.5, 67.1, 66.7, 62.1, 54.1, 47.5, 39.4, 39.0, 35.1,

30.2, 26.0, 21.1, 21.1, 21.0; HRMS (ESI): Calcd for $C_{55}H_{68}N_3O_{22}$ $[M+NH_4]^+$ 1122.4154, found 1122.4294.

6.5.8. Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->4)-deoxy-2,3,6-tri-O-acetyl-α-D-glucopyranosyl)acetyl]-L-ornithine (29)

For procedure please refer to Section **6.4.6.**

Reaction yield: 87%.



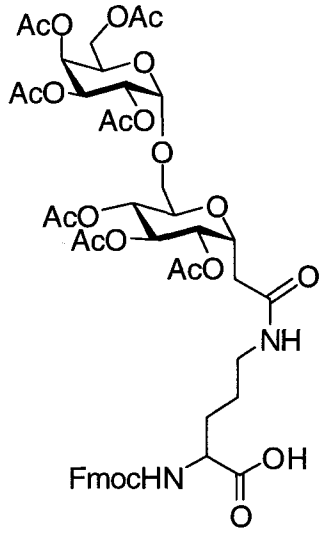
^{1}H NMR (300 MHz, $CDCl_3$) δ 7.78-7.72 (2H, m), 7.64-7.56 (2H, m), 7.42-7.35 (2H, m), 7.32-7.26 (2H, m), 6.56-6.48 (1H, m), 5.81 (1H, d, *J*=7.7 Hz), 5.37-5.30 (2H, m), 5.11 (1H, dd, *J*=7.9, 10.4 Hz), 4.98 (1H, dd, *J*=3.4, 10.4 Hz), 4.93 (1H, dd, *J*=4.5, 7.0 Hz), 4.59-4.52 (2H, m), 4.44-4.34 (3H, m), 4.24-4.17 (1H,m), 4.16-4.07 (3H, m), 3.96-3.87 (2H, m), 3.70-3.64 (1H, m), 3.39-3.25 (2H, m), 2.62 (1H, dd, *J*=8.9, 16.0 Hz), 2.42 (1H, dd, *J*=3.8, 15.2 Hz), 2.11 (3H, s), 2.07 (3H, s), 2.06 (3H, s), 2.05 (3H, s), 2.04 (3H, s), 2.03 (3H, s), 1.96 (3H, s), 1.95-1.85 (1H, m), 1.80-1.68 (1H, m), 1.67-1.52 (2H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.7, 170.5, 170.2, 170.0, 169.8, 169.6, 169.5, 156.3, 143.9, 143.7, 141.3, 127.7, 127.1, 125.1, 120.0, 101.6, 77.2, 75.8, 71.6, 70.9, 70.8, 69.1, 69.1, 68.8, 68.4, 67.0, 66.8, 61.9, 60.9, 47.1, 39.0, 35.3, 29.7, 25.3,

20.8, 20.8, 20.7, 20.6, 20.6, 20.6, 20.5; HRMS (ESI): Calcd for $C_{48}H_{62}N_3O_{22}$ $[M+NH_4]^+$ 1032.3824, found 1032.3841.

6.5.9. Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-α-D-galactopyranosyl-(1-->6)-deoxy-2,3,4-tri-O-acetyl-α-D-glucopyranosyl)acetyl]-L-ornithine (30)

For procedure please refer to Section **6.4.6.**

Reaction yield: 88%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.78-7.72 (2H, m), 7.64-7.56 (2H, m), 7.42-7.35 (2H, m), 7.30 (2H, ddd, *J*=1.0, 7.4, 7.4 Hz), 6.50-6.41 (1H, m), 5.89 (1H, d, *J*=7.7 Hz), 5.46-5.41 (1H, m), 5.32 (1H, dd, *J*=3.3, 10.8 Hz), 5.22 (1H, dd, *J*=7.8, 7.8 Hz), 5.15 (1H, d, *J*=3.6 Hz), 5.07 (1H, dd, *J*=3.7, 10.8 Hz), 5.00 (1H, dd, *J*=4.9, 8.1 Hz), 4.93 (1H, dd, *J*=7.7, 7.7 Hz), 4.64-4.55 (1H, m), 4.41-4.31 (2H, m), 4.28-4.18 (2H, m), 4.15-4.03 (2H, m), 3.99-3.91 (2H, m), 3.75 (1H, dd, *J*=5.6, 11.1 Hz), 3.59 (1H, dd, *J*=4.0, 11.1 Hz), 3.45-3.30 (1H, m), 3.30-3.15 (1H, m), 2.60 (1H, dd, *J*=9.4, 15.2 Hz), 2.49 (1H, dd, *J*=4.3, 15.2 Hz), 2.12 (3H, s), 2.09 (3H, m), 2.05 (3H, m), 2.04 (3H, s), 2.03 (6H, s), 1.98 (3H, s), 1.96-1.86 (1H, m), 1.80-1.67 (1H, m), 1.67-1.53 (2H, m); ^{13}C

NMR (75 MHz, $CDCl_3$) δ 174.3, 171.0, 170.7, 170.3, 170.3, 169.8, 169.6, 169.5, 156.3, 143.8, 143.7, 141.2, 127.7, 127.1, 125.1, 120.0, 96.3, 77.2, 76.3, 71.1, 69.4, 69.2, 69.0, 68.4, 68.2, 68.1, 67.5, 67.0, 66.5, 66.3, 61.7, 53.5, 47.1, 39.1, 34.8, 29.6, 25.3, 20.7, 20.7, 20.7, 20.7, 20.6; HRMS (ESI): Calcd for $C_{48}H_{62}N_3O_{22}$ $[M+NH_4]^+$ 1032.3824, found 1032.3842.

6.5.10. [L-Ornithine(lactose)-glycine-glycine]$_4$-glycine (31)

For procedure please refer to Section **6.4.7.**

Reaction yield: 65%.

^1H NMR (500 MHz, D_2O) δ 4.47-4.39 (2H, m), 4.38-4.28 (1H, m), 4.10-3.83 (5H, m), 3.82-3.57 (7H, m), 3.56-3.48 (1H, m), 3.28-3.15 (2H, m), 2.72-2.56 (2H, m), 1.95-1.77 (1H, m), 1.77-1.66 (1H, m), 1.66-1.45 (2H, m); ^{13}C NMR (125 MHz, D_2O) δ 102.8, 99.9, 78.5, 75.2, 72.8, 72.4, 72.0, 71.5, 70.8, 70.0, 68.4, 60.9, 59.9, 53.7, 42.2, 38.6, 32.5, 27.9; LRMS (MALDI-TOF): Calcd for $C_{94}H_{157}N_{17}O_{58}$ $[M+Na]^+$ 2476.3, found 2476.5.

6.5.11. [L-Ornithine(peracetylated melibiose)-glycine-glycine]$_4$-glycine

For procedure please refer to Section **6.4.7.**

Reaction yield: 77%.

^1H NMR (300 MHz, MeOD) δ 5.47-5.40 (1H, m), 5.38-5.25 (2H, m), 5.20-4.98 (4H, m), 4.70-4.55 (1H, m), 4.40-4.15 (2H, m), 4.18-3.73 (11H, m), 3.68-3.50 (2H, m), 3.14-2.88 (2H, m), 2.83-2.48 (2H, m), 2.18-1.95 (21H, m), 1.95-1.82 (1H, m), 1.82-1.69 (1H, m), 1.69-1.48 (2H,

m); ^{13}C NMR (75 MHz, MeOD) δ 172.2, 172.1, 171.7, 171.3, 163.2, 162.8, 120.2, 116.3, 112.4, 97.7, 72.1, 72.1, 71.6, 71.3, 71.1, 70.1, 69.5, 69.1, 67.6, 67.2, 62.9, 44.0, 40.3, 40.0, 34.9, 29.7, 27.0, 20.9, 20.7, 20.6; LRMS (MALDI-TOF): Calcd for $C_{150}H_{213}N_{17}O_{86}$ $[M+Na]^+$ 3653.36, found 3652.9.

6.5.12. [L-Ornithine(melibiose)-glycine-glycine]$_4$-glycine (32)

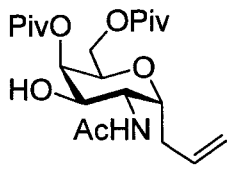
For procedure please refer to Section **6.4.7.**

Reaction yield: 80%.

^1H NMR (500 MHz, D_2O) δ 4.49-4.41 (1H, m), 4.40-4.27 (1H, m), 4.02-3.86 (6H, m), 3.82-3.66 (4H, m), 3.64-3.52 (2H, m), 3.51-3.45 (1H, m), 3.37-3.20 (2H, m), 3.07-2.85 (1H, m), 2.70-2.56 (2H, m), 1.95-1.79 (1H, m), 1.78-1.66 (1H, m), 1.66-1.45 (2H, m); ^{13}C NMR (125 MHz, D_2O) δ 173.4, 100.2, 98.4, 74.0, 73.5, 72.2, 71.2, 70.6, 70.0, 69.8, 69.4, 68.7, 66.2, 61.3, 53.9, 42.6, 39.0, 32.7, 25.1, 25.1, 62.9, 44.0, 40.3, 40.0, 34.9, 29.7, 27.0, 20.9, 20.7, 20.6; LRMS (MALDI-TOF): Calcd for $C_{94}H_{157}N_{17}O_{58}$ $[M+Na]^+$ 2476.3, found 2475.9.

6.6. Synthetic Procedures C

6.6.1. 1,2-dideoxy-2-acetamido-1-C-allyl-4,6-di-O-pivaloyl-α-D-galactopyranoside (38)[6, 7]

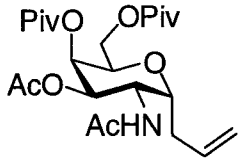


To a solution of allyl 2-acetamido-2-deoxy-6,3-dipivaloyl-α-D-glucopyranoside (3.60 g, 8.71 mmol) in 2:1 pyridine/CH_2Cl_2 (60 mL) at 0 °C, triflic anhydride (6.14 g, 21.76 mmol) was

added dwopwise. Conversion of allyl 2-acetoamido-2-deoxy-6,3-dipivaloyl-α-D-glucopyranoside to dioxolenium ion was monitored via TLC. Following complete conversion, H_2O (5 mL) was added and the reaction was allowed to warm to room temperature and it was stirred overnight. The reaction mixture was diluted with CH_2Cl_2 (60 mL), transferred to separatory funnel and the organic extract was washed with 5% aqueous HCl, water, dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography afforded (25% EtOAc/hexanes) **38** (2.92 g, 7.06 mmol) as a yellow oil in 81% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.92 (1H, d, *J*=7.7 Hz), 5.77 (1H, dddd, *J*=6.8, 6.8, 10.2, 17.0 Hz), 5.17-5.03 (3H, m), 4.70-4.55 (1H, m), 4.38 (1H, ddd, *J*=3.6, 5.2, 9.0 Hz), 4.25-4.10 (2H, m), 4.08-3.97 (2H, m), 3.20 (1H, d, *J*=4.2 Hz), 2.42-2.25 (1H, m), 2.25-2.13 (1H, m), 2.00 (3H, s), 1.24 (9H, s), 1.19 (9H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 178.2, 178.1, 170.8, 133.9, 117.5, 70.6, 70.4, 68.5, 67.8, 60.7, 52.0, 33.1, 27.1, 23.2; LRMS (ESI): Calcd for $C_{21}H_{35}NO_7$ $[M+H]^+$ 414.2, found 414.4.

6.6.2. 1,2-dideoxy-2-acetamido-1-C-allyl-3-O-acetyl-4,6-di-O-pivaloyl-α-D-galactopyranoside (39)[7]



To a solution of **38** (0.30 g, 0.73 mmol) in CH_2Cl_2 (10 mL), pyridine (1 mL), acetic anhydride (2 mL), and catalytic DMAP were added respectively. The reaction mixture was stirred overnight at ambient temperature. It was then diluted with CH_2Cl_2 (10 mL) transferred to a separatory funnel and washed successively with saturated $CuSO_4$ solution, water and brine.

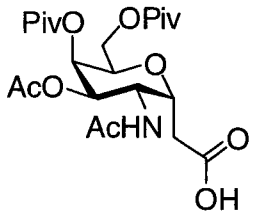
The organic extract was dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (0-25% EtOAc/hexanes) afforded product **39** (0.32 g, 0.70 mmol) in >95% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 6.02 (1H, d, *J*=8.1 Hz), 5.73 (1H, dddd, *J*=6.7, 6.7, 10.2, 17.9 Hz), 5.29 (1H, dd, *J*=2.8, 3.5 Hz), 5.20-5.00 (3H, m), 4.49-4.36 (1H, m), 4.39-4.18 (2H, m), 4.12-4.01 (1H, m), 3.96 (1H, dd, *J*=5.2, 11.2 Hz), 2.47-2.30 (1H, m), 2.30-2.15 (1H, m), 1.99 (3H, s), 1.93 (3H, s), 1.19 (9H, s), 1.14 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 178.0, 177.4, 170.6, 170.0, 133.5, 117.5, 71.3, 68.7, 68.2, 66.4, 61.1, 48.9, 31.3, 27.0, 23.2, 20.8; HRMS (ESI): Calcd for $C_{23}H_{37}NO_8$ $[M+H]^+$ 456.2597, found 456.2576.

6.6.3. 2-(1,2-dideoxy-2-acetamido-3-O-acetyl-4,6-di-O-pivaloyl-α-D-galactopyranoside)-ethanoic acid (40)[7]

For procedure please refer to Section **6.4.3.**

Reaction yield: 75%.



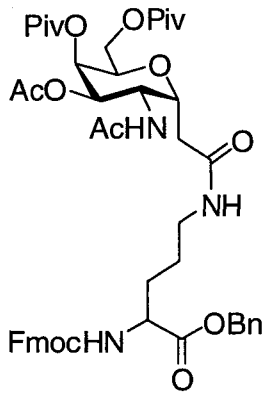
^{1}H NMR (400 MHz, $CDCl_3$) δ 8.89 (1H, br s), 6.42 (1H, dd, *J*=8.0 Hz), 5.38-5.30 (1H, m), 5.16 (1H, dd, *J*=3.0, 9.3 Hz), 4.78 (1H, dd, *J*=6.7, 11.8 Hz), 4.57-4.45 (1H, m), 4.37-4.21 (1H, m), 4.21-4.11 (1H, m), 4.02 (1H, dd, *J*=5.5, 11.0 Hz), 2.73-2.55 (2H, m), 2.03 (3H, s), 1.97 (3H, s), 1.22 (9H, s), 1.16 (9H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 178.2, 177.4, 174.0, 171.3, 170.6,

69.6, 68.1, 66.0, 60.7, 48.5, 33.7, 27.1, 27.0, 22.9, 20.8; HRMS (ESI): Calcd for $C_{22}H_{35}NO_{10}$ $[M+H]^+$ 474.2339, found 474.2326.

6.6.4. Benzyl Fluorenylmethoxycarbonyl-[2-(1,2-dideoxy-2-acetamido-3-O-acetyl-4,6-di-O-pivaloyl-α-D-galactopyranoside)acetyl]-L-ornithine (42)[7]

For procedure please refer to Section **6.4.5.**

Reaction yield: >90%.

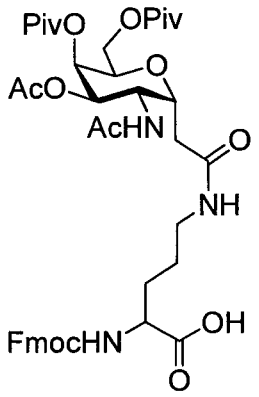


^{1}H NMR (400 MHz, $CDCl_3$) δ 7.73 (2H, d, *J*=7.5 Hz), 7.62-7.58 (2H, m), 7.41-7.22 (9H, m), 6.97 (1H, d, *J*=7.9 Hz), 6.87-6.78 (1H, m), 5.99 (1H, d, *J*=8.2 Hz), 5.38-5.32 (1H, m), 5.19-5.08 (3H, m), 4.78-4.68 (1H, m), 4.57-4.46 (1H, m), 4.45-4.12 (7H, m), 4.01 (1H, dd, *J*=6.1, 10.6 Hz), 3.35-3.10 (2H, m), 2.60 (1H, dd, *J*=9.2, 14.6 Hz), 2.46 (1H, dd, *J*=4.6, 14.8 Hz), 1.97 (3H, s), 1.93 (3H, s), 1.92-1.80 (1H, m), 1.78-1.62 (1H, m), 1.62-1.45 (2H, m), 1.22 (9H, s), 1.13 (9H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 177.9, 177.4, 172.1, 171.4, 170.8, 170.4, 170.0, 156.2, 143.8, 143.6, 141.1, 135.1, 128.5, 128.3, 128.1, 127.6, 126.9, 125.0, 119.8, 68.1, 67.1, 67.0, 65.9, 60.8, 53.8, 48.1, 47.0, 38.9, 34.6, 29.7, 29.2, 27.0, 26.9, 25.5, 22.9, 20.6; HRMS (ESI): Calcd for $C_{49}H_{61}N_3O_{13}$ $[M+H]^+$ 900.4282, found 900.4229.

6.6.5. Fluorenylmethoxycarbonyl-[2-(1,2-dideoxy-2-acetamido-3-O-acetyl-4,6-di-O-pivaloyl-α-D-galactopyranoside)acetyl]-L-ornithine (43)[7]

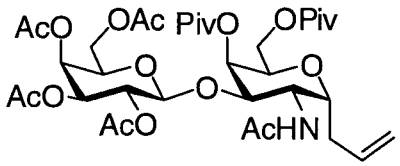
For procedure please refer to Section **6.4.6.**

Reaction yield: 83%



^{1}H NMR (400 MHz, $CDCl_3$) δ 7.72 (2H, d, *J*=7.72 Hz), 7.63-7.50 (2H, m), 7.42-7.31 (2H, m), 7.31-7.20 (2H, m), 6.13-5.97 (1H, m), 5.43-5.30 (1H, m), 5.20-5.06 (1H, m), 4.83-4.67 (1H, m), 4.55-4.10 (7H, m), 4.07-3.92 (1H, m), 3.40-3.10 (2H, m), 2.73-2.45 (2H, m), 1.96 (3H, m), 1.94 (3H, m), 2.00-1.82 (1H, m), 1.82-1.66 (1H, m), 1.66-1.45 (2H, m), 1.21 (9H, s), 1.13 (9H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 178.1, 177.4, 174.9, 171.9, 170.9, 170.6, 156.4, 143.8, 143.6, 141.2, 127.7, 127.0, 125.1, 119.9, 99.9, 68.0, 67.0, 66.0, 66.0, 61.0, 53.7, 48.2, 47.0, 39.3, 29.9, 27.1, 27.0, 22.8, 20.7; HRMS (ESI): Calcd for $C_{42}H_{56}N_3O_{13}$ $[M+H]^+$ 810.3813, found 810.3791.

6.6.6. 2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-1-C-allyl-4,6-di-O-pivaloyl-α-D-galactopyranoside (46)



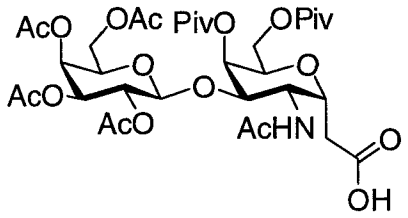
To a solution of **38** (0.20 g, 0.48 mmol) and BF_3•Et_2O (0.13 g, 0.92 mmol) in CH_2Cl_2 (15 mL) at 0 °C, hexanes were added until **38** started to precipitate out. Following a solution of trichloroacetimidate 2,3,4,6-tetra-O-acetyl-α-D-galactopyranoside[8] (0.31 g, 0.63 mmol) in CH_2Cl_2 (5 mL) was added dropwise and the reaction mixture was brought to RT.[8b, 9] After stirring for 2 hrs, solid $NaHCO_3$ was added and after 15 mins the reaction mixture was diluted with CH_2Cl_2 (15 mL) and water was added. The solution was transferred to a separatory funnel and the organic extract was washed with water, brine, dried over $MgSO_4$, and concentrated *in vacuo*. Flash column chromatography (0-15% acetone/toluene), following preparatory TLC purification (7 elutions, 3-4% MeOH/CH_2Cl_2) gave the product (0.16 g, 0.22 mmol) in 45% yield.

^{1}H NMR (500 MHz, $CDCl_3$) δ 5.79 (1H, d, *J*=8.5 Hz), 5.69 (1H, dddd, *J*=7.0, 7.0, 10.2, 17.8 Hz), 5.34 (1H, dd, *J*=0.8, 3.4 Hz), 5.21 (1H, dd, *J*=7.5, 10.5 Hz), 5.10-4.97 (4H, m), 4.84 (1H, dd, *J*=10.8, 12.8 Hz), 4.61 (1H, d, *J*=8.0 Hz), 4.15-3.98 (7H, m), 3.91-3.86 (1H, m), 2.26-2.17 (1H, m), 2.15-2.04 (7H, m), 2.01 (3H, s), 2.00 (3H, s), 1.95 (3H, s), 1.19 (9H, s), 1.17 (9H, s); ^{13}C NMR (125 MHz, $CDCl_3$) δ 177.6, 177.3, 170.2, 170.0, 169.8, 169.8, 169.3, 133.2, 117.7, 98.9, 72.7, 72.4, 70.7, 70.6, 68.8, 66.8, 65.5, 65.3, 60.8, 58.9, 49.3, 38.7, 38.5, 34.9, 29.5, 27.0, 26.9, 23.1, 20.6, 20.5, 20.4, 20.4; HRMS (ESI): Calcd for $C_{35}H_{53}NO_{16}$ $[M+H]^+$ 744.3442, found 744.3504.

6.6.7. 2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-galactopyranoside)-ethanoic acid (47)

For procedure please refer to Section **6.4.3.**

Reaction yield: 70%.

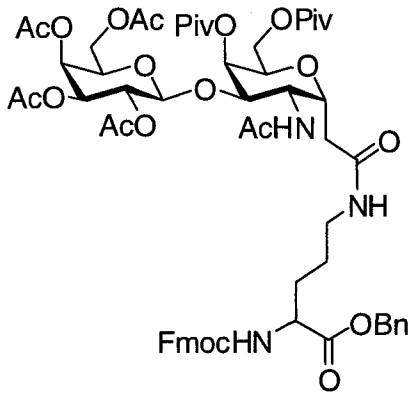


^{1}H NMR (500 MHz, $CDCl_3$) δ 6.60-6.38 (1H, m), 5.35-5.31 (1H, m), 5.21-5.13 (1H, m), 5.09-4.99 (2H, m), 4.80-4.63 (2H, m), 4.55-4.45 (1H, m), 4.24-4.21 (1H, m), 4.15-3.90 (6H, m), 2.48 (1H, dd, *J*=6.8, 16.0 Hz), 2.37 (1H, dd, *J*=6.6, 15.9 Hz), 2.09 (6H, s), 2.01 (3H, s), 1.98 (3H, s), 1.93 (3H, s), 1.18 (9H, s), 1.12 (9H, s); ^{13}C NMR (125 MHz, $CDCl_3$) δ 178.0, 177.2, 174.0, 171.1, 170.3, 170.0, 169.9, 98.6, 72.7, 71.9, 70.6, 70.4, 68.8, 66.8, 65.5, 60.8, 59.5, 49.0, 38.7, 38.5, 35.8, 27.0, 26.8, 23.0, 20.5, 20.5, 20.4; HRMS (ESI): Calcd for $C_{34}H_{51}NO_{18}$ $[M+H]^+$ 762.3148, found 762.3162.

6.6.8. Benzyl Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-galactopyranoside)acetyl]-L-ornithine (48)

For procedure please refer to Section **6.4.5.**

Reaction yield: >95%.

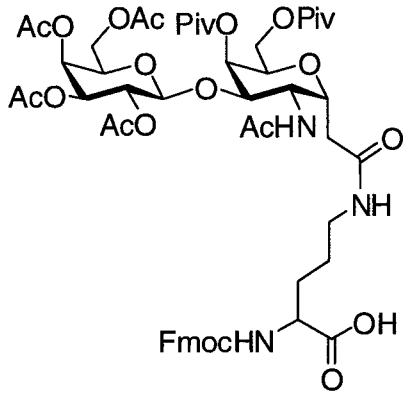


^{1}H NMR (400 MHz, $CDCl_3$) δ 7.76-7.70 (2H, m), 7.61-7.55 (2H, m), 7.40-7.22 (9H, m), 6.41 (1H, d, *J*=8.1 Hz), 6.15-6.09 (1H, m), 5.97 (1H, d, *J*=8.0 Hz), 5.34 (1H, d, *J*=3.1 Hz), 5.23-5.12 (3H, m), 5.07 (1H, dd, *J*=3.4, 10.4 Hz), 5.02 (1H, dd, *J*=3.3, 6.4 Hz), 4.70-4.60 (2H, m), 4.45.4.25 (4H, m), 4.23-4.13 (3H, m), 4.10-4.00 (3H, m), 3.93-3.82 (2H, m), 3.42-3.25 (1H, m), 3.18-3.08 (1H, m), 2.36 (1H, dd, *J*=6.5, 14.7 Hz), 2.22 (1H, dd, *J*=6.3, 14.6 Hz), 2.13 (3H, s), 2.09 (3H, s), 1.99 (3H, s), 1.93 (3H, s), 1.91 (3H, s), 1.20 (9H, s), 1.14 (9H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 178.1, 177.4, 172.4, 170.7, 170.3, 170.2, 170.0, 170.0, 169.2, 156.3, 144.0, 143.7, 141.3, 141.3, 135.3, 128.7, 128.6, 128.5, 128.2, 127.8, 127.6, 127.1, 127.0, 125.2, 125.1, 120.0, 120.0, 99.0, 73.4, 72.3, 70.8, 70.7, 68.9, 67.3, 67.0, 66.9, 65.4, 65.3, 64.2, 60.9, 59.8, 53.9, 50.1, 47.2, 38.9, 38.9, 38.7, 38.3, 29.9, 27.2, 27.0, 25.9, 23.3, 20.7, 20.6, 20.6, 20.5; HRMS (ESI): Calcd for $C_{61}H_{77}N_3O_{21}$ $[M+H]^+$ 1188.5127, found 1188.5109.

6.6.9. Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-galactopyranoside)acetyl]-L-ornithine (49)

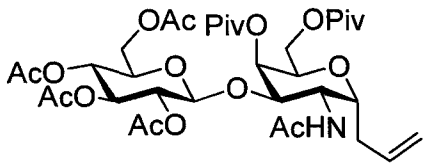
For procedure please refer to Section **6.4.6.**

Reaction yield: 83%.



^{1}H NMR (300 MHz, MeOD) δ 7.82-7.77 (2H, m), 7.70-7.65 (2H, m), 7.42-7.36 (2H, m), 7.33-7.27 (2H, m), 5.39-5.36 (1H, m), 5.33-5.28 (1H, m), 5.15-5.07 (2H, m), 4.84 (1H, d, *J*=7.1 Hz), 4.58-4.52 (1H, m), 4.40-4.32 (3H, m), 4.25-4.00 (8H, m), 3.32-3.25 (1H, m), 3.15-3.07 (1H, m), 2.53 (1H, dd, *J*=8.8, 14.2 Hz), 2.38 (1H, dd, *J*=5.2, 14.4 Hz), 2.14 (3H, s), 2.12 (3H, s), 2.02 (3H, s), 2.00 (3H, s), 1.95 (3H, s), 1.92-1.81 (1H, m), 1.75-1.63 (1H, m), 1.63-1.52 (2H, m), 1.25 (9H, s), 1.19 (9H, s); ^{13}C NMR (75 MHz, MeOD) δ 178.1, 177.9, 177.0, 172.0, 170.6, 170.5, 170.4, 170.0, 170.0, 157.0, 143.9, 143.7, 141.1, 127.3, 126.7, 124.8, 119.5, 99.3, 72.6, 70.8, 70.2, 68.7, 67.2, 66.4, 60.9, 60.6, 55.1, 48.8, 48.4, 46.9, 38.7, 38.5, 38.3, 35.3, 29.4, 26.2, 26.1, 25.4, 21.3, 19.5, 19.1, 19.0, 19.0; HRMS (ESI): Calcd for $C_{54}H_{75}N_4O_{21}$ $[M+NH_4]^+$ 1115.4923, found 1115.4872.

6.6.10. 2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-1-C-allyl-4,6-di-O-pivaloyl-α-D-galactopyranoside (50)



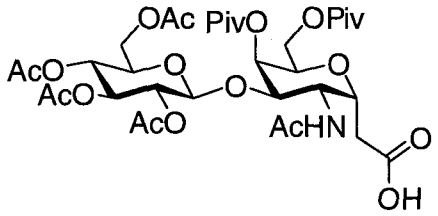
To a solution of **38** (0.15 g, 0.36 mmol), **2a** (0.21 g, 0.51 mmol) and 2,6-lutidine (0.08 g, 0.73 mmol) in CH_2Cl_2 (20 mL) with 4Å MS at 0 °C, solid AgOTf (0.14 g, 0.54 mmol) was added in 3 equal portions over 15 mins.[10] The reaction mixture was brought to RT and after stirring for 2 hrs, solid $NaHCO_3$ was added. After 15 mins the reaction mixture was diluted with CH_2Cl_2 (20 mL), transferred to a separatory funnel and the organic extract was washed with water, brine, dried over $MgSO_4$, and concentrated *in vacuo*. Flash column chromatography (0-15% acetone/toluene), following preparatory TLC purification (7 elutions, 3-4% MeOH/CH_2Cl_2) gave the product (0.20 g, 0.27 mmol) in 74% yield.

^{1}H NMR (400 MHz, $CDCl_3$) δ 5.94 (1H, d, *J*=7.4 Hz), 5.70 (1H, dddd, *J*=6.7, 6.7, 9.9, 17.7 Hz), 5.25-5.00 (6H, m), 4.95-4.82 (1H, m), 4.69 (1H, d, *J*=8.0 Hz), 4.24 (1H, dd, *J*=4.0, 12.4 Hz), 4.17-3.83 (6H, m), 3.75-3.66 (1H, m), 2.30-2.17 (1H, m), 2.14-2.04 (2H, m), 2.05 (3H, s) 2.04 (3H, s), 2.01 (3H, s), 1.99 (3H, s), 1.19 (9H, s), 1.18 (9H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 177.7, 177.5, 170.6, 170.1, 170.1, 170.0, 169.5, 133.3, 118.0, 98.5, 72.8, 72.6, 71.9, 71.1, 68.0, 65.7, 65.5, 65.5, 61.6, 58.8, 49.3, 38.9, 38.7, 35.1, 27.2, 27.0, 23.3, 20.7, 20.6, 20.6; HRMS (ESI): Calcd for $C_{35}H_{53}NO_{16}$ $[M+H]^+$ 744.3442, found 744.3396.

6.6.11. 2-(2,3,4,6-tetra-O-acetyl-β-D-galactopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-glucopyranoside)-ethanoic acid (51)

For procedure please refer to Section **6.4.3.**

Reaction yield: 65%.

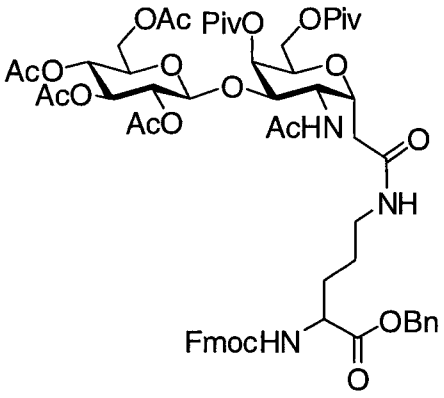


^{1}H NMR (300 MHz, $CDCl_3$) δ 6.40 (1H, d, *J*=8.4 Hz), 5.33-5.21 (1H, m), 5.17-5.01 (2H, m), 4.86-4.71 (2H, m), 4.60-4.51 (1H, m), 4.34-4.23 (2H, m), 4.13-3.97 (4H, m), 3.80-3.71 (1H, m), 2.53 (1H, dd, *J*=6.8, 15.8 Hz), 2.42 (1H, dd, *J*=6.9, 15.9 Hz), 2.14 (3H, s), 2.07 (6H, s), 2.03 (3H, s), 2.01 (3H, s), 1.21 (9H, s), 1.17 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 178.5, 177.9, 174.2 , 171.4, 171.0, 170.7, 170.4, 169.9, 98.8, 77.7, 73.4, 73.2, 72.7, 72.2, 71.5, 68.4, 65.9, 64.4, 62.0, 59.7, 53.9, 49.5, 39.3, 39.1, 36.2, 27.5, 27.4, 23.6, 21.1, 21.0, 20.9; HRMS (ESI): Calcd for $C_{34}H_{51}NO_{18}$ $[M+H]^+$ 762.3148, found 762.3158.

6.6.12. Benzyl Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-galactopyranoside)acetyl]-L-ornithine (52)

For procedure please refer to Section **6.4.5.**

Reaction yield: >95%.

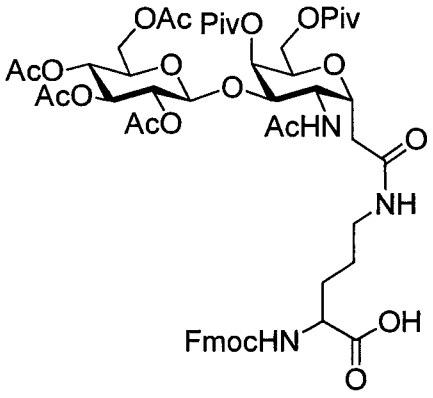


^{1}H NMR (500 MHz, $CDCl_3$) δ 7.79-7.73 (2H, m), 7.63-7.56 (2H, m), 7.44-7.26 (9H, m), 6.49-6.37 (1H, m), 6.15-6.05 (1H, m), 5.90 (1H, d, *J*=8.0 Hz), 5.30-5.00 (6H, m), 4.77-4.61 (2H, m), 4.48-4.32 (4H, m), 4.28-3.95 (8H, m), 3.80-3.71 (1H, m), 3.40-3.10 (3H, m), 2.47-2.33 (1H, m), 2.32-2.21 (1H, m), 2.14 (3H, s), 2.04 (3H, s), 2.02 (3H, s), 1.98 (3H, s), 1.96 (3H, s), 1.93-1.83 (1H, m), 1.77-1.66 (1H, m), 1.64-1.46 (2H, m), 1.21 (9H, s), 1.16 (9H, s); ^{13}C NMR (125 MHz, $CDCl_3$) δ 178.0, 177.4, 172.3 , 170.7, 170.5, 170.0, 169.9, 169.4, 169.2, 156.2, 143.9, 143.7, 141.3, 135.2, 128.6, 128.5, 128.3, 127.7, 127.0, 125.1, 125.1, 120.0, 120.0, 98.6, 73.3, 72.7, 72.4, 71.7, 71.1, 68.0, 67.3, 67.0, 65.3, 64.1, 61.6, 59.6, 50.1, 47.1, 38.9, 38.8, 38.7, 38.2, 30.0, 27.1, 27.0, 25.7, 23.3, 20.6, 20.5, 20.5; HRMS (ESI): Calcd for $C_{61}H_{77}N_3O_{21}$ $[M+H]^+$ 1188.5127, found 1188.5113.

6.6.13. Fluorenylmethoxycarbonyl-[2-(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-galactopyranoside)acetyl]-L-ornithine (53)

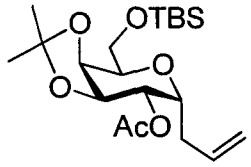
For procedure please refer to Section **6.4.6.**

Reaction yield: 88%.



^{1}H NMR (300 MHz, $CDCl_3$) δ 7.75 (2H, d, *J*=7.5 Hz), 7.66 (2H, dd, *J*=7.5, 12.9 Hz), 7.39 (2H, dd, *J*=7.4, 7.4 Hz), 7.33-7.24 (2H, m), 6.62 (1H, d, *J*=8.2 Hz), 6.48 (1H, d, *J*=8.4 Hz), 6.03-5.97 (1H, m), 5.68 (1H, dd, *J*=9.6, 9.6 Hz), 5.12 (1H, dd, *J*=9.8, 9.8 Hz), 5.06-4.96 (2H, m), 4.73-4.61 (2H, m), 4.52-4.30 (4H, m), 4.29-4.17 (3H, m), 4.17-4.03 (4H, m), 3.93-3.83 (1H, m), 3.83-3.73 (1H, m), 3.73-3.55 (1H, m), 3.06-2.91 (1H, m), 2.45 (1H, dd, *J*=4.3, 13.3 Hz), 2.27-2.17 (1H, m), 2.21 (3H, s), 2.10 (3H, s), 2.05-1.95 (1H, m), 2.02 (6H, s), 1.97 (3H, s), 1.82-1.45 (3H, m), 1.20 (9H, s), 1.18 (9H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 178.1, 177.6, 174.3, 171.7, 170.9, 170.5, 170.2, 169.4, 168.9, 156.8, 144.1, 143.7, 141.2, 127.7, 127.0, 126.9, 125.2, 125.2, 120.0, 98.6, 77.2, 73.2, 73.0, 72.8, 71.4, 71.1, 67.5, 66.9, 65.3, 64.5, 61.5, 59.2, 53.4, 49.2, 47.2, 38.8, 38.7, 38.5, 29.7, 27.1, 26.9, 23.3, 20.7, 20.6, 20.6, 20.4; HRMS (ESI): Calcd for $C_{54}H_{72}N_3O_{21}$ $[M+H]^+$ 1098.4658, found 1098.4666.

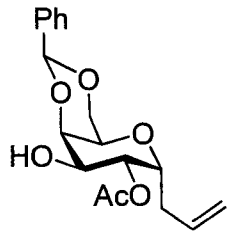
6.6.14. 2-O-acetyl-1-deoxy-C-allyl-3,4-O-isopropylidene-6-O-tert-butyldimethylsilyl-α-D-galactopyranoside (54)



To a solution of **6** (0.70 g, 1.95 mmol) in CH_2Cl_2 (10 mL), pyridine (1 mL), acetic anhydride (2 mL) and catalytic DMAP were added respectively. The reaction mixture was stirred overnight, then diluted with CH_2Cl_2 (10 mL) and washed successively with saturated $CuSO_4$ solution, water and brine. The organic extract was dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (25% EtOAc/hexanes) afforded compound **54** (0.78 g, 1.95 mmol) as in quantitative yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 5.74 (1H, dddd, *J*=7.0, 7.0, 10.5, 17.1 Hz), 5.05 (1H, dd, *J*=1.6, 7.2 Hz), 5.00 (1H, br s), 4.90 (1H, dd, *J*=2.8, 2.8 Hz), 4.36 (1H, dd, *J*=1.4, 7.2 Hz), 4.19 (1H, dd, *J*=3.0, 7.2 Hz), 4.09 (1H, ddd, *J*=2.5, 7.5, 7.5 Hz), 3.91 (1H, ddd, *J*=1.4, 5.9, 7.6 Hz), 3.80-3.66 (2H, m), 2.36-2.23 (1H, m), 2.20-2.07 (1H, m), 2.08 (3H, s), 1.46 (3H, s), 1.29 (3H, s), 0.87 (9H, s), 0.05 (6H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 169.7, 133.6, 117.4, 109.4, 72.0, 71.4, 70.3, 69.9, 69.3, 62.4, 35.3, 26.7, 25.8, 24.3, 20.9, 18.3, -5.4, -5.5; HRMS (ESI): Calcd for $C_{20}H_{37}O_6Si$ $[M+H]^+$ 401.2359, found 401.2353.

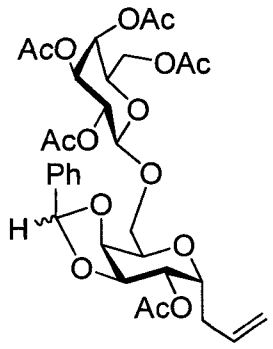
6.6.15. 2-O-acetyl-1-deoxy-C-allyl-4,6-O-benzylidene-α-D-galactopyranoside (56)



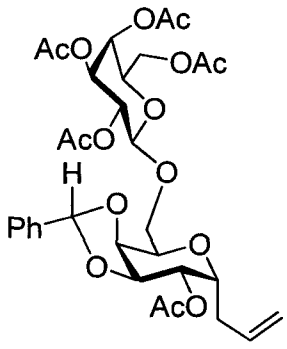
A solution of acetal **54** (0.78 g, 1.95 mmol) in 80% acetic acid (30 mL) was heated for 1 hr at 80 °C. The reaction mixture was then cooled and the solvent was co-evaporated with toluene. The residue was taken up in CH_3CN/DMF (9:1) (20 mL), followed by addition of benzaldehyde dimethyl acetal (0.44 g, 2.92 mmol) and catalytic amount of p-TsOH. After stirring for 2 hrs, the reaction mixture was neutralized with triethylamine and concentrated *in vacuo*. Flash column chromatography (25% EtOAc/hexanes) afforded **56** (0.49 g, 1.46 mmol) in overall 75% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.55-7.47 (2H, m), 7.42-7.33 (3H, m), 5.81 (1H, dddd, *J*=6.3, 7.0, 10.2, 16.5 Hz), 5.55 (1H, s), 5.32 (1H, dd, *J*=6.0, 10.1 Hz), 5.19-5.06 (2H, m), 4.38 (1H, ddd, *J*=4.6, 5.7, 10.6 Hz), 4.29 (1H, dd, *J*=1.1, 3.8 Hz), 4.24 (1H, dd, *J*=1.5, 12.5 Hz), 4.04 (1H, dd, *J*=1.8, 12.5 Hz), 3.93 (1H, dd, *J*=3.7, 10.1 Hz), 3.58 (1H, dd, *J*=1.9, 2.9 Hz), 2.60-2.28 (3H, m), 2.10 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.7, 137.3, 134.0, 129.2, 128.2, 126.3, 117.1, 101.4, 76.0, 73.2, 71.2, 69.6, 68.1, 63.0, 29.8, 21.0; HRMS (ESI): Calcd for $C_{18}H_{23}O_6$ $[M+H]^+$ 335.1494, found 335.1498.

6.6.16. 2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1-->6)-2-O-acetyl-1-deoxy-C-allyl-3,4-O-isopropylidene-α-D-galactopyranoside (57)

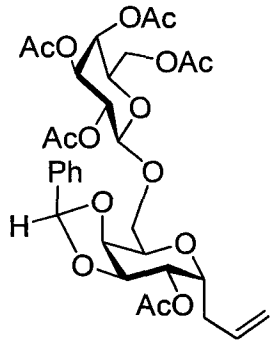


To a solution of **56** (20 mg, 0.06 mmol), **2a** (40 mg, 0.10 mmol) and 2,6-lutidine (12 mg, 0.12 mmol) in CH_2Cl_2 (5 mL) with 4Å MS at -30 °C, solid AgOTf (40 mg, 0.16 mmol) was added. The reaction mixture was brought to 5 °C over 2 hrs of stirring, then solid $NaHCO_3$ was added. After 15 mins the reaction mixture was diluted with CH_2Cl_2 (10 mL), transferred to a separatory funnel and the organic extract was washed with water, brine, dried over $MgSO_4$, and concentrated *in vacuo*. Preparatory TLC purification (3 elutions, 10% acetone/toluene) gave the two diasteriomers (15 mg, 0.02 mmol) in total 37% yield.



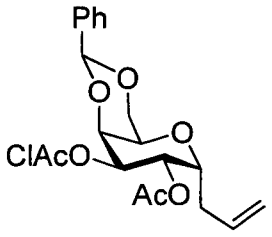
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.55-7.47 (2H, m), 7.44-7.37 (3H, m), 5.86-5.68 (1H, m), 5.78 (1H, s), 5.39 (1H, dd, *J*=1.0, 3.3 Hz), 5.23 (1H, dd, *J*=7.9, 10.4 Hz), 5.14 (1H, dd, *J*=2.8, 2.8 Hz), 5.13-5.05 (2H, m), 5.01 (1H, dd, *J*=3.4, 10.5 Hz), 4.62 (1H, d, *J*=8.0 Hz), 4.35 (1H, dd, *J*=2.8, 7.7 Hz), 4.33-4.23 (2H, m), 4.18-4.03 (4H, m), 3.89 (1H, ddd, *J*=1.0, 6.6, 6.7 Hz), 3.78 (1H, dd,

J=7.1, 10.1 Hz), 2.43-2.28 (1H, m), 2.28-2.15 (1H, m), 2.15 (3H, s), 2.12 (3H, s), 2.07 (3H, s), 2.01 (3H, s), 1.99 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.4, 170.2, 170.2, 169.7, 169.4, 138.2, 133.5, 129.2, 128.5, 126.1, 117.8, 104.5, 101.5, 73.3, 72.8, 71.1, 70.9, 70.7, 70.0, 69.8, 68.9, 68.8, 67.1, 61.2, 57.3, 34.4, 29.7, 20.9, 20.8, 20.7, 20.7, 20.6; LRMS (ESI): Calcd for $C_{32}H_{40}O_{15}$ $[M+Na]^+$ 687.7, found 687.2.



^{1}H NMR (500 MHz, $CDCl_3$) δ 7.45-7.32 (5H, m), 5.90-5.71 (1H, m), 5.39 (1H, dd, *J*=0.9, 3.5 Hz), 5.27-5.18 (2H, m), 5.17-5.07 (2H, m), 5.01 (1H, dd, *J*=3.4, 10.4 Hz), 4.63 (1H, d, *J*=7.9 Hz), 4.42 (1H, dd, *J*=4.3, 7.1 Hz), 4.30-4.20 (2H, m), 4.18-4.03 (4H, m), 3.90 (1H, ddd, *J*=1.0, 6.7, 6.7 Hz), 3.82 (1H, dd, *J*=8.6, 12.0 Hz), 2.45-2.30 (1H, m), 2.30-2.16 (1H, m), 2.15 (3H, s), 2.11 (3H, s), 2.04 (6H, s), 1.99 (3H, s); ^{13}C NMR (100 MHz, $CDCl_3$) δ 170.4, 170.2, 170.1, 169.5, 169.5, 135.8, 133.5, 129.8, 128.5, 126.9, 117.6, 104.4, 101.7, 73.7, 72.6, 70.9, 70.7, 70.2, 69.8, 69.4, 69.2, 68.9, 67.1, 61.2, 35.5, 29.7, 20.9, 20.8, 20.7, 20.6, 20.6; LRMS (ESI): Calcd for $C_{32}H_{40}O_{15}$ $[M+Na]^+$ 687.7, found 687.3.

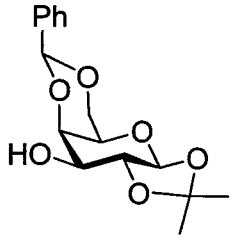
6.6.17. 2-O-acetyl-1-deoxy-C-allyl-4,6-O-benzylidene-3-O-chloroacetyl-α-D-galactopyranoside (58)



To a solution of **56** (20 mg, 0.06 mmol) in 1:9 pyridine/CH_2Cl_2 (5 mL) at 0 °C, chloroacetic anhydride (15 mg, 0.09 mmol) was added and the reaction mixture was stirred for 2 hrs at this temperature.[11] The reaction mixture was diluted with CH_2Cl_2 (10 mL), transferred to separatory funnel and washed with water, $CuSO_4$, brine, and dried over $MgSO_4$. The organic residue was coevaporated with toluene to give pure **58** (21 mg, 0.05 mmol) in 87% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.55-7.43 (2H, m), 7.42-7.33 (3H, m), 5.80 (1H, dddd, *J*=6.2, 7.0, 10.2, 16.3 Hz), 5.57 (1H, dd, *J*=5.9, 10.4 Hz), 5.50 (1H, s), 5.23 (1H, dd, *J*=3.6, 10.4 Hz), 5.19-5.07 (2H, m), 4.59-4.46 (2H, m), 4.24 (1H, dd, *J*=1.5, 12.5 Hz), 4.09 (2H, s), 4.03 (1H, dd, *J*=1.7, 12.5 Hz), 3.62 (1H, dd, *J*=1.5, 2.8 Hz), 2.59-2.42 (1H, m), 2.42-2.25 (1H, m), 2.05 (3H, s); LRMS (ESI): Calcd for $C_{20}H_{24}O_7Cl$ $[M+H]^+$ 411.9, found 411.1.

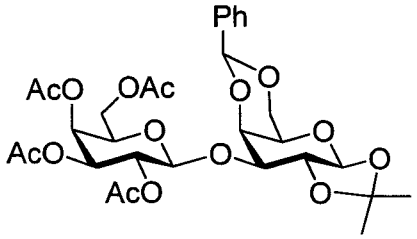
6.6.18. 4,6-O-benzylidene-1,2-O-isopropylidene-β-D-galactopyranoside (60)



To a solution of D-galactose (5.0 g, 27.75 mmol) in DMF (100 mL), benzaldehyde dimethyl acetal (4.2 g, 27.75) and catalytic p-TsOH (0.1 g) were added.[12] The reaction mixture was stirred vigorously overnight and then concentrated under reduced pressure. The intermediate was crystallized from CH_2Cl_2, filtered and dried *in vacuo* in the presence of P_2O_5. 4,6-O-benzylidene-D-galactopyranoside (0.16 g, 0.60 mmol) was then suspended in anhydrous acetone (30 mL) and freshly distilled dimethoxy propane (0.08 g, 0.78 mmol) and catalytic p-TsOH were (0.01 g) added.[13] The reaction mixture was stirred overnight at room temperature, then neutralized with saturated $NaHCO_3$ solution and concentrated under reduced pressure. The residue was taken up in CH_2Cl_2 (10 mL), washed with water, dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (5-20% EtOAc/hexanes) afforded **60** (0.14 g, 0.44 mmol) as a colourless syrup in 74% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 7.49-7.44 (2H, m), 7.40-7.33 (3H, m), 5.85 (1H, d, *J*=4.3 Hz), 5.54 (1H, s), 4.35 (1H, dd, *J*=1.4, 12.7 Hz), 4.28 (1H, dd, *J*=2.4, 4.0 Hz), 4.15 (1H, dd, *J*=4.3, 5.5 Hz), 4.08 (1H, dd, *J*=2.2, 12.7 Hz), 4.03-3.98 (1H, m), 3.82 (1H, dd, *J*=2.1, 3.6 Hz), 2.88 (1H, d, *J*=7.5 Hz), 1.54 (3H, s), 1.39 (3H, s); ^{13}C NMR (75 MHz, $CDCl_3$) δ 137.3, 129.2, 128.3, 126.1, 109.4, 100.8, 77.6, 73.6, 70.6, 70.5, 64.4, 27.8, 26.5; HRMS (ESI): Calcd for $C_{16}H_{24}NO_6$ $[M+NH_4]^+$ 326.1603, found 326.1600.

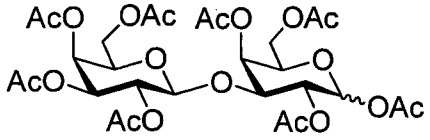
6.6.19. 2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1-->3)-4,6-O-benzylidene-1,2-O-isopropylidene-β-D-galactopyranoside (61a)



To a solution of **60** (50 mg, 0.17 mmol), AgOTf (62 mg, 0.24 mmol) and 2,6-lutidine (21 mg, 0.29 mmol) in CH_2Cl_2 (10 mL) with 4Å MS, solid **2a** (93 mg, 0.23 mmol) was added in 3 equal portions over 30 mins.[10] After stirring for 1.5 hr, solid $NaHCO_3$ was added. After 15 mins the reaction mixture was diluted with CH_2Cl_2 (10 mL), transferred to a separatory funnel and the organic extract was washed with water, brine, dried over $MgSO_4$, and concentrated *in vacuo*. Flash column chromatography (25% ethyl acetate/hexanes) afforded a mixture of α/β product (76 mg, 0.12 mmol) in 70% yield (only 30 mg of pure β product was isolated and characterized).

^1H NMR (500 MHz, $CDCl_3$) δ 7.54-7.50 (2H, m), 7.42-7.37 (3H, m), 5.74 (1H, s), 5.61 (1H, d, *J*=5.0 Hz), 5.38 (1H, dd, *J*=1.0, 3.4 Hz), 5.22 (1H, dd, *J*=8.0, 10.5 Hz), 5.02 (1H, dd, *J*=3.5, 10.2 Hz), 4.64 (1H, dd, *J*=2.4, 8.3 Hz), 4.58 (1H, d, *J*=8.0 Hz), 4.42 (1H, dd, *J*=2.4, 5.0 Hz), 4.27 (1H, dd, *J*=1.8, 8.3 Hz), 4.15-4.08 (3H, m), 4.07-4.03 (1H, m), 3.88 (1H, ddd, *J*=1.0, 6.7, 6.7 Hz), 3.74 (1H, dd, *J*=7.4, 11.3 Hz), 2.13 (3H, s), 2.09 (3H, s), 1.99 (3H, s), 1.98 (3H, s), 1.55 (3H, s), 1.34 (3H, s); HRMS (ESI): Calcd for $C_{30}H_{42}NO_{15}$ $[M+NH_4]^+$ 656.2527, found 656.2536.

6.6.20. 2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1-->3)-1,2,4,6-tetra-O-acetyl-D-glucopyranosyl (62)



A solution of acetal **61a** (30 mg, 0.05 mmol) in 80% acetic acid (10 mL) was heated overnight at 85 °C.[4] The reaction mixture was then cooled and the solvent was evaporated. The residue was taken up in CH_2Cl_2 (5 mL), followed by addition of pyridine (0.1 mL), acetic anhydride (0.2 mL) and catalytic DMAP. After overnight stirring the reaction mixture was diluted with CH_2Cl_2 and washed successively with saturated $CuSO_4$ solution, water and brine. The organic extract was dried over $MgSO_4$ and concentrated *in vacuo*. Flash column chromatography (25% EtOAc/hexanes) afforded compound **62** (30 mg, 0.04 mmol) in overall 88% yield.

^{1}H NMR (300 MHz, $CDCl_3$) δ 6.36 (1H, br s), 5.68 (1H, d, *J*=8.3 Hz), 5.48 (1H, br s), 5.41 (1H, d, *J*=3.1 Hz), 5.39-5.27 (5H, m), 5.20-5.11 (2H, m), 5.06 (1H, dd, *J*=3.2, 10.4 Hz), 5.02-4.94 (2H, m), 4.50 (1H, d, *J*=7.9 Hz), 4.46 (1H, d, *J*=7.9 Hz),4.27 (1H, dd, *J*=6.1 Hz), 4.20-4.06 (4H, m), 3.99 (1H, dd, *J*=6.2, 6.2 Hz), 3.88 (1H, dd, *J*=6.2, 11.9 Hz), 3.84-3.75 (3H, m), 3.63 (1H, dd, *J*=7.0, 10.6 Hz), 2.18 (3H, s), 2.17-2.13 (12H, m), 2.12 (3H, s), 2.07-2.03 (15H, m), 2.02 (3H, s), 2.00-1.95 (12H, m); ^{13}C NMR (75 MHz, $CDCl_3$) δ 170.4, 170.3, 170.2, 170.2, 170.1, 170.1, 170.0, 169.9, 169.9, 169.5, 169.4, 169.1, 168.9, 101.0, 100.4, 92.2, 89.6, 72.9, 71.0, 70.9, 70.8, 70.8, 70.7, 70.0, 68.4, 68.3, 67.9, 67.9, 67.5, 67.1, 67.1, 67.0, 66.9, 66.5, 65.6, 61.3, 61.1, 20.9, 20.8, 20.7, 20.7, 20.7, 20.7, 20.6, 20.6, 20.6; HRMS (ESI): Calcd for $C_{28}H_{42}NO_{19}$ $[M+NH_4]^+$ 696.2351, found 696.2341.

6.6.21. [L-Ornithine(1,2-dideoxy-2-acetamido-3-O-acetyl-4,6-di-O-pivaloyl-α-D-galactopyranoside)-glycine-glycine]$_4$-glycine

For procedure please refer to Section **6.4.7.**

Reaction yield: 78%.

^{1}H NMR (300 MHz, MeOD) δ 5.40 (1H, br s), 5.08 (1H, d, *J*=10.2 Hz), 4.69-4.56 (1H, m), 4.56-4.43 (1H, m), 4.35-4.15 (2H, m), 4.14-3.70 (7H, m), 2.78-2.57 (1H, m), 2.48-2.28 (1H, m), 2.10-1.40 (10H, m), 1.20 (9H, s), 1.11 (9H, s); ^{13}C NMR (75 MHz, MeOD) δ 180.1, 180.0, 179.7, 179.7, 174.4, 173.4, 173.4, 172.6, 164.6, 164.5, 164.0, 163.6, 163.6, 121.0, 117.1, 74.0, 73.9, 70.3, 70.1, 68.5, 62.9, 62.6, 44.7, 41.0, 40.8, 40.6, 35.4, 30.5, 28.4, 27.8, 23.5, 21.6; LRMS (MALDI-TOF): Calcd for $C_{126}H_{201}N_{21}O_{50}$ $[M+Na]^+$ 2833.0, found 2832.7.

6.6.22. [L-Ornithine(galactose)-alanine-alanine]$_4$-glycine (63)

For procedure please refer to Section **6.4.7.**

Reaction yield: 75%.

^{1}H NMR (500 MHz, D_2O) δ 4.51-4.42 (1H, m), 4.37-4.18 (3H, m), 4.00 (1H, dd, *J*=6.1, 9.8 Hz), 3.98-3.94 (1H, m), 3.83-3.77 (1H, m), 3.75-3.69 (2H, m), 3.69-3.63 (2H, m), 3.28-3.13 (2H, m), 2.65 (1H, dd, *J*=10.7, 14.9 Hz), 2.57 (1H, dd, *J*=3.8, 15.0 Hz), 1.90-1.75 (1H, m), 1.75-1.64 (1H, m), 1.64-1.47 (3H, m), 1.42-1.27 (7H, m); ^{13}C NMR (125 MHz, D_2O) δ 176.1, 174.3, 174.2, 173.4, 173.3, 117.3, 115.0, 99.9, 72.8, 72.4, 69.5, 68.6, 67.4, 60.7, 53.3, 49.4, 48.7, 43.1, 38.6, 32.3, 28.2, 24.6, 16.4; LRMS (MALDI-TOF): Calcd for $C_{78}H_{133}N_{17}O_{38}$ $[M+Na]^+$ 1940.0, found 1939.1.

6.6.23. [L-Ornithine(N-acetyl-galactosamine)-glycine-glycine]$_4$-glycine (64)

For procedure please refer to Section **6.4.7.**

Reaction yield: 87%.

^{1}H NMR (500 MHz, D_2O) δ 4.59-4.51 (1H, m), 4.39-4.27 (2H, m), 4.21 (1H, dd, *J*=5.9, 10.8 Hz), 4.10-3.90 (5H, m), 3.84 (1H, dd, *J*=2.8, 11.0 Hz), 3.80 (1H, dd, *J*=5.5, 6.4 Hz), 3.73-3.63 (2H, m), 3.28-3.12 (2H, m), 2.66 (1H, dd, *J*=10.8, 14.5 Hz), 2.44 (1H, dd, *J*=4.0, 14.8 Hz), 1.99 (3H, s), 1.93-1.78 (1H, m), 1.78-1.65 (1H, m), 1.65-1.46 (2H, m); ^{13}C NMR (125 MHz, D_2O) δ 174.5, 172.9, 172.7, 171.8, 171.5, 170.2, 162.7, 99.9, 72.4, 70.9, 68.0, 67.0, 60.8, 53.6, 52.7, 49.2, 46.5, 42.4, 42.2, 38.6, 38.4, 33.0, 28.0, 24.7, 21.8, 8.1; LRMS (MALDI-TOF): Calcd for $C_{78}H_{129}N_{21}O_{38}$ $[M+Na]^+$ 1992.0, found 1991.1.

6.6.24. [L-Ornithine(β-D-Gal-(1-->3)-D-GalNAc)-glycine-glycine]$_4$-glycine (65)

For procedure please refer to Section **6.4.7.**

Reaction yield: 72%.

^{1}H NMR (500 MHz, D_2O) δ 4.58-4.47 (2H, m), 4.40-4.27 (2H, m), 4.24-4.19 (1H, m), 4.04-3.82 (13H, m), 3.81-3.59 (6H, m), 3.56-3.48 (1H, m), 3.25-3.14 (2H, m), 2.67 (1H, dd, *J*=11.1, 14.0 Hz), 2.45 (1H, dd, *J*=3.3, 15.0 Hz), 2.06-2.01 (2H, m), 1.99 (2H, br s), 1.93-1.78 (1H, m), 1.78-1.65 (1H, m), 1.65-1.45 (2H, m); ^{13}C NMR (125 MHz, D_2O) δ 175.0, 174.8, 174.8, 174.4, 173.9, 172.5, 171.8, 171.6, 171.5, 103.8, 99.9, 76.3, 74.9, 72.7, 72.4, 70.5, 68.4, 67.5, 60.9, 60.2, 53.6,

53.2, 48.1, 42.5, 42.3, 42.2, 41.5, 38.7, 33.6, 29.5, 27.9, 24.6, 21.8, 21.6, 21.5, 17.6, 16.1; LRMS (MALDI-TOF): Calcd for $C_{102}H_{169}N_{21}O_{58}$ $[M+Na]^+$ 2640.5, found 2640.2.

6.6.25. [L-Ornithine(2,3,4,6-tetra-O-acetyl-β-D-glucopyranosyl-(1-->3)-1,2-dideoxy-2-acetamido-4,6-di-O-pivaloyl-α-D-galactopyranoside)-glycine-glycine]$_4$-glycine

For procedure please refer to Section **2.4.7.**

Reaction yield: 77%.

^{1}H NMR (300 MHz, MeOD) δ 5.37-5.22 (2H, m), 5.07-4.99 (1H, m), 4.56-4.47 (1H, m), 4.47-4.21 (4H, m), 4.21-4.01 (5H, m), 4.01-3.75 (7H, m), 3.17-3.03 (1H, m), 2.60-2.45 (1H, m), 2.43-2.31 (1H, m), 2.10 (3H, br s), 2.06 (3H, s), 2.03 (3H, s), 2.01 (3H, s), 1.97 (3H, s), 1.94-1.80 (1H, m), 1.80-1.66 (1H, s), 1.66-1.45 (2H, s), 1.23 (9H, s), 1.18 (9H, s); ^{13}C NMR (75 MHz, MeOD) δ 179.5, 178.6, 173.4, 172.4, 172.3, 172.2, 171.6, 171.3, 163.2, 162.9, 119.4, 117.1, 100.7, 100.6, 100.6, 79.3, 74.3, 74.3, 74.2, 72.9, 72.6, 69.7, 68.8, 63.0, 61.9, 61.9, 55.1, 50.2, 44.0, 44.0, 40.0, 39.8; LRMS (MALDI-TOF): Calcd for $C_{174}H_{265}N_{21}O_{82}$ $[M+Na]^+$ 3986.1, found 3985.6.

6.6.26. [L-Ornithine(β-D-Glc-(1-->3)-D-GalNAc)-glycine-glycine]4-glycine (66)

For procedure please refer to Section **2.4.7.**

Reaction yield: 91%.

^{1}H NMR (500 MHz, D_2O) δ 4.60-4.51 (2H, m), 4.42-4.28 (2H, m), 4.26-4.20 (1H, m), 4.05-3.83 (7H, m), 3.82-3.62 (3H, m), 3.52-3.36 (3H, m), 3.35-3.27 (1H, m), 3.25-3.15 (2H, m), 2.75-2.62 (1H, m), 2.52-2.41 (1H, m), 1.95-1.80 (1H, m), 1.80-1.66 (1H, m), 1.66-1.46 (2H, m); ^{13}C NMR (125 MHz, D_2O) δ 174.4, 172.5, 103.3, 99.9, 76.7, 76.6, 75.6, 75.4, 75.4, 72.8, 72.5, 69.4, 69.4, 69.3, 67.4, 60.4, 53.6, 48.1, 48.0, 42.4, 42.2, 38.7, 33.5, 28.0, 27.9, 24.6, 21.8; LRMS (MALDI-TOF): Calcd for $C_{102}H_{169}N_{21}O_{58}$ $[M+Na]^+$ 2640.5, found 2640.2.

6.7. Assessing Antifreeze Activity

6.7.1. Recrystallization Inhibition (RI)

Recrystallization-inhibition (RI) activity was assessed using the splat cooling technique previously described by Knight and co-workers.[14] Briefly, the latter technique involves generating a frozen wafer from the analyte which is dissolved in phosphate buffered saline (PBS) solution; a 10 uL drop of this solution is then dropped from a pipette through a two metre high plastic tube (10 cm in diameter) onto a block of polished aluminum pre-cooled to approximately -80 °C. The droplet freezes instantly on the polished aluminum block and is approximately 1 cm in diameter and 20 μm thick. This wafer is then carefully removed from the surface of the block and transferred to a cryostage (refrigerated microscope stage) held at -6.4 °C. The wafer is left to anneal at this temperature for a period of 30 mins and then photographed between crossed polarizing filters using a digital camera (Nikon CoolPix 5000) fitted to the microscrope. During flash freezing, ice crystals spontaneously nucleated from the super-cooled solution. These initial crystals were relatively homogenous in size and quite small. During the annealing cycle recrystallization occurred, resulting in a dramatic increase in ice crystal size. A quantitative

measure of the difference in recrystallization inhibition of two compounds X and Y is the difference in the dynamics of the ice crystal size distribution.

RI analysis via mean elliptical method:

A total of six different images are taken from each wafer; three from the middle and three from the edge. The images are then analyzed using the mean elliptical method. In this method, the ten largest ice crystals were chosen from the field of view (FOV) in each image. Selection of these crystals was arbitrary in that they were chosen after a visual inspection of the image.[15] The two dimensional surface area of each of these ten crystals was then calculated via approximation of the crystal as an elliptical area. The major and minor elliptical axes were defined by the two largest orthogonal dimensions across the ice grain surface. The surface area of each ice grain was then calculated based on the formula: $A = \pi ab$, in which A represented area; a and b represented the length of the major and minor elliptical axes. Totalling all individual measurements for each FOV produces a value for the average grain surface area referred to as the mean largest grain size (MLGS). Error was calculated using standard error of the mean (SEM). T-tests were performed to a 95% confidence level.

Analysis via domain recognition software program:

For each sample including PBS, three drops were performed and three images were taken from the middle of each wafer. Image analysis of the ice wafers was performed using a novel domain recognition software (DRS) program that was developed at the Steacie Institute for Molecular Sciences (SIMS) of the National Research Council of Canada (NRCC).[16] The software was written in C using Microsoft Visual Studio 6.0 on a Pentium class personal computer running

Microsoft Windows 2000 or XP. This processing employed the Microsoft Windows Graphical User Interface to allow a user to visually demarcate and store the vertices of ice crystal domains in a digital micrograph. These data were then used to calculate the domain areas. To eliminate the need to fully process each micrograph, an algorithm was developed to randomly display a number of x/y locations. The algorithm made use of a built-in pseudo random number generator [rand(x)] and was written so that no two locations were closer than 1/10th the field of view of the micrograph. The formula for the area of a polygon that is not self-intersecting and contains no holes is then given by:

$$A = \frac{1}{2}\sum_{i=0}^{N-1}\left(x_i y_{i+1} - x_{i+1} y_i\right)$$

where N is the number of vertices, and x_0, y_0 to x_{N-1}, y_{N-1} are the vertices circumventing the polygon in a clockwise direction. The point x_0, y_0 is assumed to be equivalent to the point x_N, y_N. In the FOV of each image, twelve locations (ice crystal domains) were randomly chosen and analyzed. All data was compiled, further analyzed and plotted using Microsoft Excel.

6.7.2. Thermal Hysterisis (TH)

Nanoliter osmometry was performed using a nanoliter osmometer (Clifton Technical Physics, Hartford, NY) as described by Chakrabartty and Hew.[17] All measurements were made in doubly distilled water. Ice crystal morphology was observed through a Leitz compound microscope equipped with an Olympus 20X (infinity corrected) objective, Leitz Periplan 32X

photo eyepiece and a Hitachi KP-M2U CCD camera connected to a Toshiba MV13K1 TV/VCR system. Still images were captured directly using a Nikon CoolPix digital camera.

6.8. References

1. (a) Murphy, A. V., Ph.D. Thesis, SUNY Binghamton, Binghamton **2004**. (b) Eniade, A., Murphy, A. V., Landreau, G., Ben, R. N., *Bioconj. Chem.* **2001**, 12, 817.
2. Roe, B. A., Boojamra, C. G., Griggs, J. L., Bertozzi, C., *J. Org. Chem.* **1996**, 61, 6442.
3. (a) Carlsen, H. J., Katsuki, T., Martin, V. S., Sharpless, B. K., *J. Org. Chem.* **1981**, 46, 3936. (b) Plietker, B., *Synthesis* **2005**, 15, 2453.
4. Bundle, D. R., Eichler, E., Lowary, T. L., *Can. J. Chem.* **2002**, 80, 1112.
5. Fields, G. B., Fields, C. G., *J. Am. Chem. Soc.* **1991**, 113, 4202.
6. Cipolla, L., La Ferla, B., Lay, L., Peri, F., Nicotra, F., *Tetrahedron: Asymmetry* **2000**, 11, 295.
7. (a) Bouvet, V., *Ph.D. Thesis*, University of Ottawa, Ottawa, **2006**. (b) Bouvet, V. R., Ben, R. N., *J. Org. Chem.* **2006**, 71, 3619.
8. (a) de la Funente, J. M., Penades, S., *Tet. Asym.* **2002**, 13, 1879. (b) Liao, W., Locke R. D., Matta, K. L., *Chem. Commun.* **2000**, 369.
9. (a) Kinzy, W., Schmidt, R. R., *Carbohydr. Res.* **1987**, 164, 265. (b) Schmidt, R. R., Toepfer, A., *Tet. Lett.* **1991**, 32, 3353.
10. (a) Koenigs, W., Knorr, E., *Ber. Dtsch. Chem. Ges.* **1901**, 34, 957. (b) Garegg, P. J., Norberg, T., *Acta. Chem. Scand. B* **1979**, 33, 116. (c) Banoub, J., Bundle, D. R., *Can. J. Chem.* **1979**, 57, 2091.
11. Belot, F., Jacquinet, J.-C., *Carbohydr. Res.* **2000**, 325, 93.
12. (a) Lonn, H., *Carbohydr. Res.* **1985**, 139, 105. (b) Duclos Jr., *Chem. Phys. Lipids* **2001**, 111, 111.
13. Busse, H., Hakoda, M., Stanley, M., Streicher, H., *J. Carb. Chem.* **2007**, 26, 159.
14. Knight, C. A., Hallett, J., DeVries, A.L., *Cryobiology* **1988**, 25, 55.

15. (a) Eniade, A., Purushotham, M., Ben, R.N., Wang, J.B., Horwath, K., *Cell Biochem. Biophys.* **2003**, 38, 115. (b) Liu, S., Ben, R.N., *Org. Lett.* **2005**, 7, 2385.

16. Jackmann, J., Noestheden, M., Moffat, D., Pezacki, J., Findlay, S., Ben, R. N., *Biochem. Biophys. Res. Commun.* **2007**, 354, 340.

17. Chakrabartty, A., Hew, C.L., *Eur. J. Biochem.* **1991**, 202, 1057.

Claims to Original Research

The successful syntheses of novel C-linked AFGP analogues **22** $[O(Tal)GG]_4G$ – Chapter 3, **31** $[O(Lac)GG]_4G$ and **32** $[O(Meli)GG]_4G$ – Chapter 4, **63** $[O(Gal)AA]_4G$, **64** $[O(GalNAc)GG]_4G$, **65** $[O(Glc\text{-}GalNAc)GG]_4G$ and **66** $[O(Gal\text{-}GalNAc)GG]_4G$ – Chapter 5, were presented. Analogues **31**, **32**, **65** and **66** were the first analogues that contained a disaccharide moiety. None of the analogues demonstrated any detectable thermal hysteresis (TH) or dynamic ice shaping (DIS) however, analogues **32**, **63** and **66** possessed potent ice recrystallization inhibition (RI) activity.

The recrystallization inhibition activities of analogues **19** $[O(Gal)GG]_4G$, **20** $[O(Glc)GG]_4G$, **21** $[O(Man)GG]_4G$ and **22** $[O(Tal)GG]_4G$ were evaluated and correlated to the partial molar compressibility (PMC) values of their carbohydrate residues. Analogue **19** showed the most potent RI activity following **20**, **21** and **22** respectively. This work demonstrated that partial molar compressibility of the carbohydrates (which describes the fit of a carbohydrate into the three-dimensional hydrogen-bonded network of water) is inversely proportional to RI activity. Collectively, the relative stereochemistry of the carbohydrate moiety plays an important role in determining the potency of recrystallization inhibition and is optimal when it matches that of galactose. This work was described in the following publication: Czechura, P., Tam, R. Y., Dimitrijevic, E., Murphy, A. V., Ben, R. N., *J. Am. Chem. Soc.* **2008**, 130, 2928.

The RI activities of analogues **31** $[O(Lac)GG]_4G$ and **32** $[O(Meli)GG]_4G$ were also assessed, and compared to the PMC values of their corresponding carbohydrate units. Analogue **32** showed superior RI activity to that of the most potent monosaccharide analogue **19**. Yet, the RI activity of analogue **31** was equal to that of **19** (within experimental error). This was

unexpected as the PMC values of lactose and melibiose were nearly doubled in comparison to D-galactose. Nonetheless this indicated that not only does the stereochemistry of the individual carbohydrates influence RI activity but also the concentration of carbohydrate units and the regiochemistry of the disaccharides. Since the PMC values did not correlate well with RI data, the hydration index (HI) values of carbohydrates (which predict RI more accurately then PMC or hydration numbers) were used in attempt to explain the observed results. This work was described in the following publication: Tam, R. Y., Ferreira, S. S., Czechura, P., Chaytor, J. L., Ben, R. N., *J. Am. Chem. Soc.* **2008**, 130, 17494. Unfortunately, hydration indices of the carbohydrates also failed to correlate with trends in RI activity of the glycopeptides. As a result, the first method for estimating HI of glycopeptides was developed. The predicted HI values of the AFGP analogues **19-22**, **31** and **32** were inversely proportional to their RI activities. This demonstrated that the model for estimating the hydration indices of glycopeptides is even more accurate in predicting RI then HI of the carbohydrates alone. Consequently, it is feasible that this model could be utilized as a tool to engineer predetermine RI of future AFGP analogues.

Other structural features of C-linked AFGP analogues were studied with analogues **19**, **63**, **64**, **65**, **66** and previously synthesized $[O(GalNAc)AA]_4G$. Firstly, it was noted that the C-2 N-acetyl groups of the carbohydrate moiety and the L-alanines both have a positive effect on RI activity and should therefore be incorporated into the future design of the AFGP analogues. Secondly, it was deemed that the labour intensive synthesis of complex disaccharides is not necessary to obtain potent recrystallization inhibitors. Thirdly, it has been identified that stereochemistry of the terminal carbohydrate is important for effective RI activity and is the best when it matches that of galactose. Finally, our model for estimating hydration indices predicted correct trends for RI activity of **19**, **63**, **64**, **65**, **66** and $[O(GalNAc)AA]_4G$ but its accuracy was

limited. This indicates that our method for estimating RI activity needs further attention and has to be refined or substituted with an advanced computational method.

Collectively, the abovementioned data suggests that optimal structure of the C-linked AFGP analogue should include: L-alanines as a choice of the amino acids, L-ornithine as the amino acid linking the peptide to the carbohydrate, and carbohydrate(s) with C-2 N-acetyl moiety and stereochemistry of its terminal sugar matching that of galactose.

The solution conformation of analogues **22**, **31**, **32**, **63**, **64**, **65** and **66** was evaluated using Circular Dichroism. Analogues **22**, **31**, **32**, and **63** exhibit secondary structures similar to that of the native AFGP8; a large percentage of random coil structure with varying degrees of other more ordered secondary structures. While analogues **64**, **65** and **66** also show a large percentage of random coil, the deconvolution results also imply that they possess a significant percent of more ordered structure. The more defined conformation of these analogues is clearly due to the presence of the C-2 N-acetyl moiety of the carbohydrate.

FUTURE WORK

The structure-function studies described in this dissertation verify that hydration is related to ice RI activity of the C-linked AFGP analogues. Also, we identified several key structural features that affect hydration of the glycoproteins and their RI activity. Thus, to improve the design and consequently the activity of our AFGP analogues, following future work should be realized:

1. To determine whether RI activity of analogue $[O(Meli)GG]_4G$ can be improved following analogues should be synthesized: C-linked $[O(Meli)AA]_4G$, $[S(Meli)GG]_4G$ and $[S(Meli)AA]_4G$.

2. To complete investigating the importance of the C-2 N-acetyl group of the carbohydrate in the C-linked AFGP analogues, synthesis of $[O(GalGal)GG]_4G$ should be completed and its antifreeze activity should be compared to $[O(GalGalNAc)GG]_4G$. Moreover, the methyl group of the C-2 N-acetyl moiety of $[O(GalNAc)GG]_4G$ and $[O(GalNAc)AA]_4G$ should be modified to include: isopropyl group (to determine whether by increasing the number of methyls in the analogue the RI activity can be improved), $-CF_3$ (to evaluate the effect of the electron withdrawing groups on the function of the amide bond) and –Bn/Ph (to examine the effect of the aromatic groups on RI activity).

3. To improve the method for predicting hydration and subsequently RI activity of the AFGP analogues following steps should be undertaken:

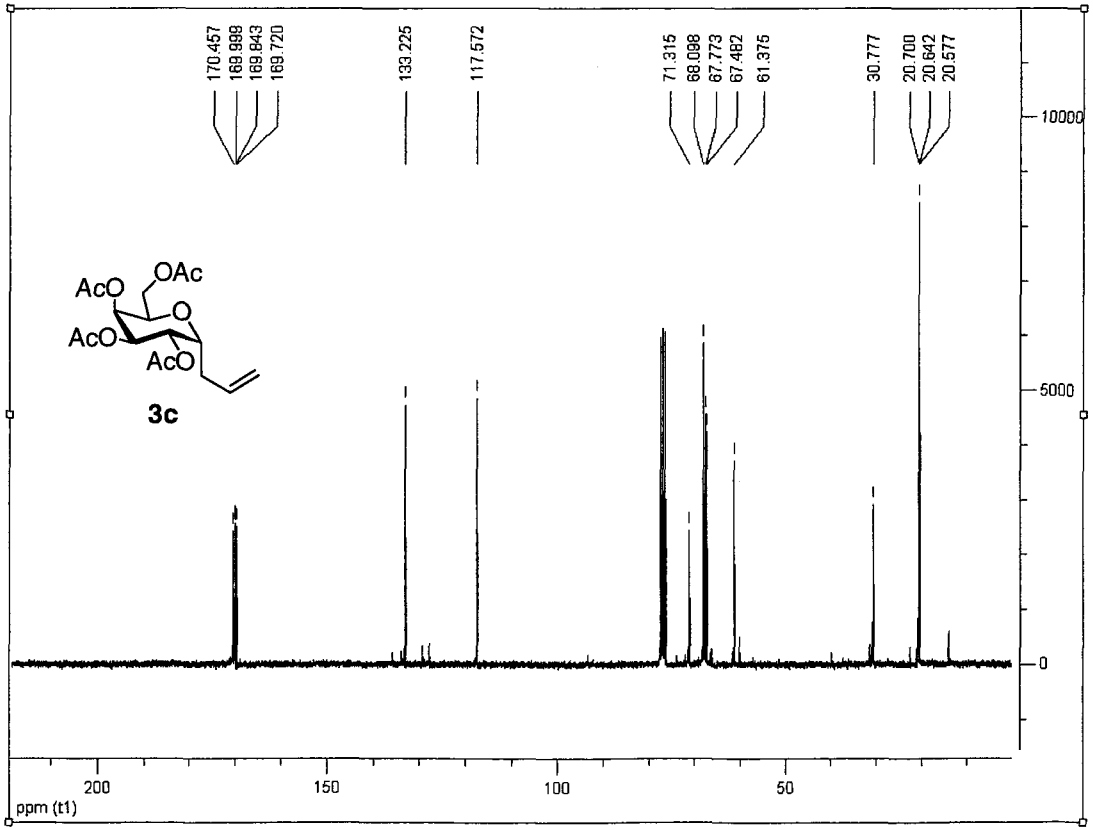
a. RI activity of amino acids should be assessed and compared to their hydration values presented in the literature

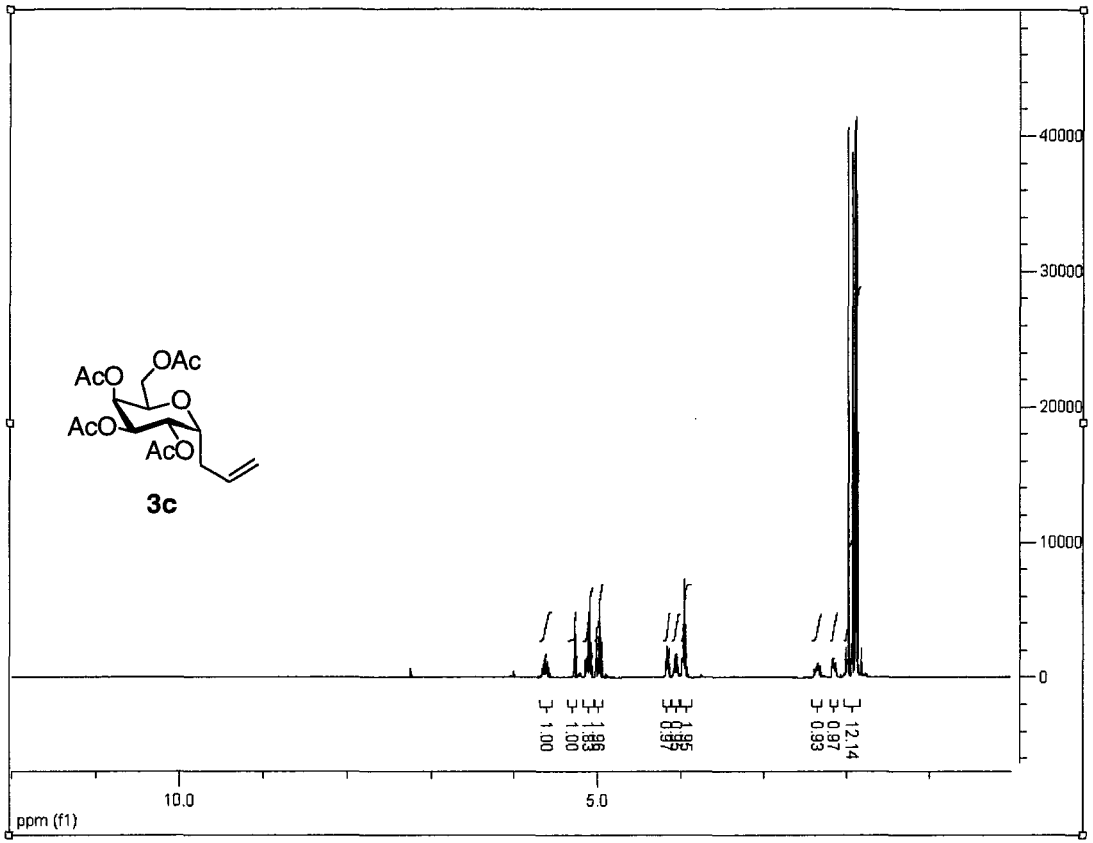
b. PMVs and hydration numbers of the most active analogues and of the key components of our analogues should be determined experimentally and compared to the estimated values

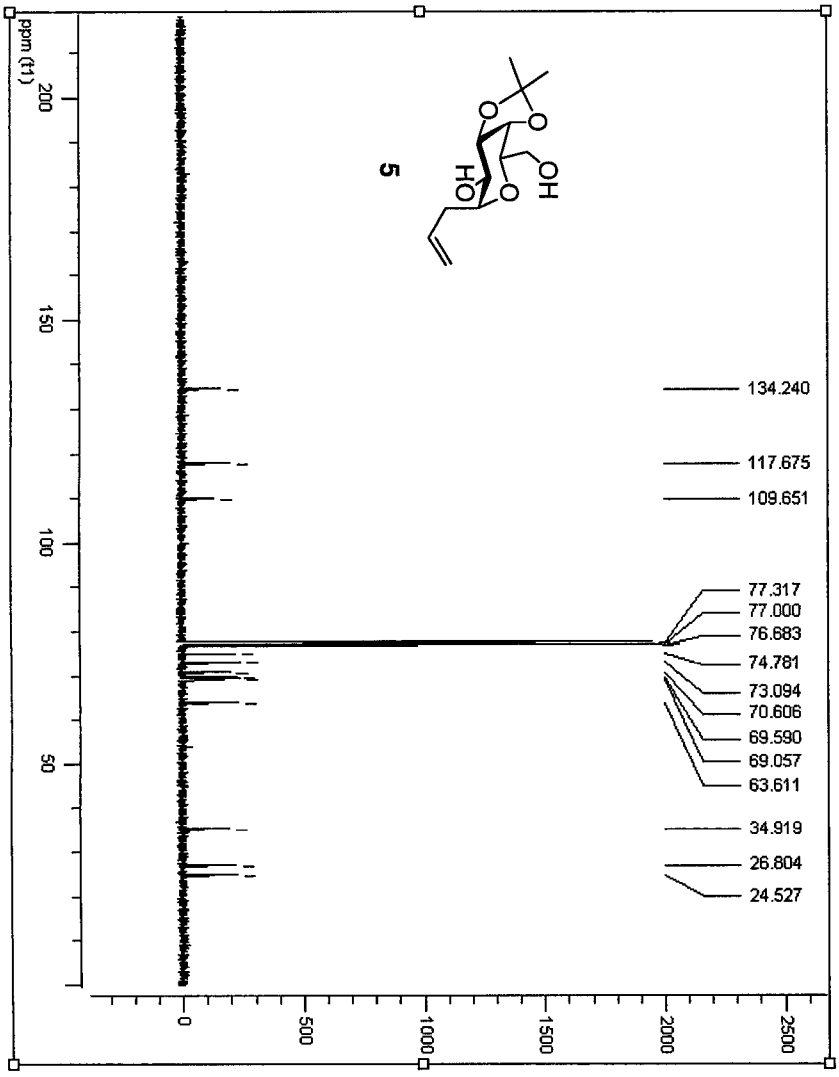
c. computational modeling with respect to hydration and conformation of our analogues should be routinely performed and compared to the RI activity.

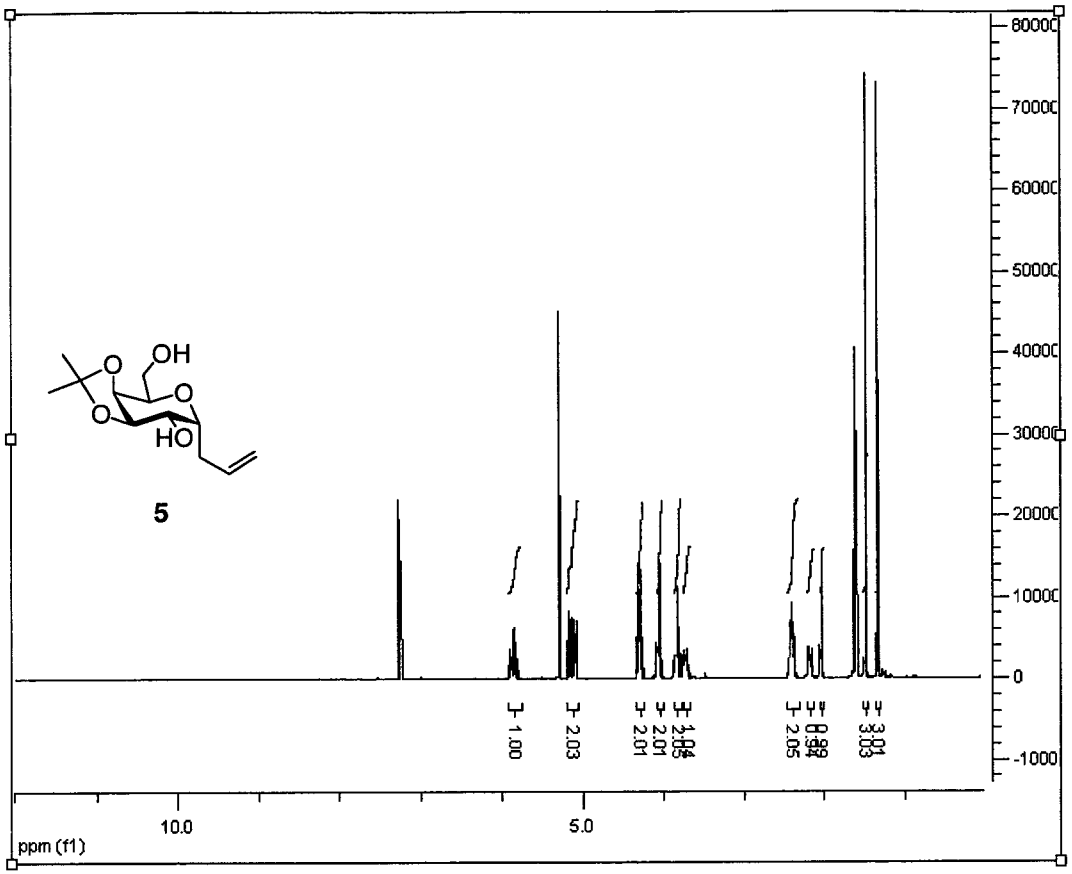
APPENDIX

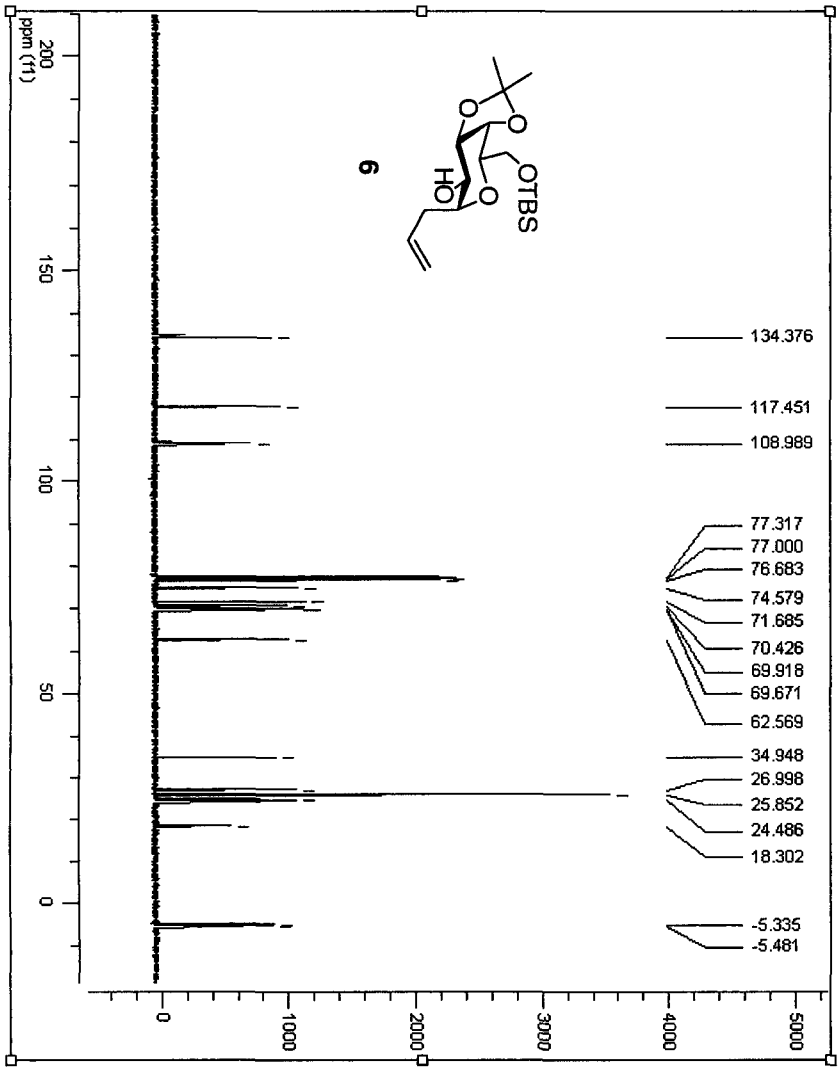
Selected ^{1}H and ^{13}C NMR Spectra

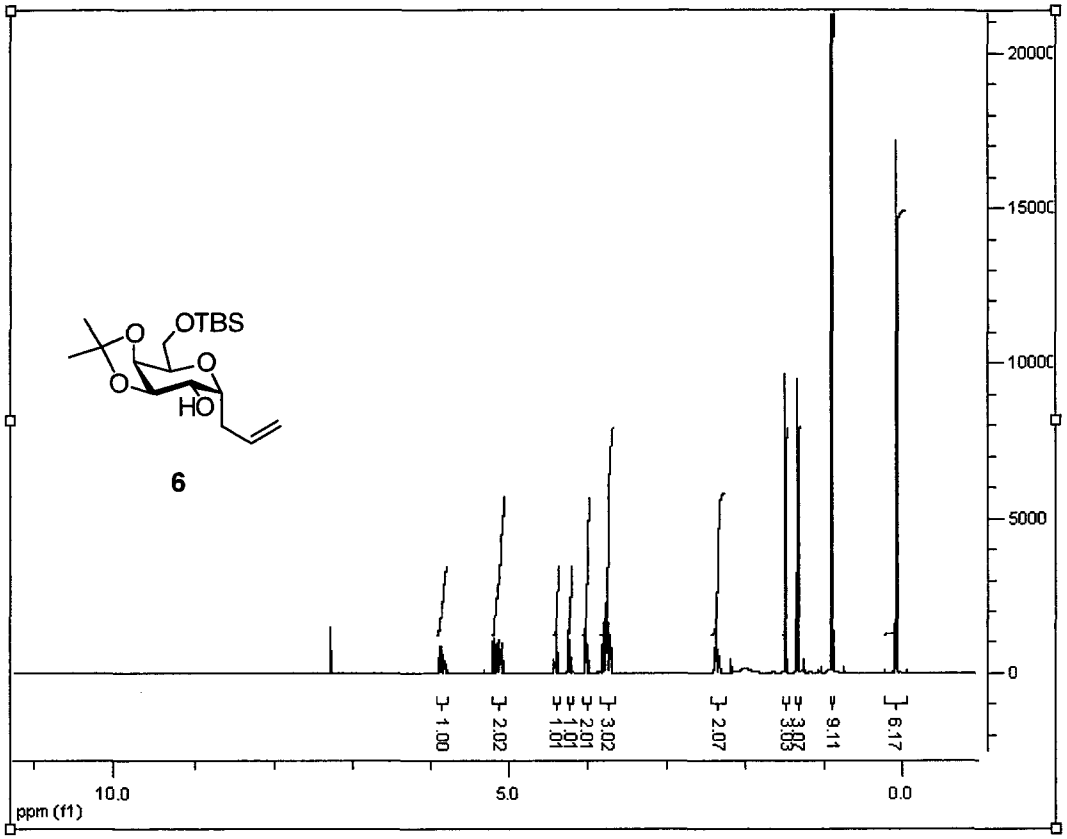

170.457
169.998
169.843
169.720
133.225
117.572
71.315
68.098
67.773
67.482
61.375
30.777
20.700
20.642
20.577
AcO
OAc
O
AcO
AcO
3c
10000
5000
0
200
150
100
50
ppm (t1)

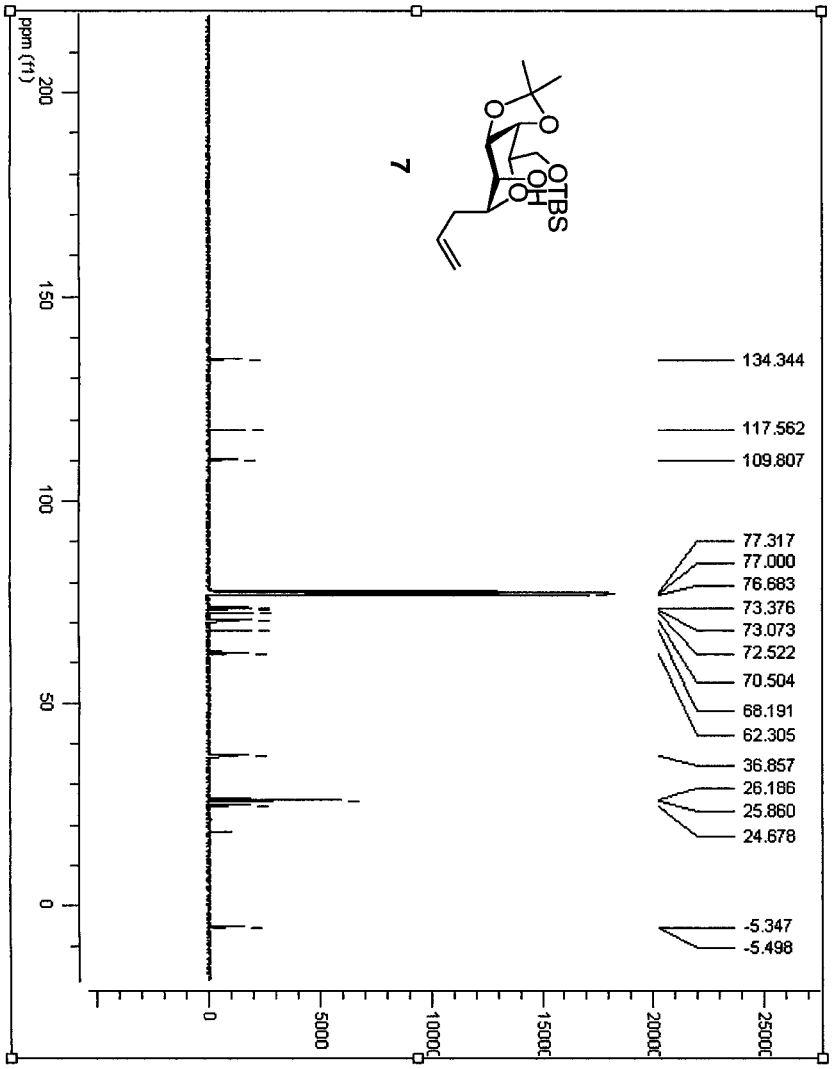

AcO
OAc
O
AcO
AcO
3c
40000
30000
20000
10000
0
1.00
1.00
1.96
1.83
1.95
0.97
0.97
0.93
12.14
10.0
5.0
ppm (f1)

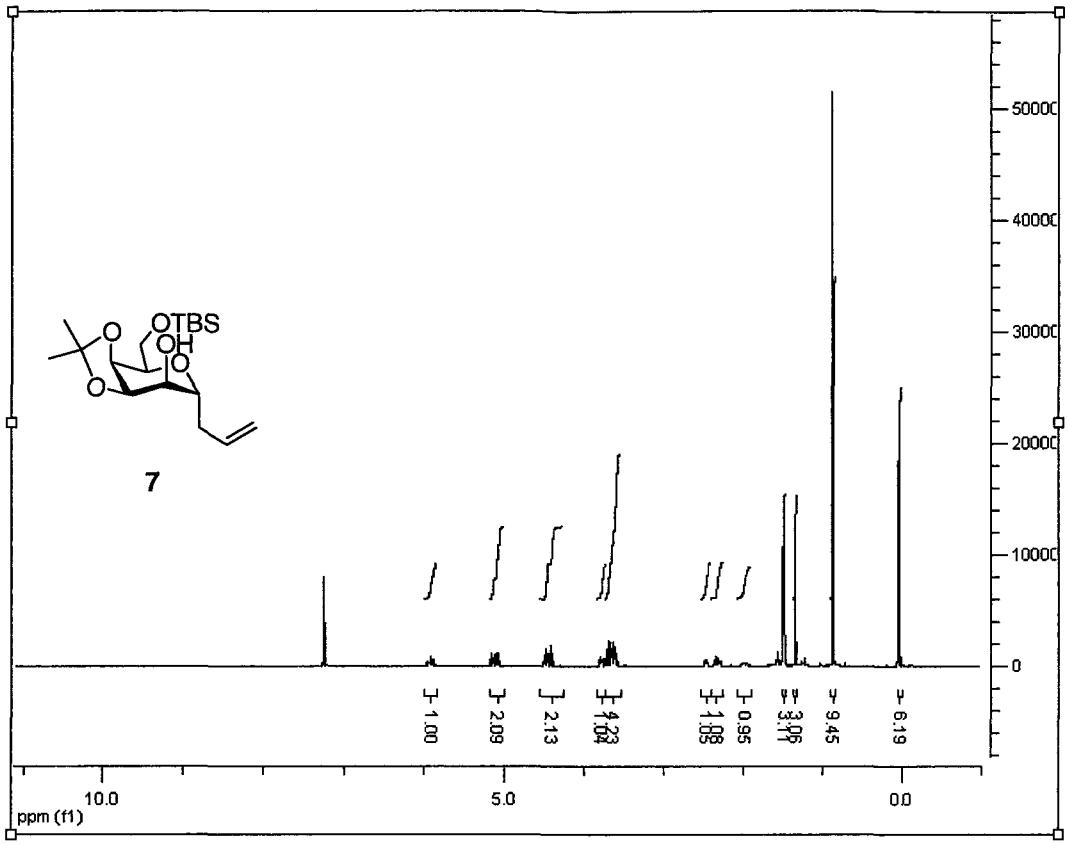

ppm (f1)
200
150
100
50
0
500
1000
1500
2000
2500
5
HO
OH
O
O
O
O
134.240
117.675
109.651
77.317
77.000
76.683
74.781
73.094
70.606
69.590
69.057
63.611
34.919
26.804
24.527

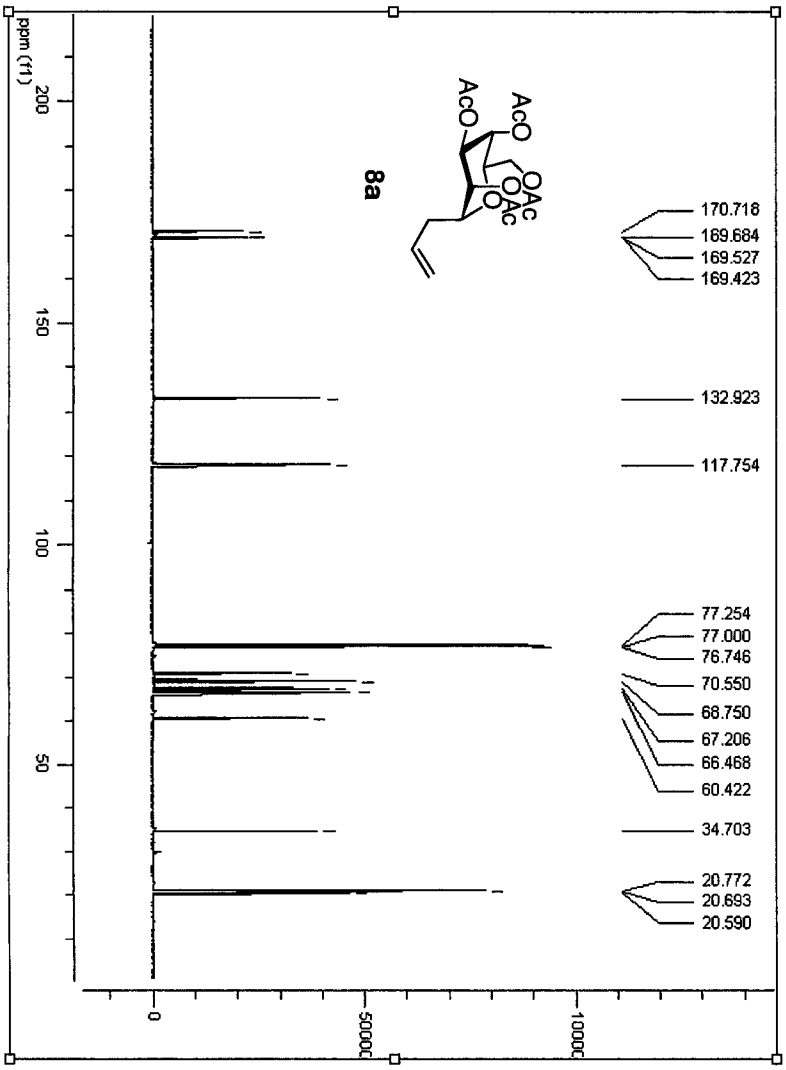

OH
O
O
O
HO
5
80000
70000
60000
50000
40000
30000
20000
10000
0
-1000
1.00
2.03
2.01
2.05
3.01
3.03
10.0
5.0
ppm (f1)

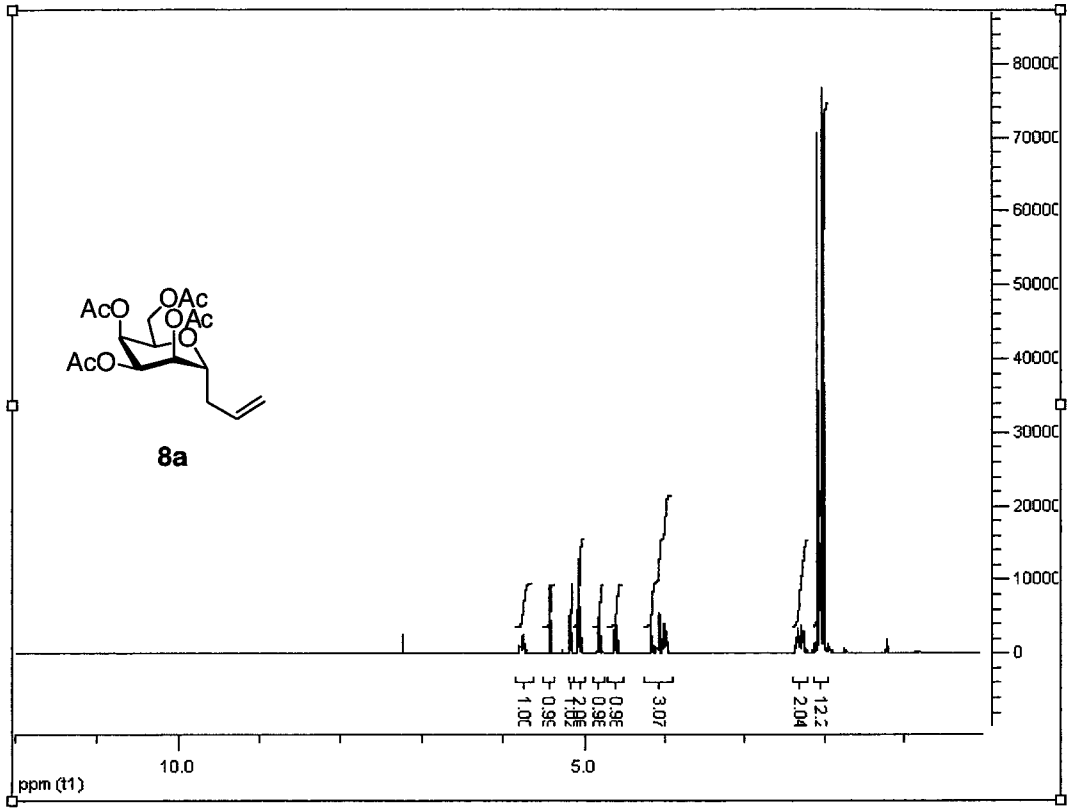

ppm (f1)
200
150
100
50
0
0
1000
2000
3000
4000
5000
6
HO
OTBS
134.376
117.451
108.989
77.317
77.000
76.683
74.579
71.685
70.426
69.918
69.671
62.569
34.948
26.998
25.852
24.486
18.302
-5.335
-5.481

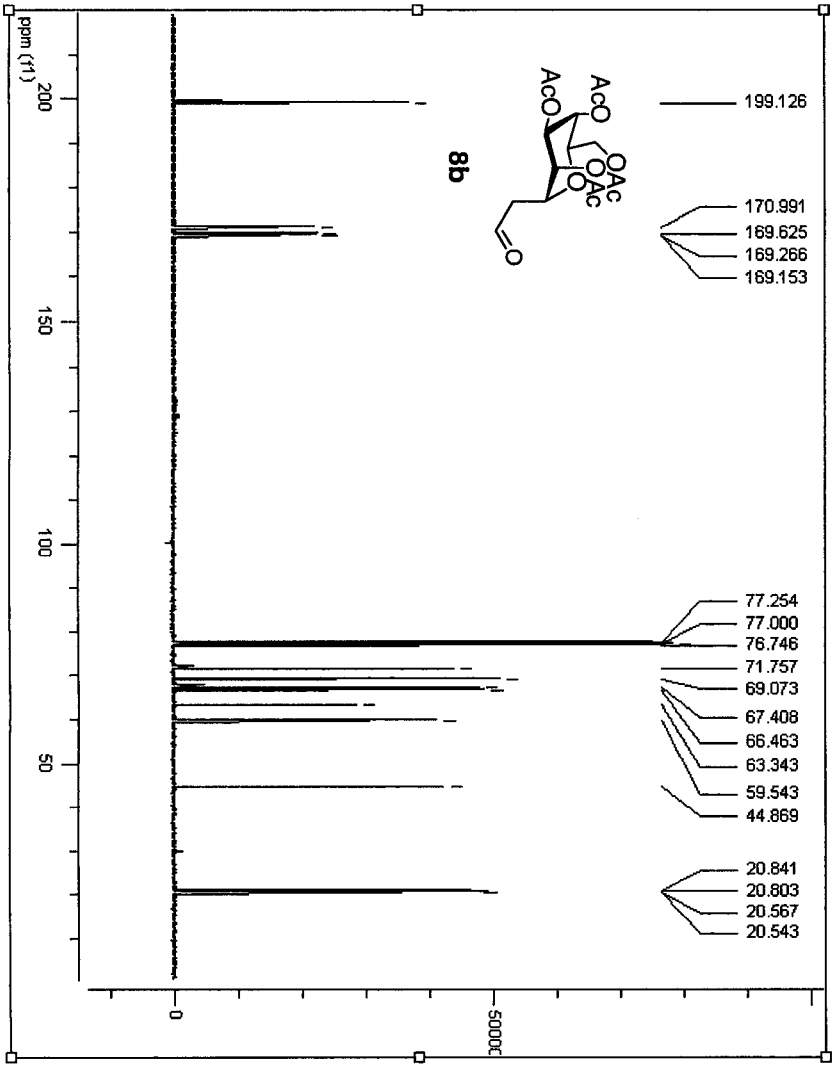

OTBS
O
O
O
HO
6
20000
15000
10000
5000
0
1.00
2.02
1.01
1.01
2.01
3.02
2.07
3.03
3.03
9.11
6.17
10.0
5.0
0.0
ppm (f1)

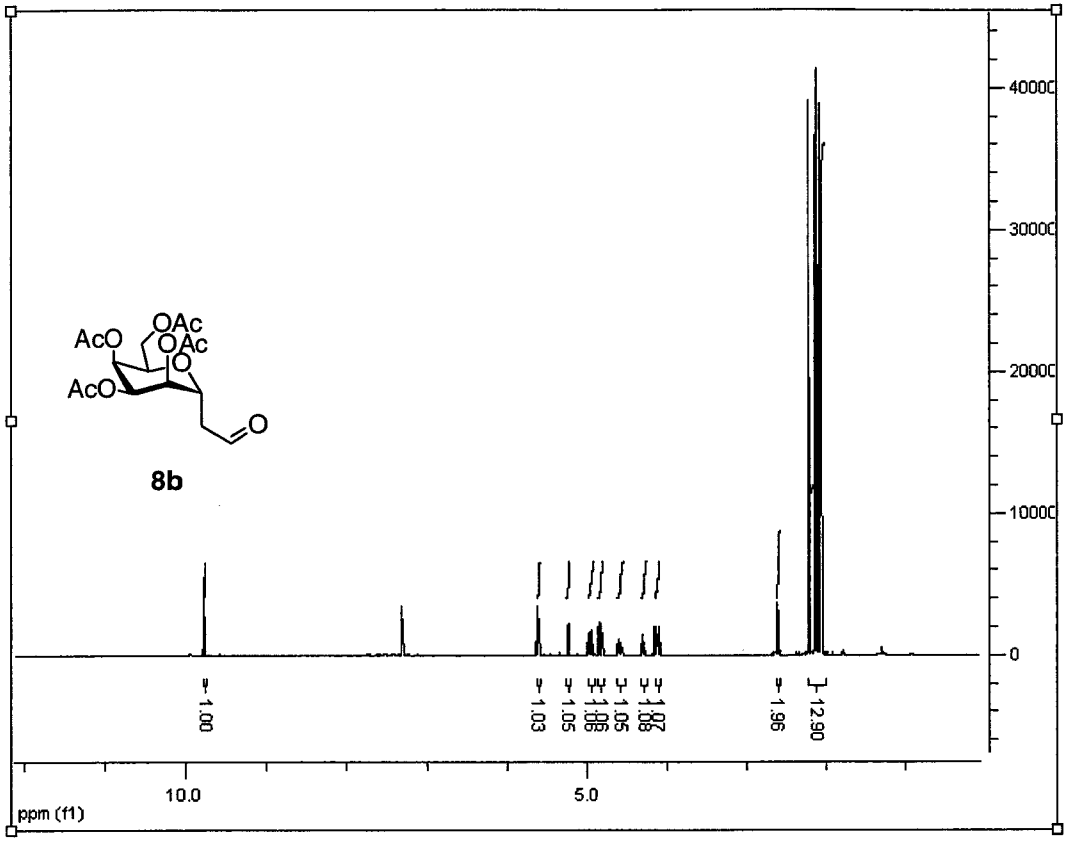

7
OTBS
OH
O
O
O
134.344
117.562
109.807
77.317
77.000
76.683
73.376
73.073
72.522
70.504
68.191
62.305
36.857
26.186
25.860
24.678
-5.347
-5.498
ppm (f1)
200
150
100
50
0
0
5000
10000
15000
20000
25000

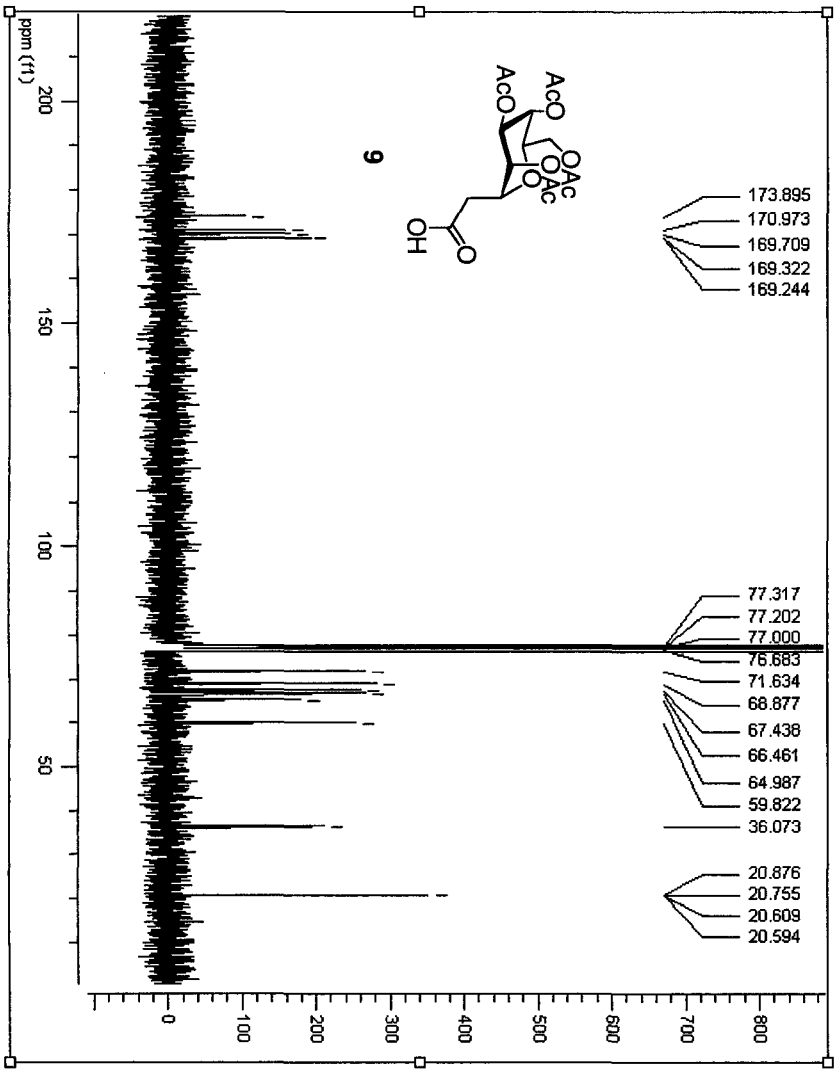

OTBS
OH
O
O
O
7
50000
40000
30000
20000
10000
0
1.00
2.09
2.13
4.23
1.08
0.95
3.06
9.45
6.19
10.0
5.0
0.0
ppm (f1)

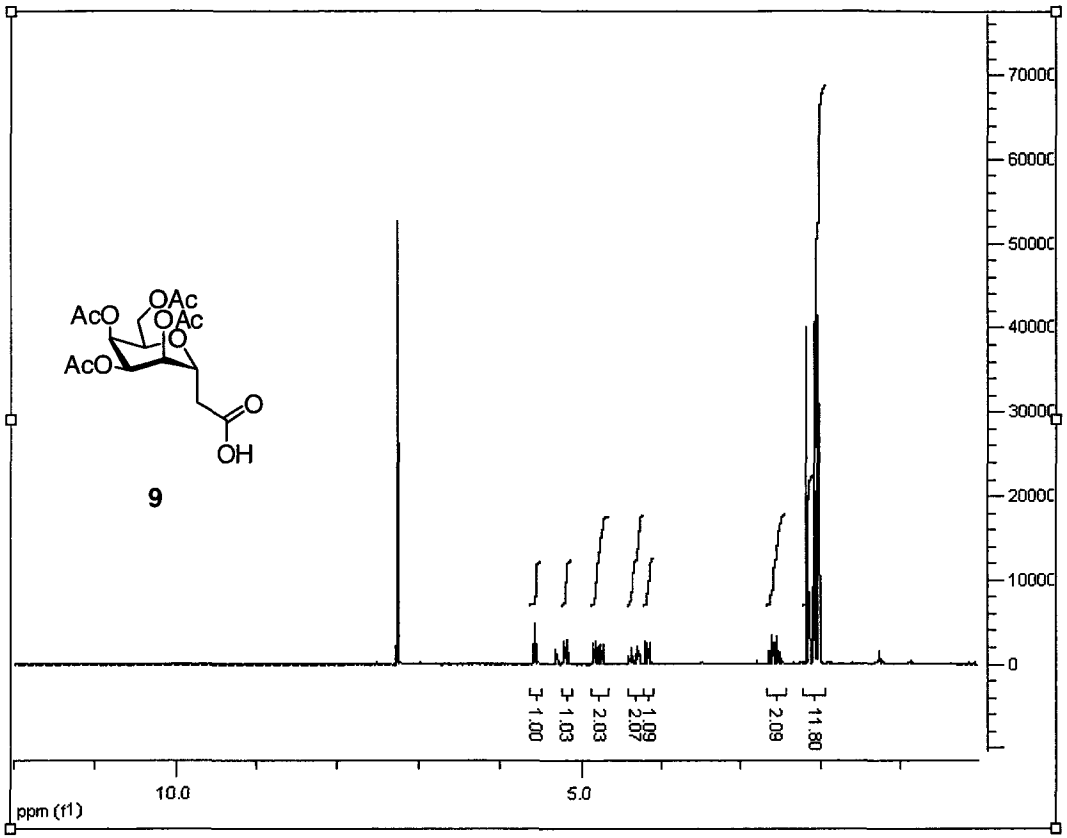

ppm (f1)
200
150
100
50
0
50000
100000
AcO
AcO
OAc
OAc
O
8a
170.718
169.684
169.527
169.423
132.923
117.754
77.254
77.000
76.746
70.550
68.750
67.206
66.468
60.422
34.703
20.772
20.693
20.590

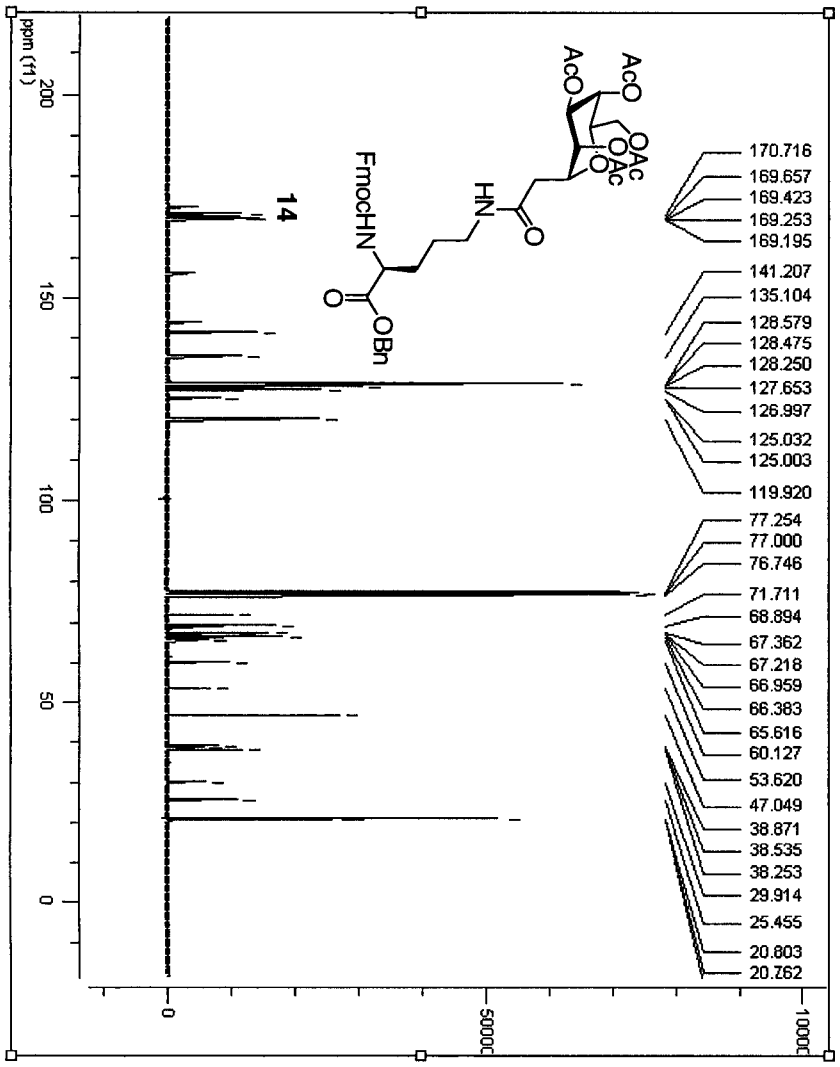

AcO
OAc
OAc
O
AcO
8a
80000
70000
60000
50000
40000
30000
20000
10000
0
1.00
0.95
7.82
0.96
0.96
3.07
12.2
2.04
10.0
5.0
ppm (t1)

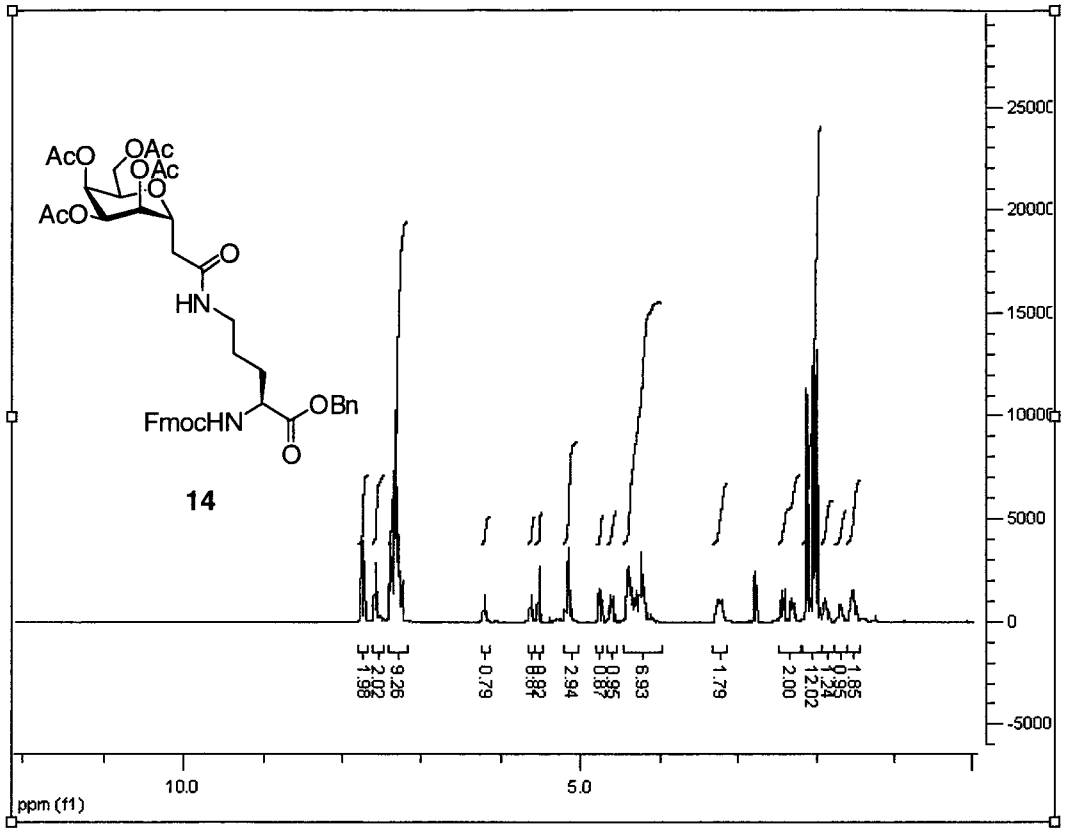

ppm (f1)
200
150
100
50
0
5000
8b
AcO
AcO
OAc
OAc
O
O
199.126
170.991
169.625
169.266
169.153
77.254
77.000
76.746
71.757
69.073
67.408
66.463
63.343
59.543
44.869
20.841
20.803
20.567
20.543

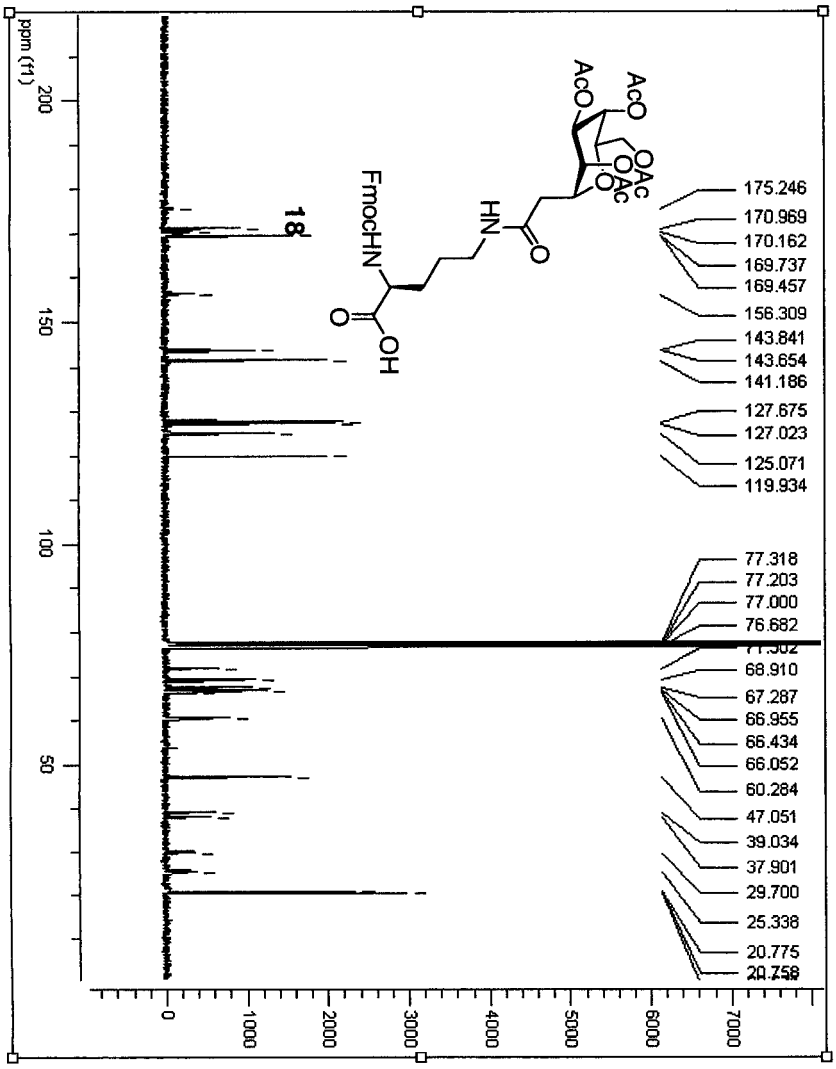

AcO
OAc
OAc
O
AcO
O
8b
40000
30000
20000
10000
0
1.00
1.03
1.05
1.09
1.05
1.08
1.07
1.96
12.90
10.0
5.0
ppm (f1)

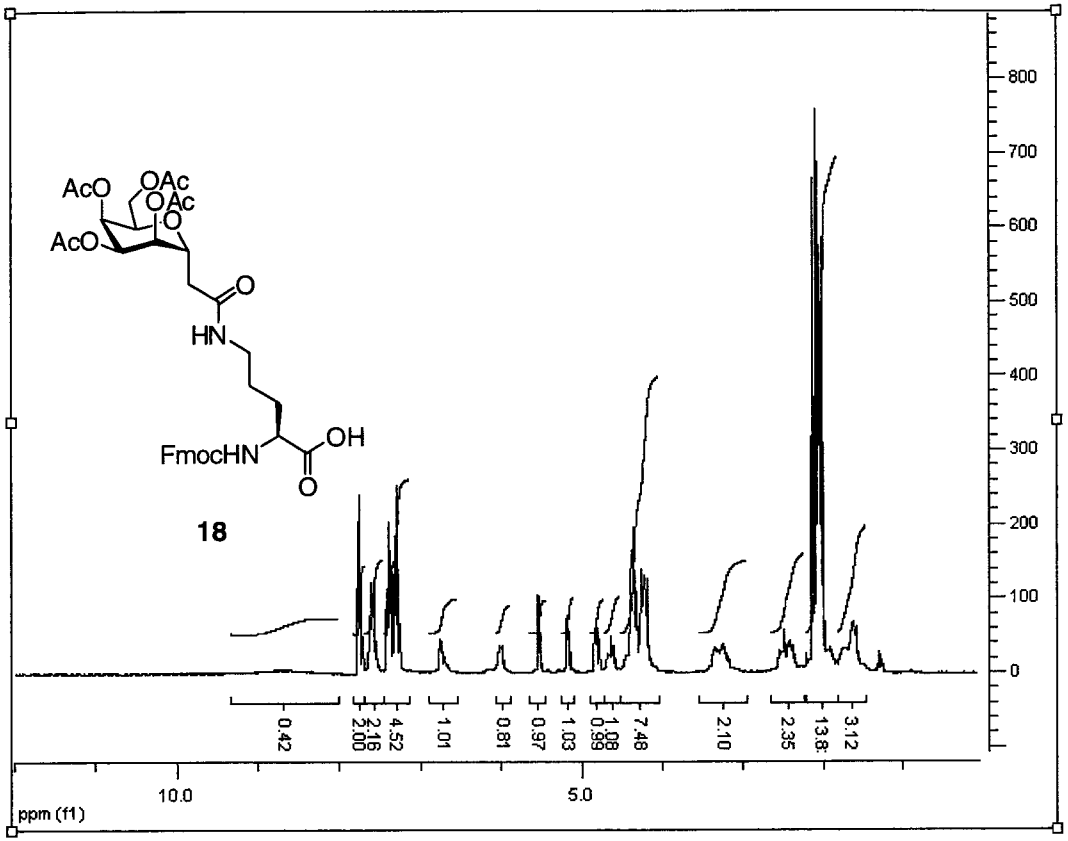

ppm (f1)
200
150
100
50
0
100
200
300
400
500
600
700
800
AcO
AcO
OAc
OAc
O
OH
O
9
173.895
170.973
169.709
169.322
169.244
77.317
77.202
77.000
76.683
71.634
68.877
67.438
66.461
64.987
59.822
36.073
20.876
20.755
20.609
20.594

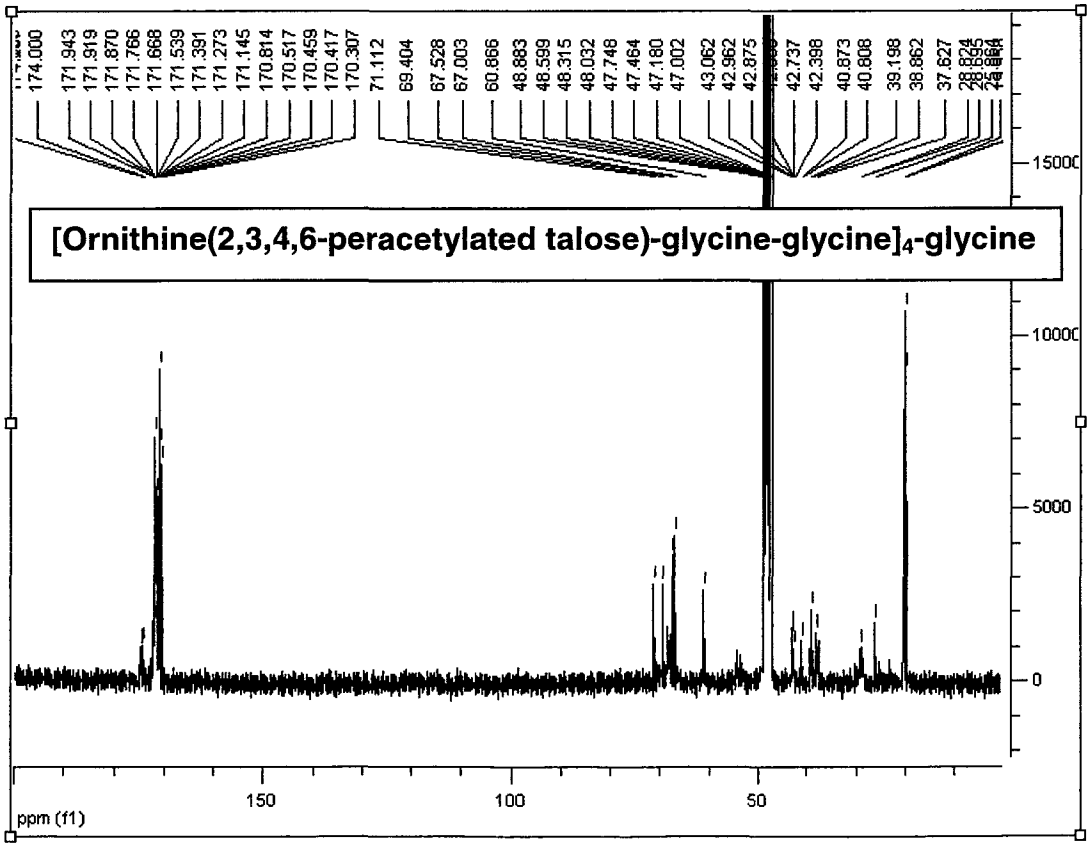

AcO
OAc
OAc
O
AcO
O
OH
9
1.00
1.03
2.03
2.07
1.09
2.09
11.80
10.0
5.0
ppm (f1)
70000
60000
50000
40000
30000
20000
10000
0

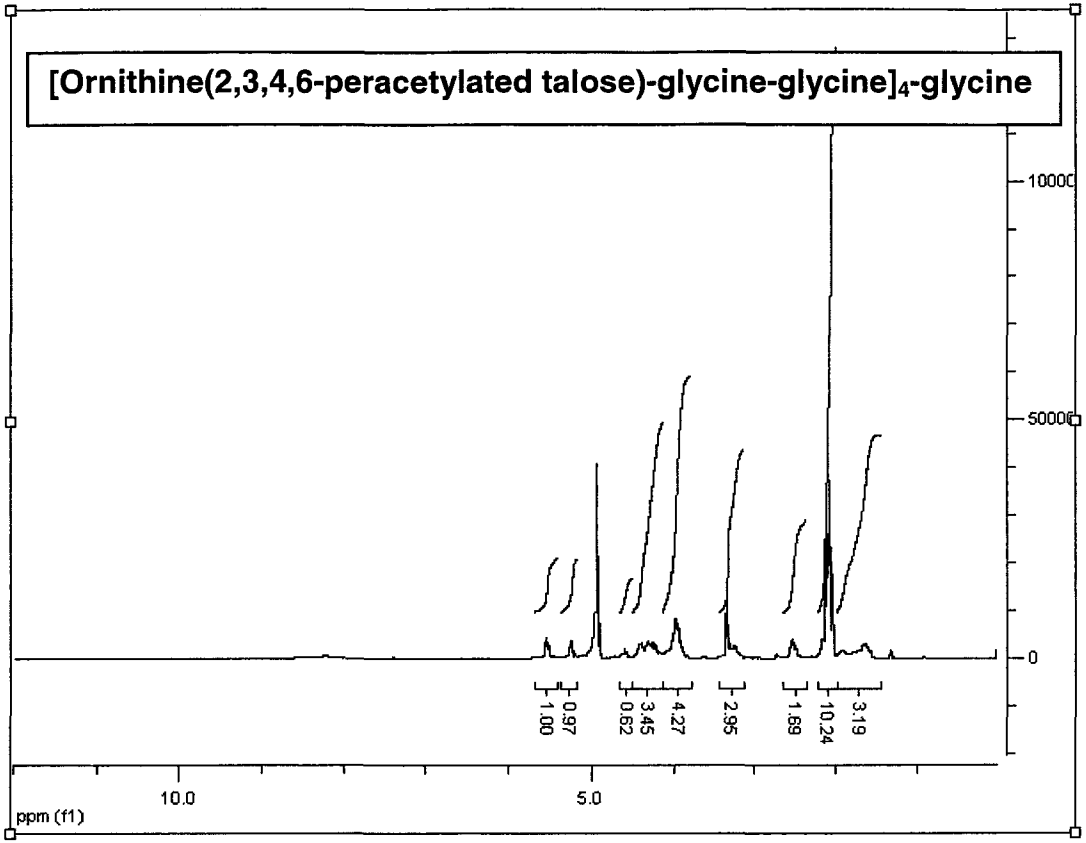

14
AcO
AcO
OAc
OAc
O
HN
O
FmocHN
O
OBn
170.716
169.657
169.423
169.253
169.195
141.207
135.104
128.579
128.475
128.250
127.653
126.997
125.032
125.003
119.920
77.254
77.000
76.746
71.711
68.894
67.362
67.218
66.959
66.383
65.616
60.127
53.620
47.049
38.871
38.535
38.253
29.914
25.455
20.803
20.762
ppm (f1)
200
150
100
50
0
0
5000
10000

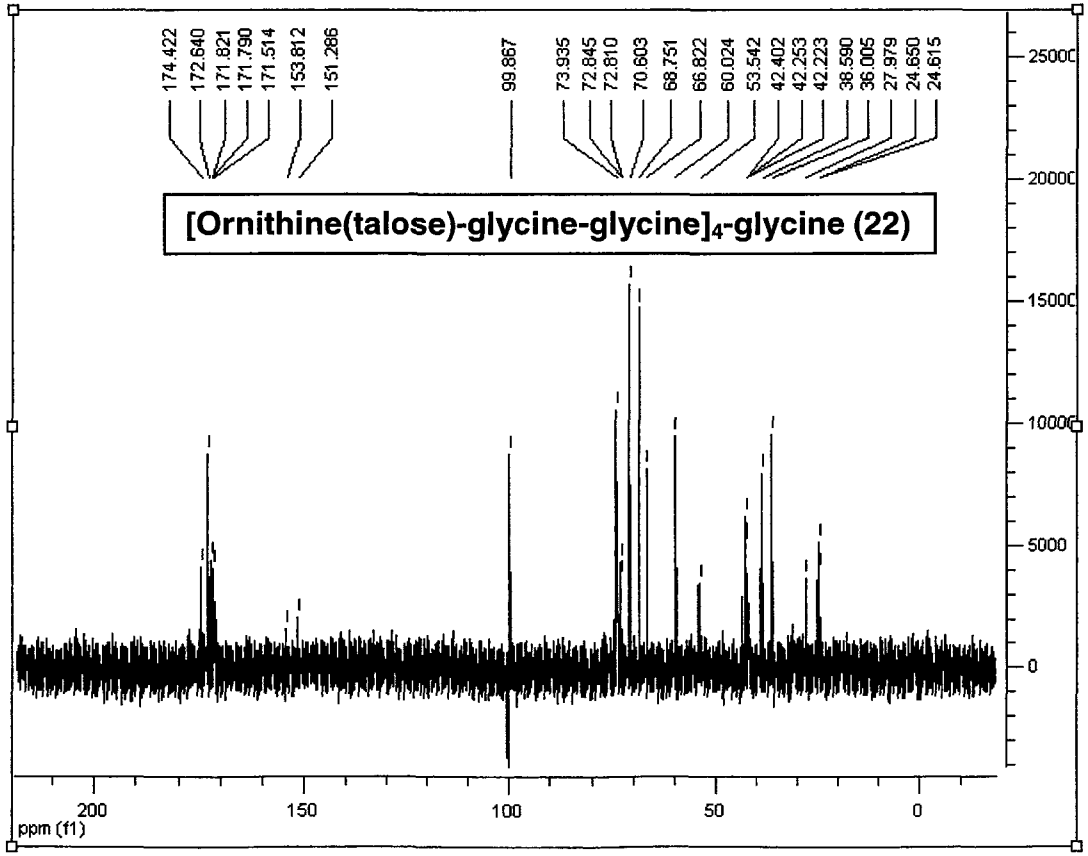

AcO
OAc
OAc
O
AcO
O
HN
FmocHN
OBn
O
14
25000
20000
15000
10000
5000
0
-5000
1.98
2.02
9.26
0.79
0.87
0.87
2.94
0.85
0.87
6.93
1.79
2.00
12.02
1.24
0.95
1.85
10.0
5.0
ppm (f1)

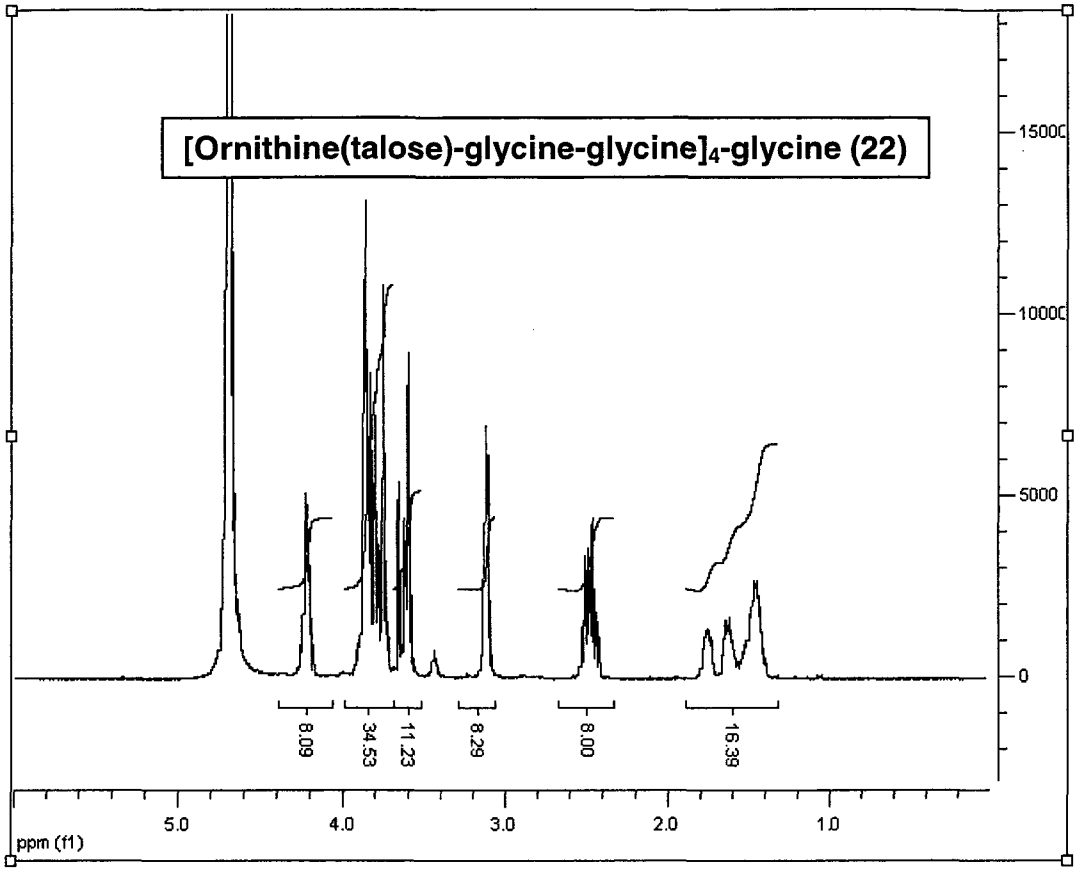

18'
FmocHN
HN
OH
AcO
OAc
175.246
170.969
170.162
169.737
169.457
156.309
143.841
143.654
141.186
127.675
127.023
125.071
119.934
77.318
77.203
77.000
76.682
68.910
67.287
66.955
66.434
66.052
60.284
47.051
39.034
37.901
29.700
25.338
20.775
ppm (f1)
200
150
100
50
0
1000
2000
3000
4000
5000
6000
7000

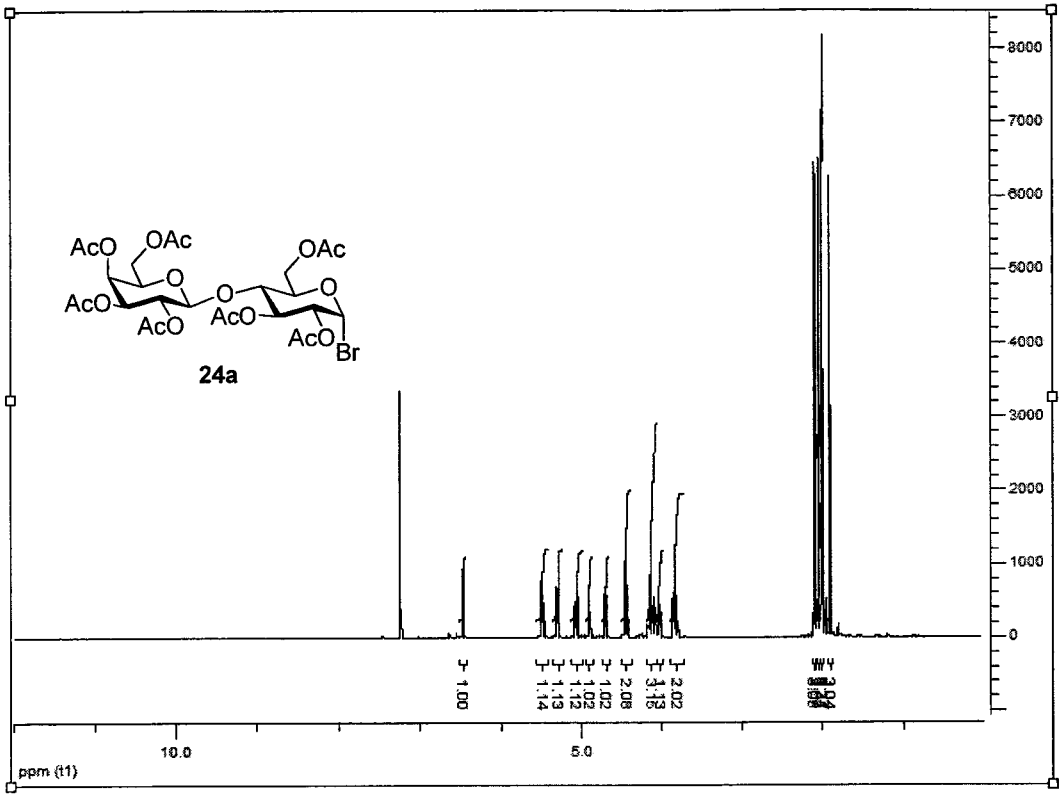

AcO
OAc
OAc
AcO
O
HN
O
FmocHN
OH
O
18
800
700
600
500
400
300
200
100
0
0.42
2.16
2.00
4.52
1.01
0.81
0.97
1.03
1.08
0.99
7.48
2.10
2.35
13.8
3.12
10.0
5.0
ppm (f1)

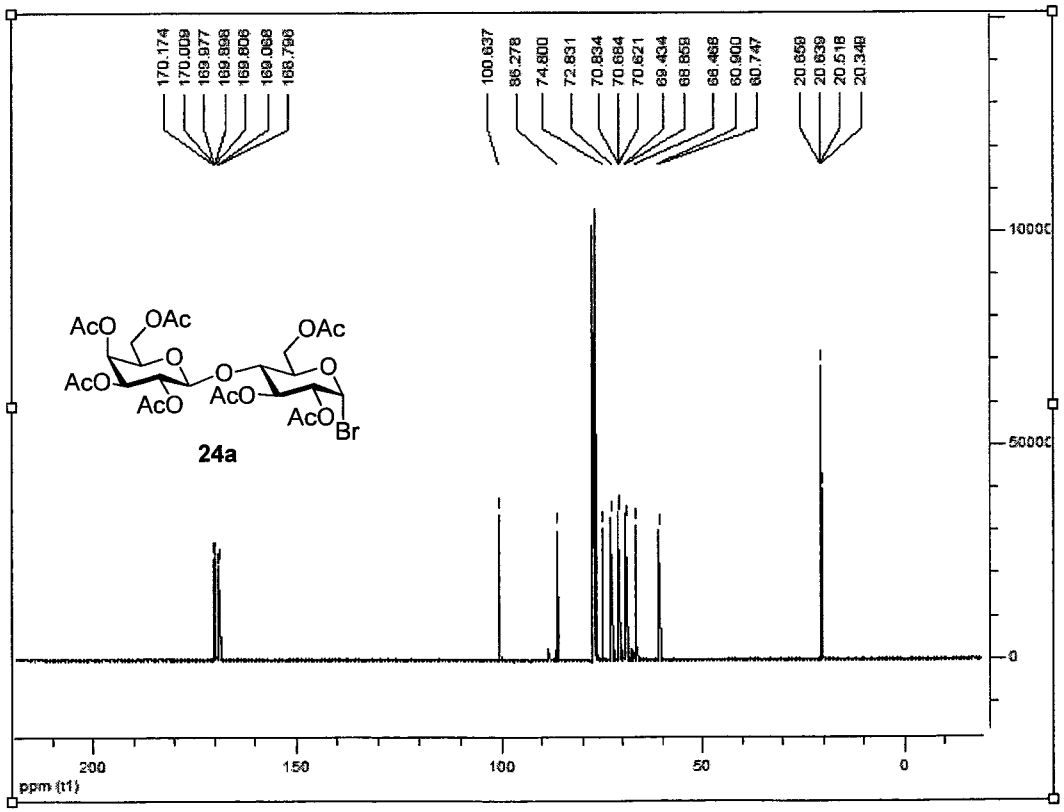

174.000
171.943
171.919
171.870
171.766
171.668
171.539
171.391
171.273
171.145
170.814
170.517
170.459
170.417
170.307
71.112
69.404
67.528
67.003
60.866
48.883
48.599
48.315
48.032
47.748
47.464
47.180
47.002
43.062
42.962
42.875
42.737
42.398
40.873
40.808
39.198
38.862
37.627
15000
[Ornithine(2,3,4,6-peracetylated talose)-glycine-glycine]4-glycine
10000
5000
0
150
100
50
ppm (f1)

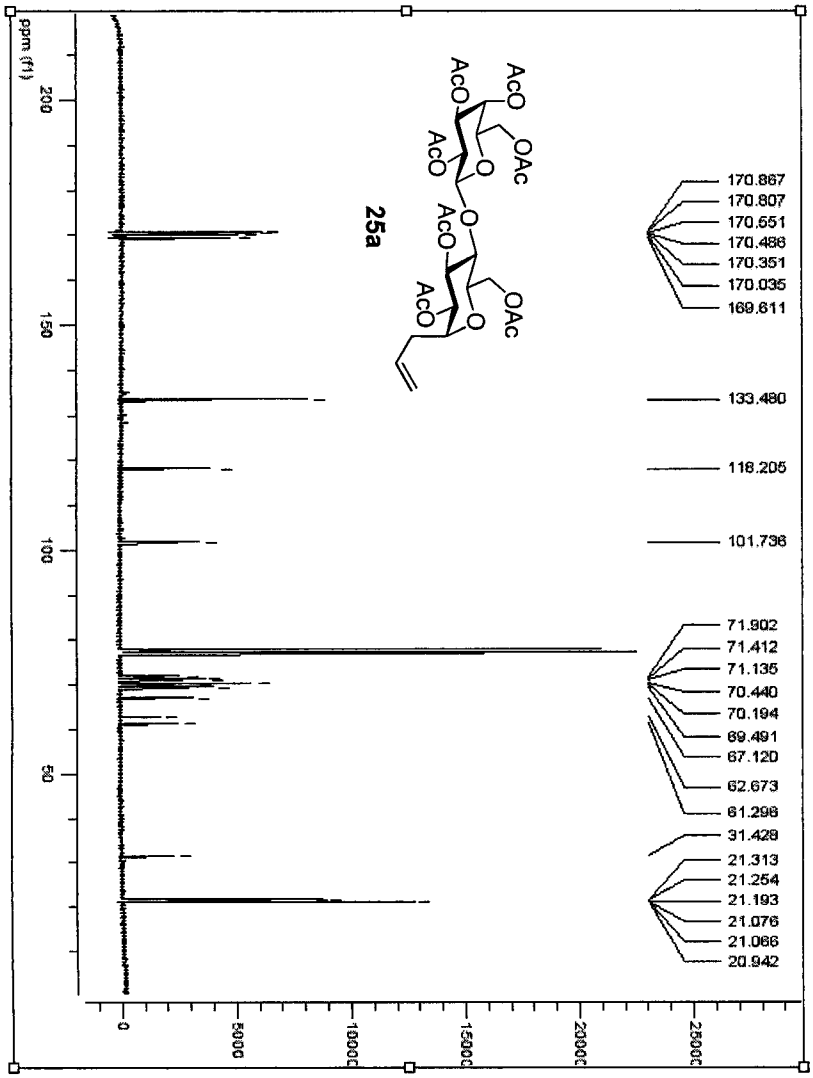

[Ornithine(2,3,4,6-peracetylated talose)-glycine-glycine]4-glycine
10000
50000
0
1.00
0.97
0.62
3.45
4.27
2.95
1.69
10.24
3.19
10.0
5.0
ppm (f1)

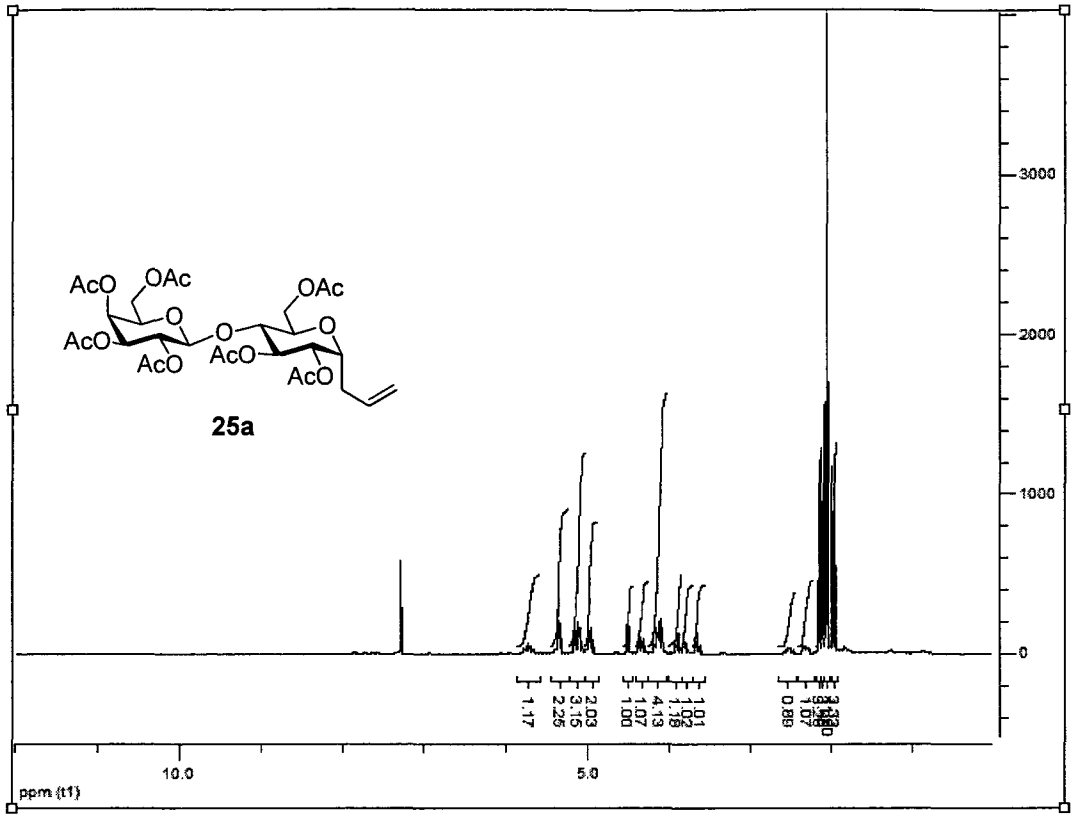

174.422
172.640
171.821
171.790
171.514
153.812
151.286
99.867
73.935
72.845
72.810
70.603
68.751
66.822
60.024
53.542
42.402
42.253
42.223
38.590
36.005
27.979
24.650
24.615
[Ornithine(talose)-glycine-glycine]4-glycine (22)
25000
20000
15000
10000
5000
0
200
150
100
50
0
ppm (f1)

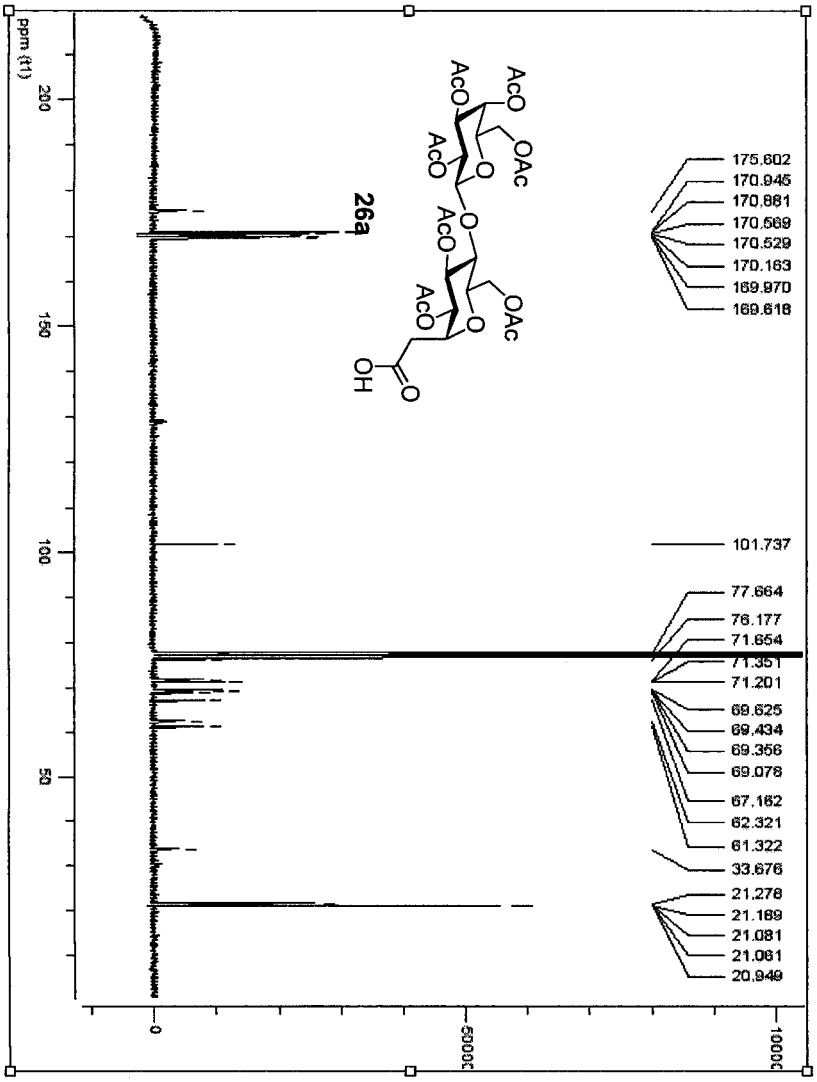

[Ornithine(talose)-glycine-glycine]4-glycine (22)
15000
10000
5000
0
8.09
34.53
11.23
8.29
8.00
16.39
5.0
4.0
3.0
2.0
1.0
ppm (f1)

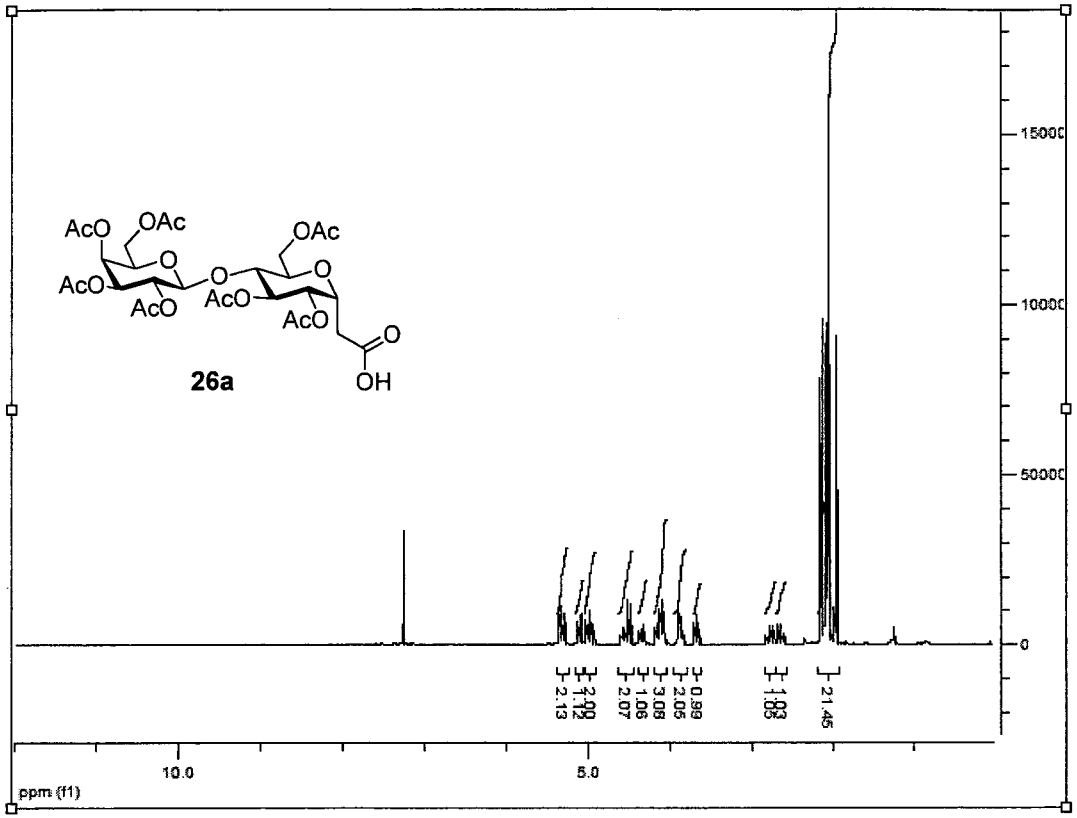

AcO
OAc
AcO
AcO
AcO
OAc
AcO
Br
24a
8000
7000
6000
5000
4000
3000
2000
1000
0
1.00
1.14
1.13
1.12
1.02
2.08
3.13
2.02
10.0
5.0
ppm (t1)

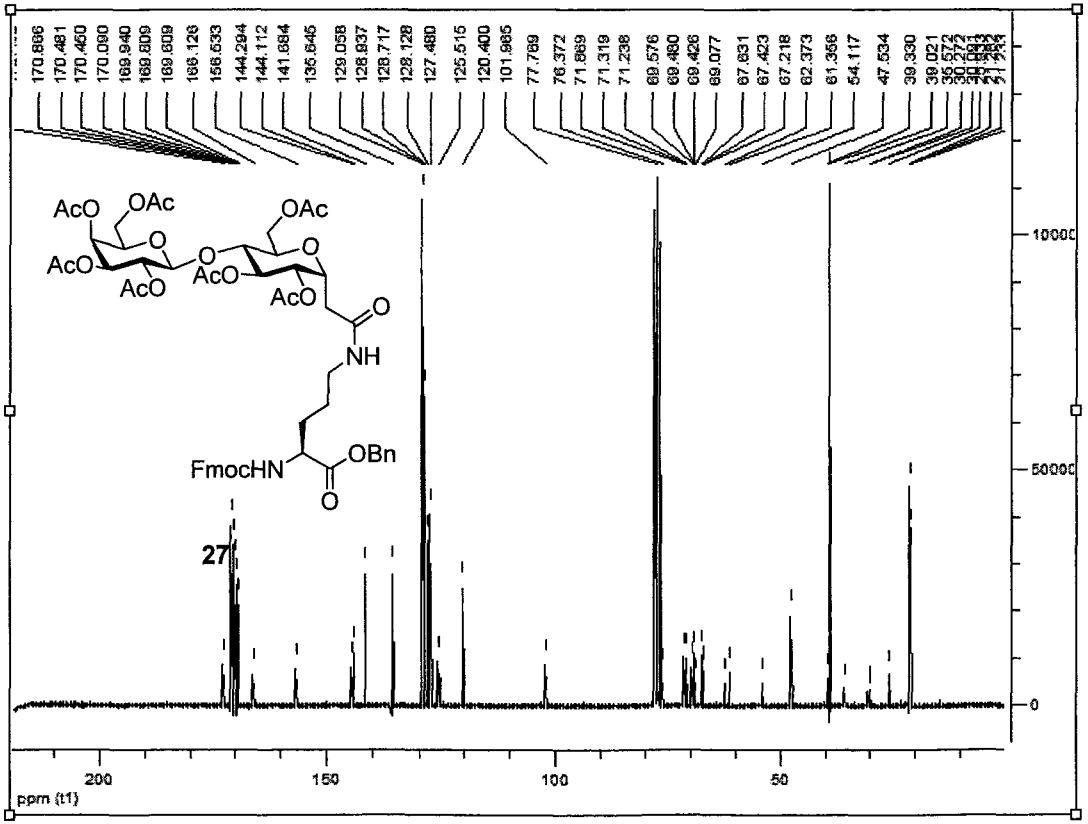

170.174
170.009
169.977
169.898
169.806
169.088
168.796
100.637
95.278
74.800
72.831
70.834
70.684
70.621
69.434
68.859
68.488
60.900
60.747
20.859
20.639
20.518
20.349
AcO
OAc
AcO
AcO
O
O
AcO
OAc
AcO
Br
24a
10000
50000
0
200
150
100
50
0
ppm (t1)

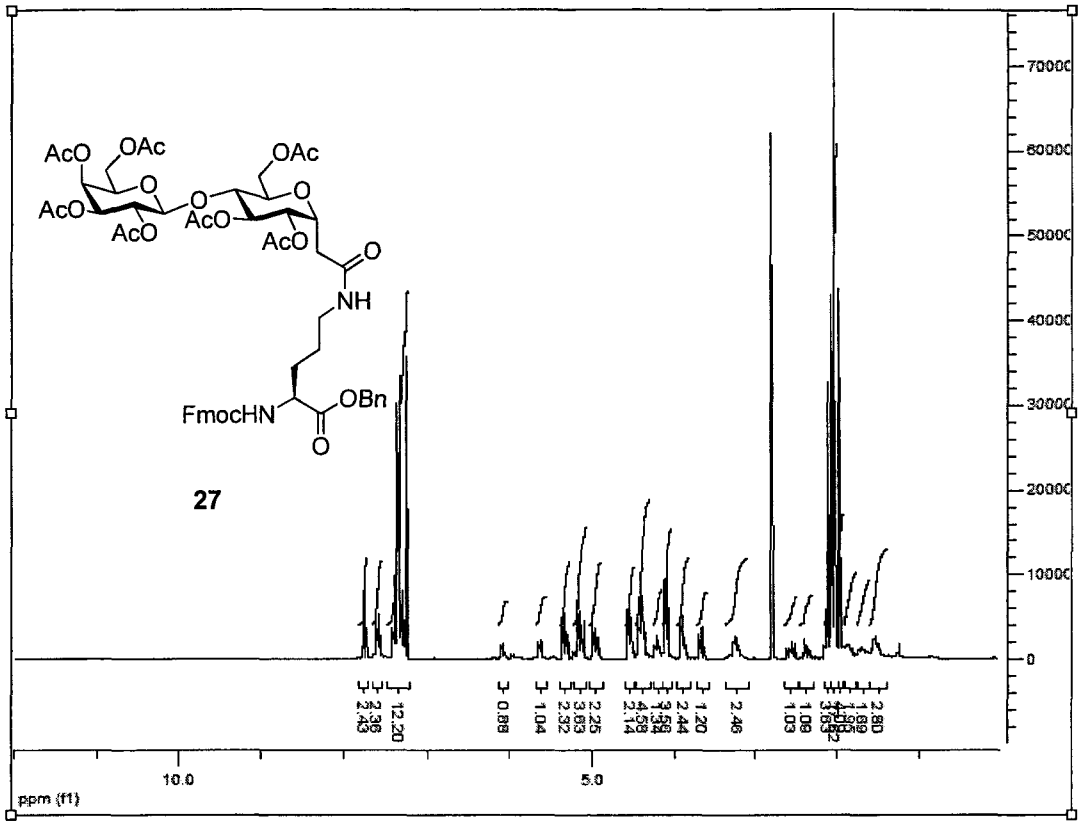

ppm (f1)
200
150
100
50
0
5000
10000
15000
20000
25000
AcO
AcO
AcO
AcO
OAc
OAc
O
O
O
O
25a
170.867
170.807
170.551
170.486
170.351
170.035
169.611
133.480
118.205
101.736
71.902
71.412
71.135
70.440
70.194
69.491
67.120
62.673
61.298
31.428
21.313
21.254
21.193
21.076
21.066
20.942

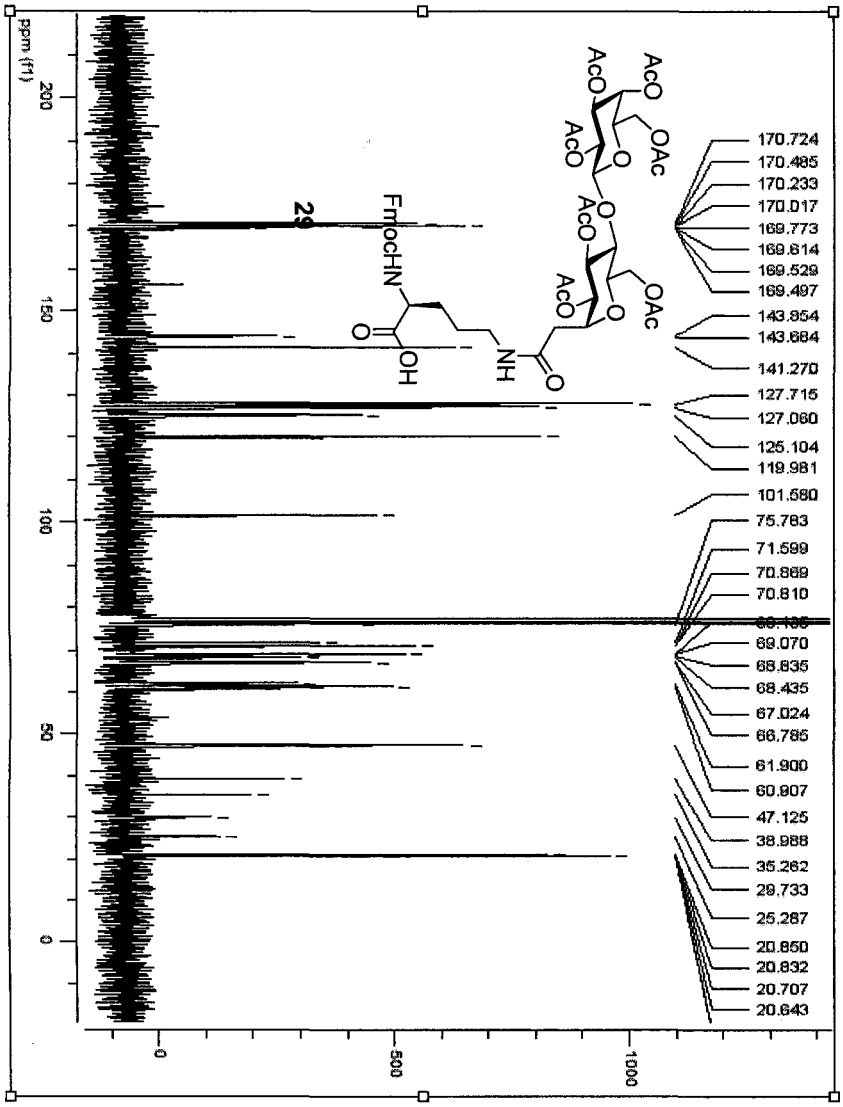

AcO
OAc
AcO
AcO
OAc
AcO
AcO
25a
3000
2000
1000
0
1.17
2.25
3.15
2.03
1.00
1.07
4.13
1.18
1.02
1.01
0.89
10.0
5.0
ppm (t1)

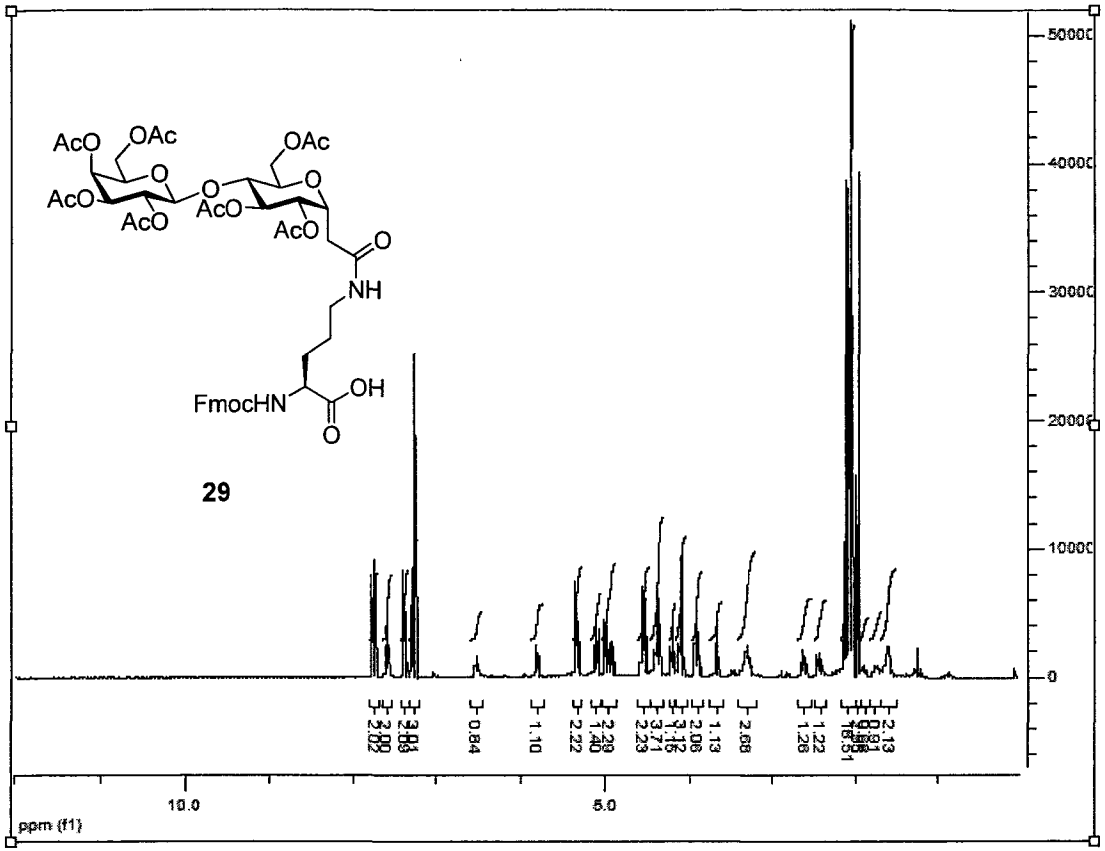

26a
AcO
AcO
AcO
OAc
O
AcO
AcO
O
AcO
OAc
O
OH
O
175.602
170.945
170.881
170.569
170.529
170.163
169.970
169.618
101.737
77.664
76.177
71.654
71.351
71.201
69.625
69.434
69.356
69.078
67.162
62.321
61.322
33.676
21.278
21.189
21.081
21.061
20.949
ppm (f1)
200
150
100
50
0
5000
10000


AcO
OAc
AcO
AcO
O
AcO
OAc
AcO
O
OAc
O
O
OH
26a
15000
10000
50000
0
2.13
1.12
3.09
2.07
1.06
3.08
2.05
0.89
1.03
1.03
21.45
10.0
5.0
ppm (f1)

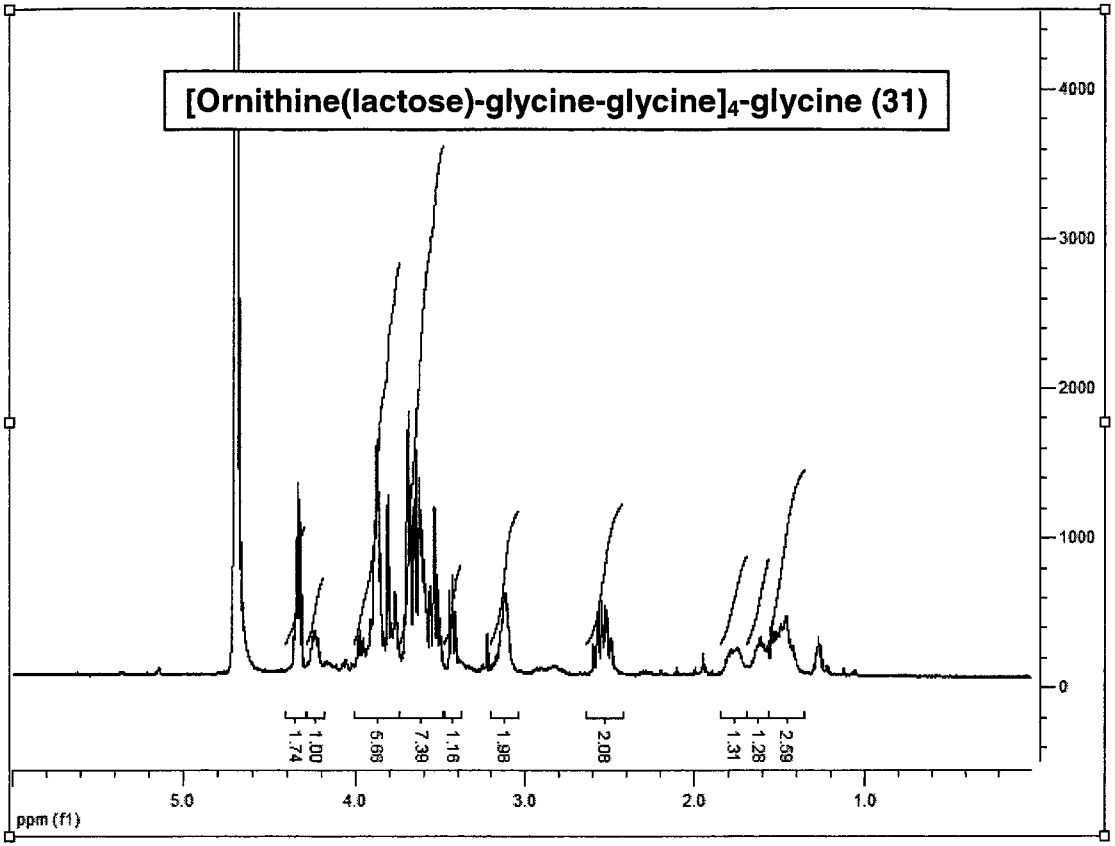

170.886
170.481
170.450
170.090
169.940
169.809
169.609
166.126
156.533
144.294
144.112
141.684
135.645
129.058
128.837
128.717
128.128
127.480
125.515
120.400
101.985
77.789
76.372
71.869
71.319
71.238
69.576
69.480
69.426
69.077
67.831
67.423
67.218
62.373
61.356
54.117
47.534
39.330
39.021
AcO
OAc
AcO
AcO
AcO
OAc
AcO
O
NH
FmocHN
OBn
O
27
10000
50000
0
200
150
100
50
ppm (t1)

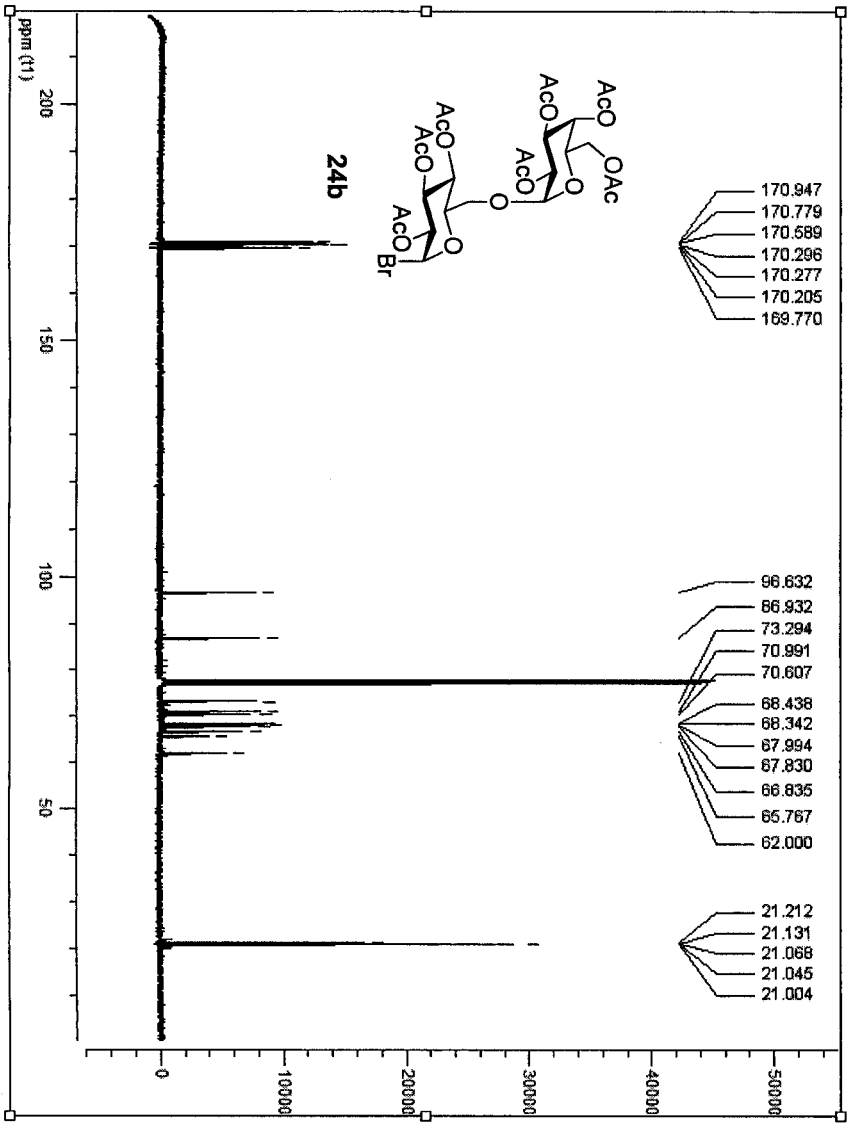

AcO
OAc
AcO
AcO
AcO
OAc
AcO
AcO
O
NH
FmocHN
OBn
O
27
70000
60000
50000
40000
30000
20000
10000
0
10.0
5.0
ppm (f1)

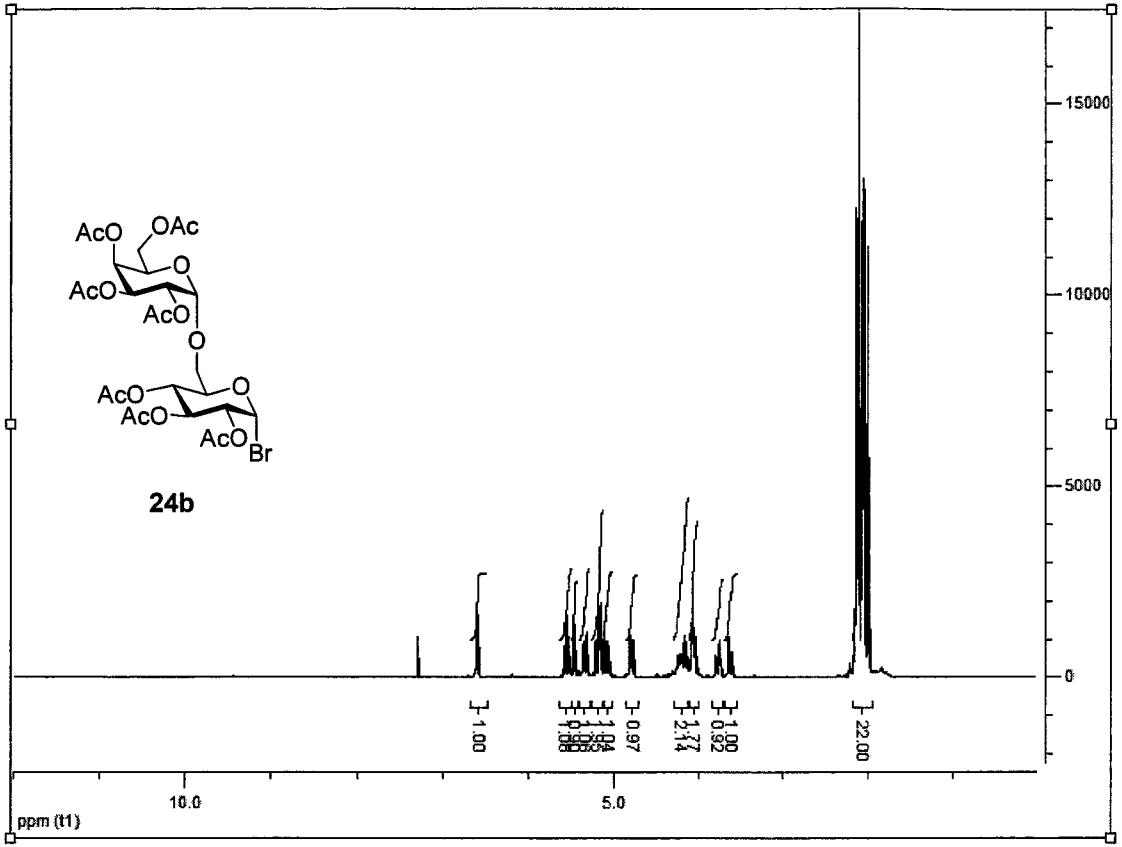

ppm (f1)
200
150
100
50
0
0
500
1000
170.724
170.485
170.233
170.017
169.773
169.614
169.529
169.497
143.854
143.684
141.270
127.715
127.060
125.104
119.981
101.580
75.783
71.599
70.869
70.810
69.070
68.835
68.435
67.024
66.785
61.900
60.907
47.125
38.988
35.262
29.733
25.287
20.850
20.832
20.707
20.643
FmocHN
OH
NH
AcO
OAc

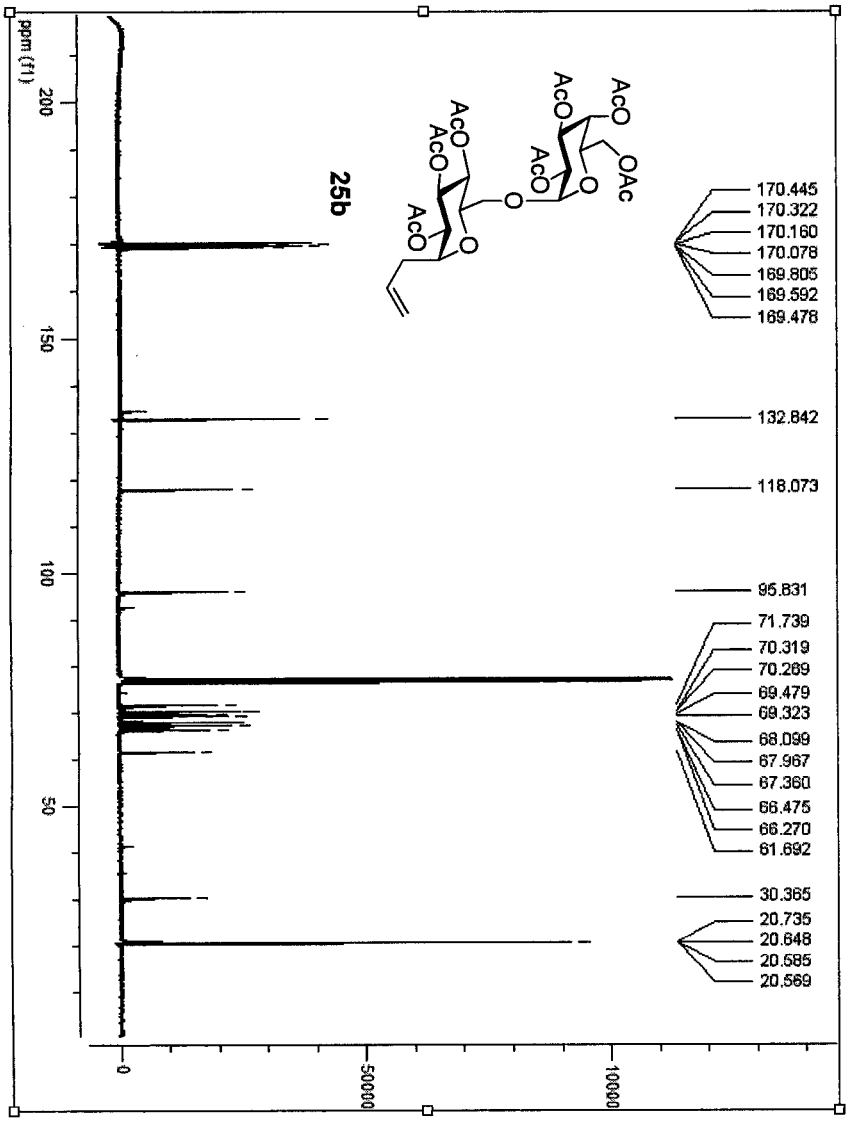

AcO
OAc
AcO
AcO
O
AcO
OAc
AcO
O
NH
FmocHN
OH
O
29
10.0
5.0
ppm (f1)

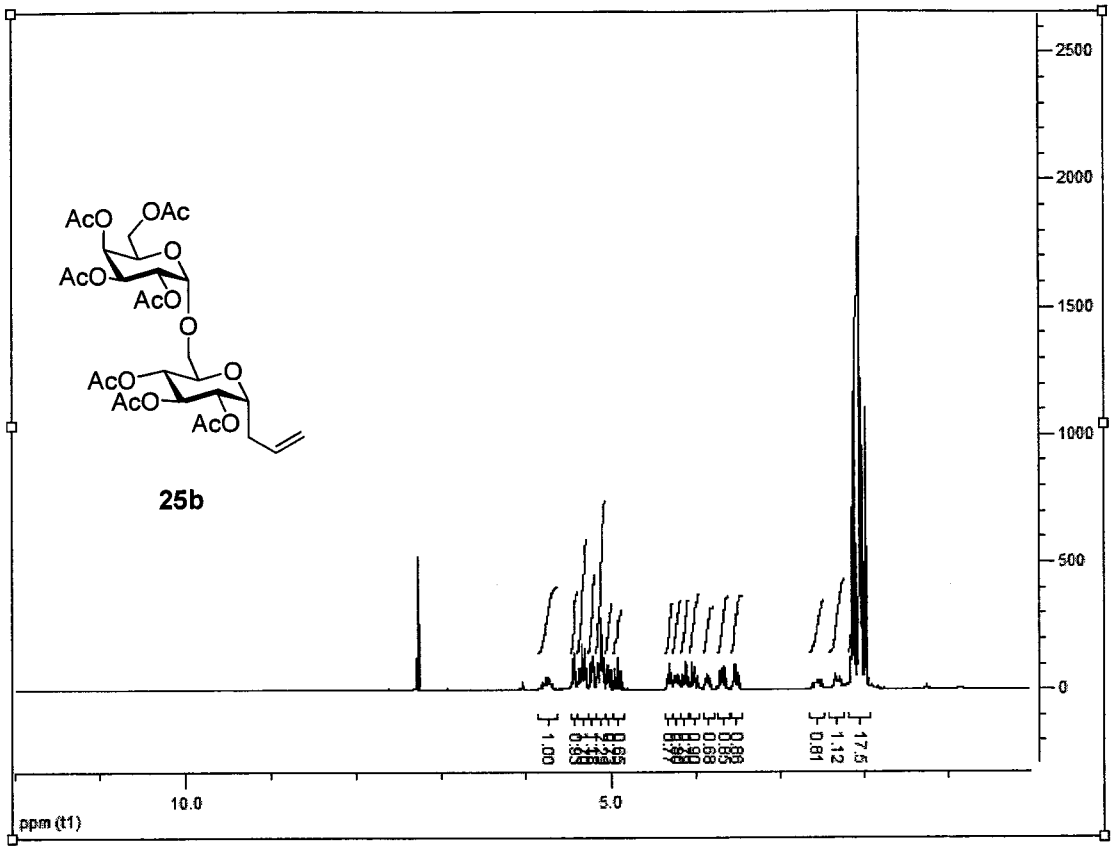
[Ornithine(lactose)-glycine-glycine]4-glycine (31)
102.788
99.867
78.526
75.240
72.759
72.381
71.992
71.495
70.812
69.980
68.397
60.885
59.944
53.712
42.209
38.575
32.494
ppm (f1)
200
150
100
50
0
0
1000
2000
3000
4000
5000
6000

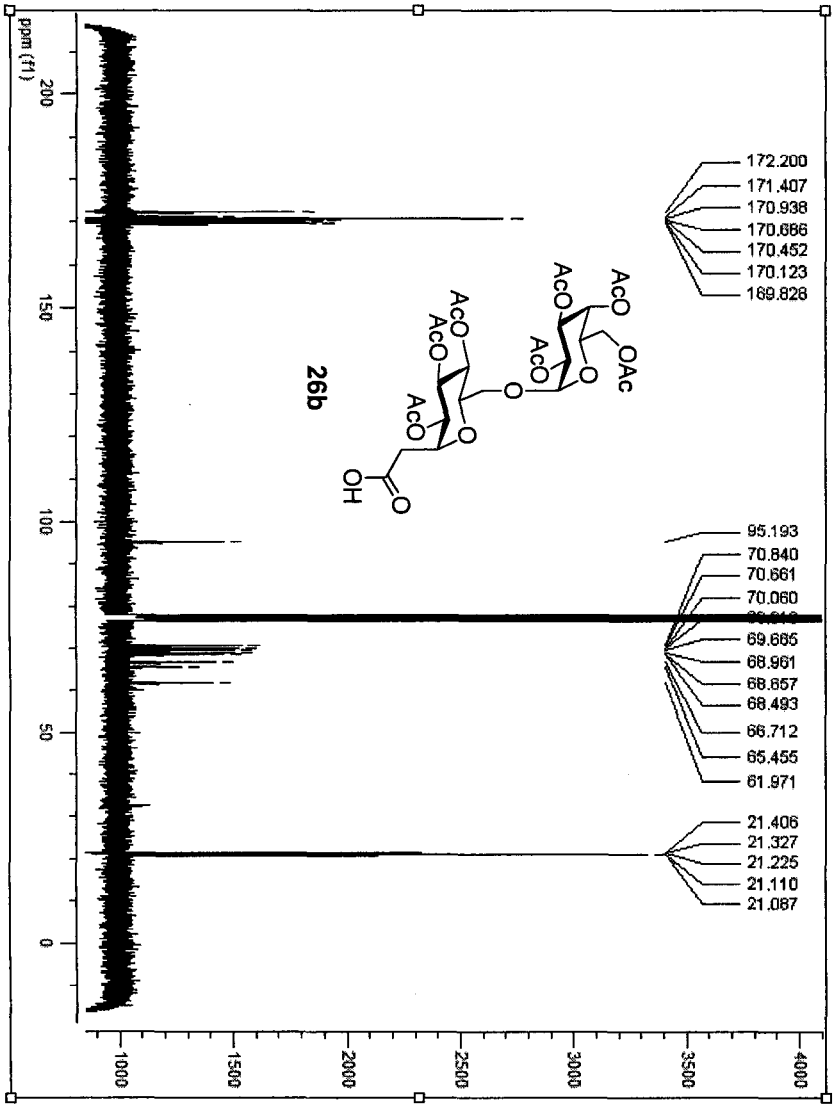

[Ornithine(lactose)-glycine-glycine]4-glycine (31)
4000
3000
2000
1000
0
1.74
1.00
5.66
7.39
1.16
1.98
2.08
1.31
1.28
2.59
5.0
4.0
3.0
2.0
1.0
ppm (f1)

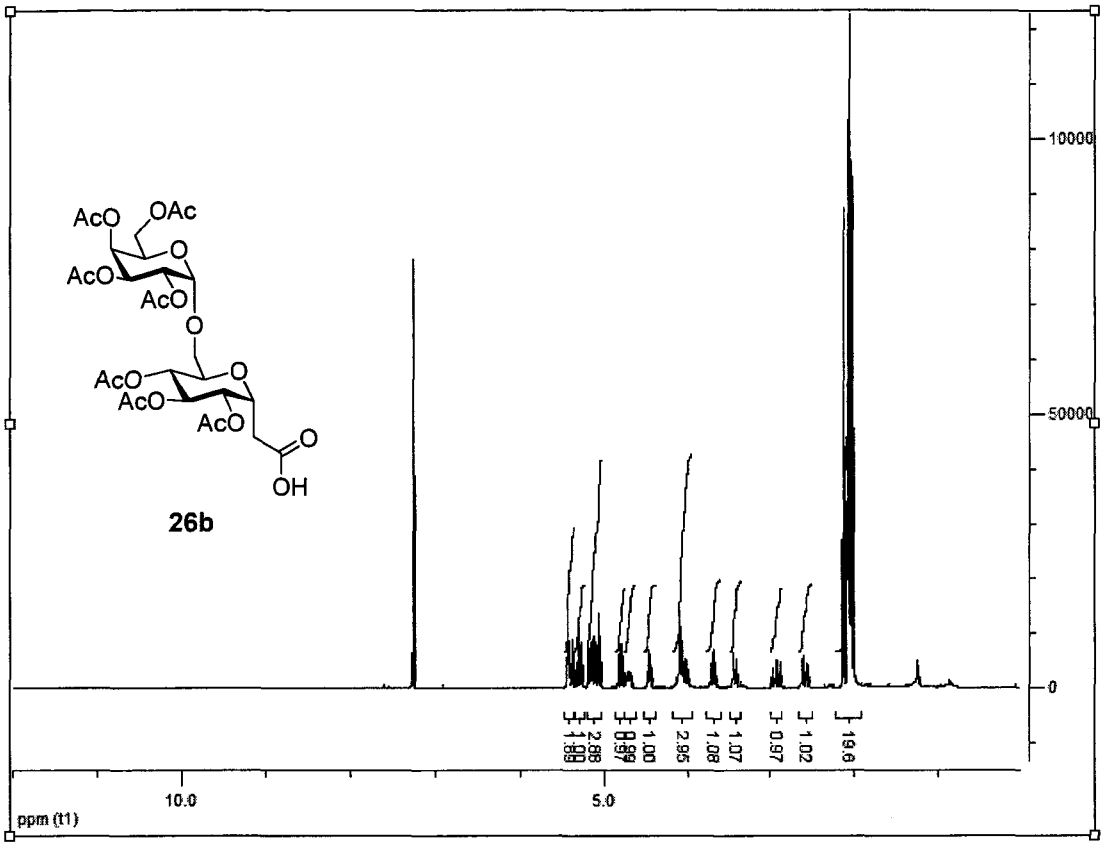

24b
170.947
170.779
170.589
170.296
170.277
170.205
169.770
96.632
86.932
73.294
70.991
70.607
68.438
68.342
67.994
67.830
66.835
65.767
62.000
21.212
21.131
21.068
21.045
21.004
ppm (t1)
200
150
100
50
0
10000
20000
30000
40000
50000

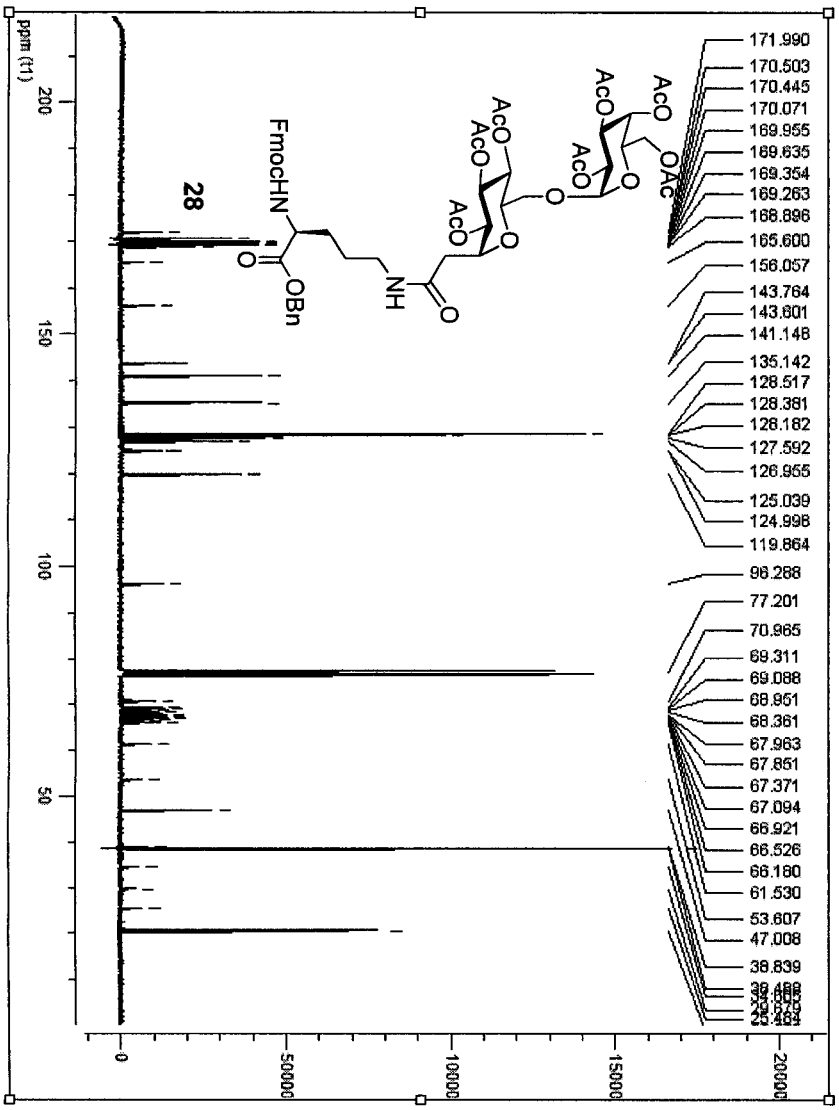

AcO
OAc
O
AcO
AcO
O
AcO
O
AcO
AcO
Br
24b
15000
10000
5000
0
1.00
0.97
2.14
1.00
0.92
22.00
10.0
5.0
ppm (t1)

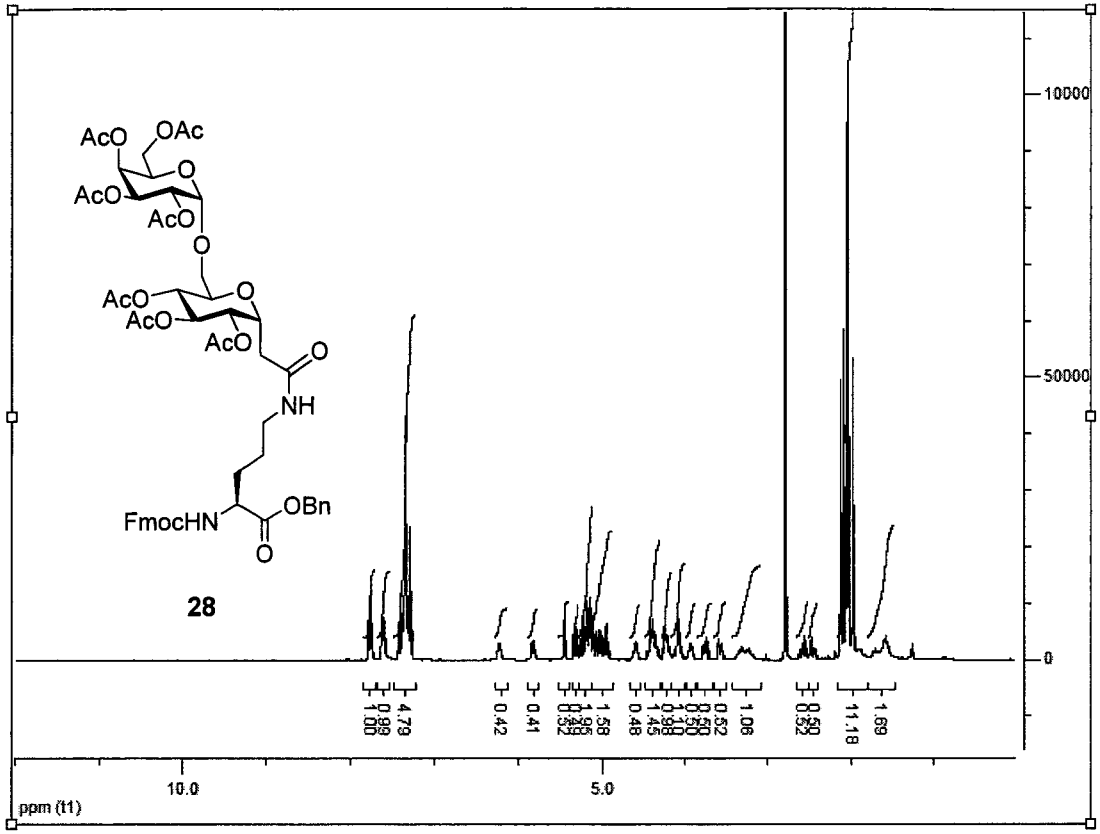

25b
AcO
AcO
AcO
AcO
AcO
AcO
OAc
O
O
O
170.445
170.322
170.160
170.078
169.805
169.592
169.476
132.842
118.073
95.831
71.739
70.319
70.269
69.479
69.323
68.099
67.967
67.360
66.475
66.270
61.692
30.365
20.735
20.648
20.585
20.569
ppm (f1)
200
150
100
50
0
50000
100000

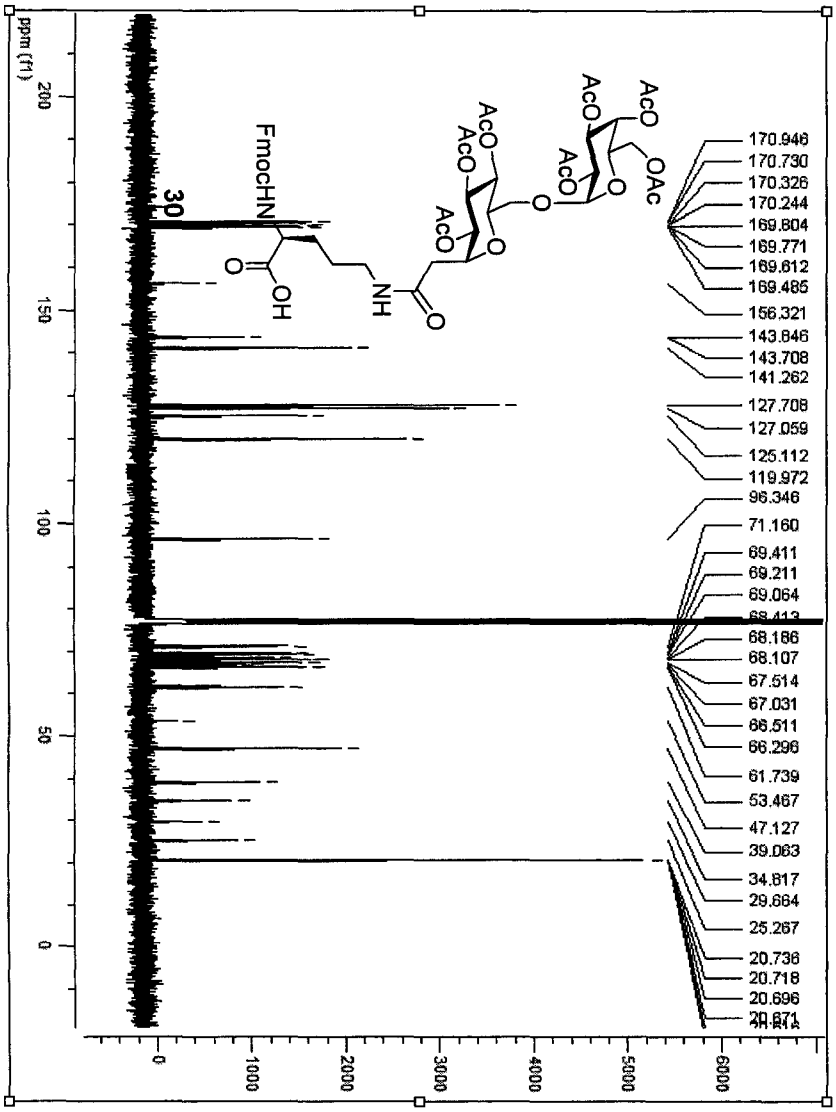

AcO
OAc
AcO
AcO
O
AcO
AcO
O
AcO
O
AcO
25b
2500
2000
1500
1000
500
0
1.00
0.81
1.12
17.5
10.0
5.0
ppm (t1)

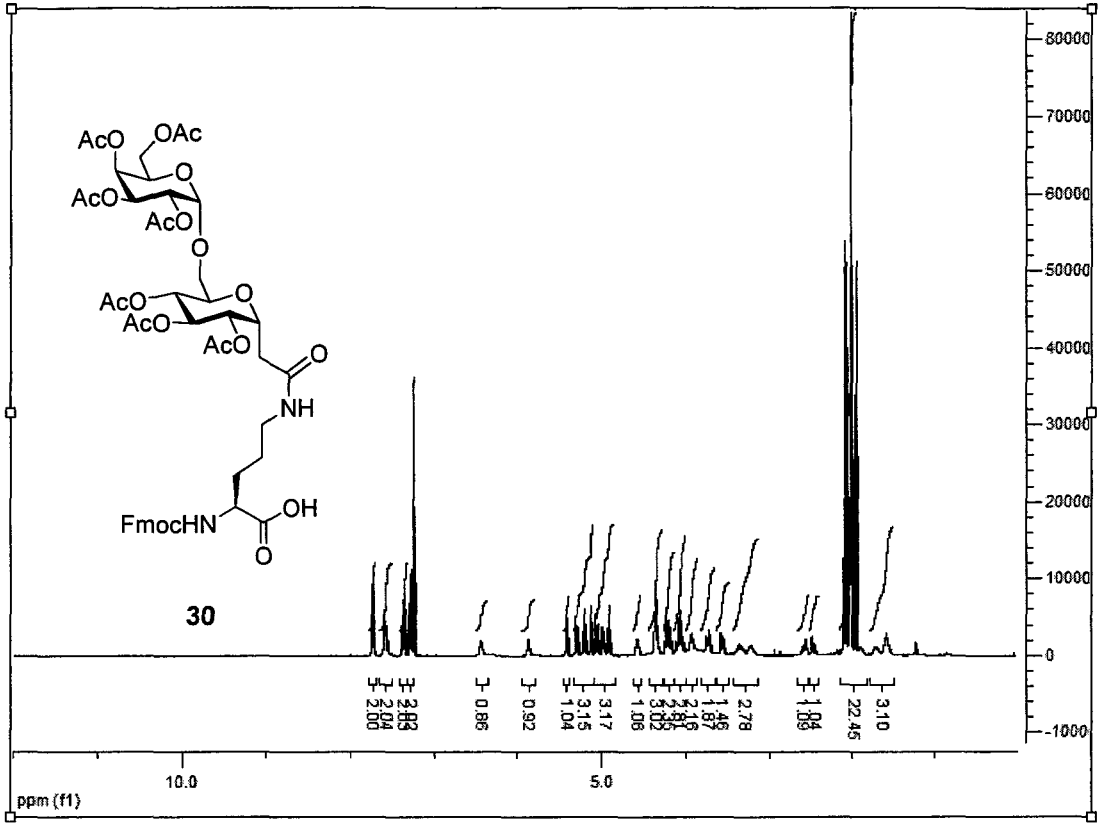

ppm (f1)
200
150
100
50
0
1000
1500
2000
2500
3000
3500
4000
172.200
171.407
170.938
170.666
170.452
170.123
169.828
95.193
70.840
70.661
70.060
69.665
68.961
68.657
68.493
66.712
65.455
61.971
21.406
21.327
21.225
21.110
21.087
26b
AcO
OAc
OH
O

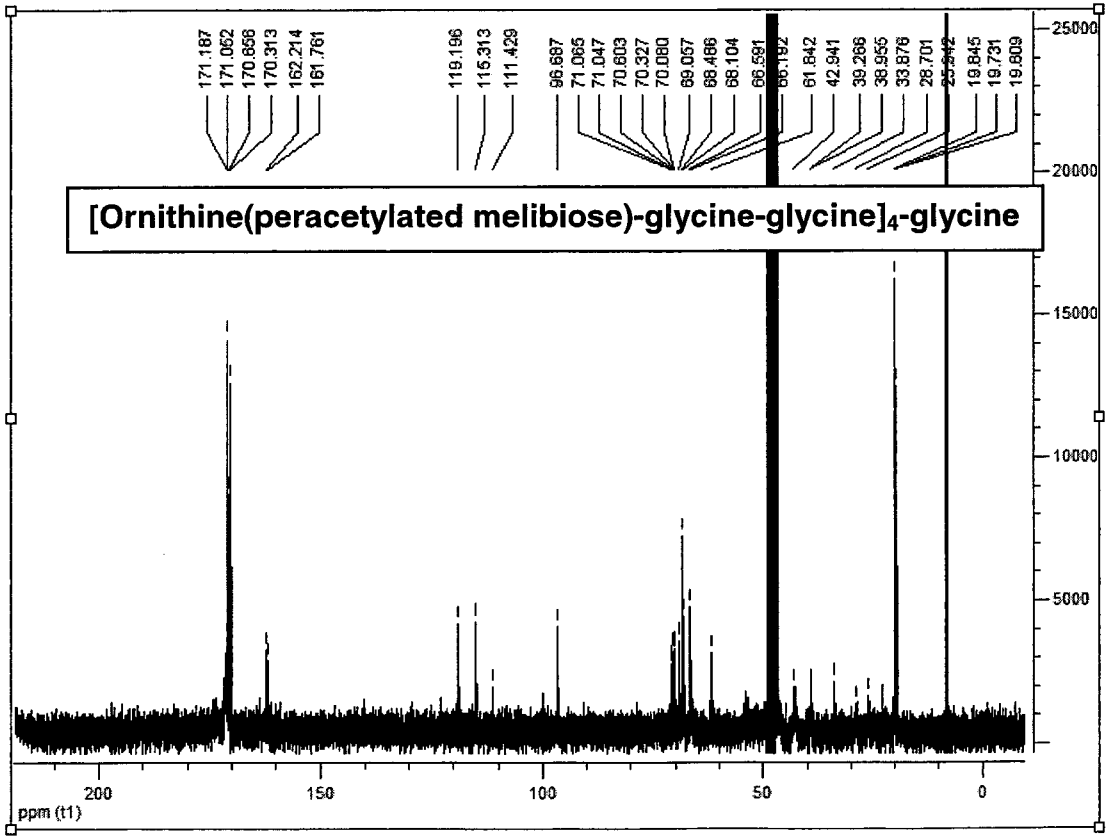

AcO
OAc
AcO
AcO
O
O
AcO
AcO
AcO
O
O
OH
26b
10000
50000
0
1.99
2.88
0.99
1.00
2.95
1.08
1.07
0.97
1.02
19.0
10.0
5.0
ppm (t1)

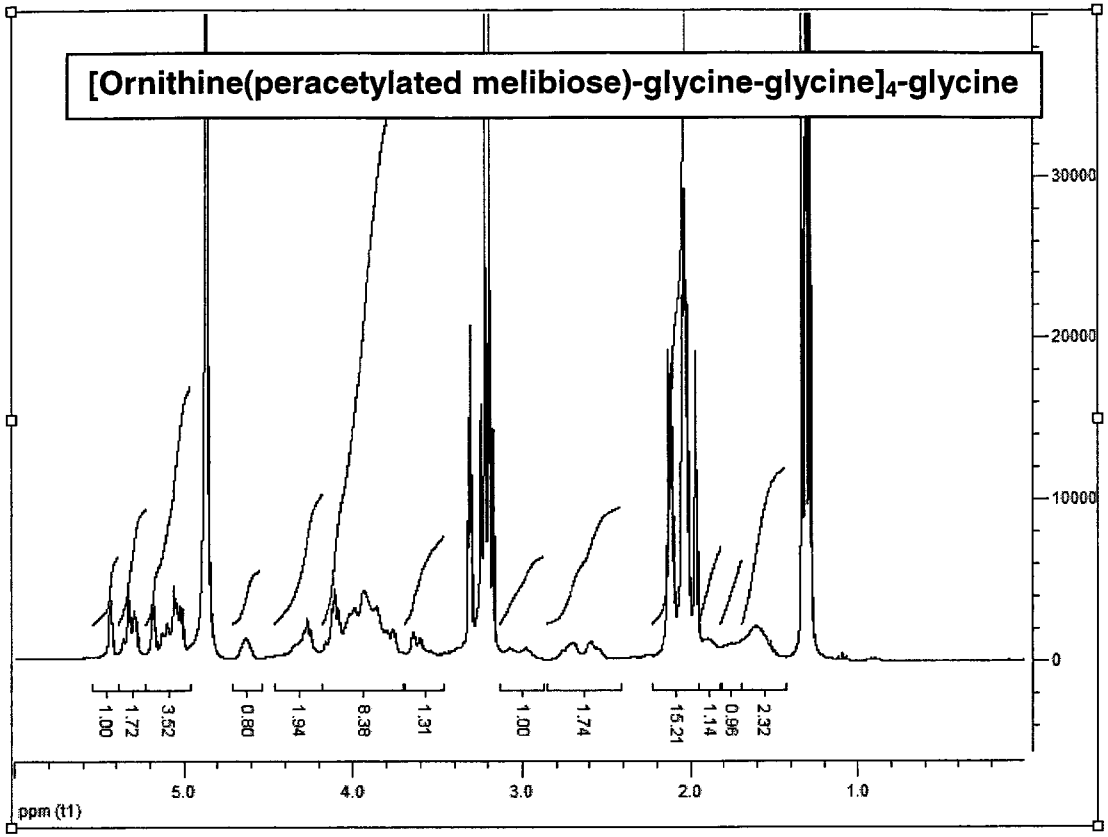

28
FmocHN
OBn
NH
AcO
OAc
171.990
170.503
170.445
170.071
169.955
169.635
169.354
169.263
168.896
165.600
156.057
143.764
143.601
141.148
135.142
128.517
128.381
128.162
127.592
126.955
125.039
124.996
119.864
98.288
77.201
70.965
69.311
69.088
68.951
68.361
67.963
67.851
67.371
67.094
66.921
66.526
66.180
61.530
53.607
47.008
38.839
ppm (f1)
200
150
100
50
0
50000
100000
150000
200000

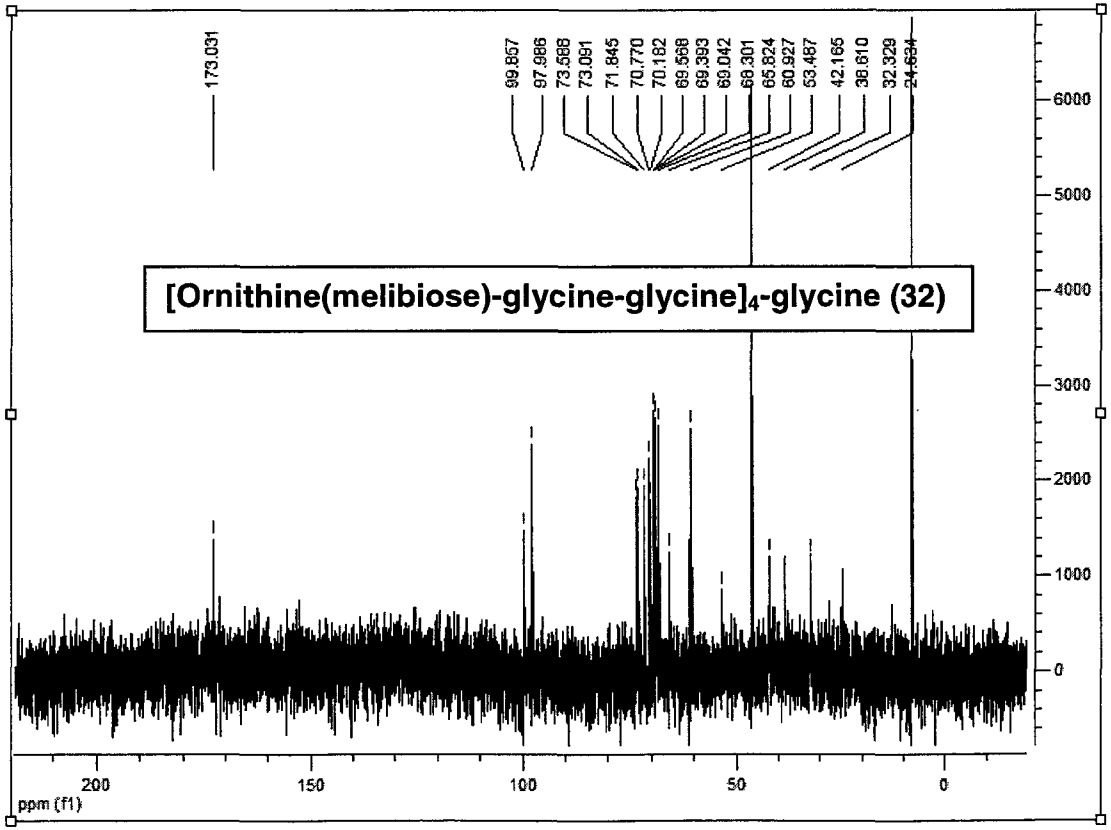

AcO
OAc
AcO
AcO
O
AcO
AcO
AcO
O
O
NH
FmocHN
OBn
O
28
10000
50000
0
4.79
0.42
0.41
1.58
0.48
1.06
11.18
1.69
10.0
5.0
ppm (t1)

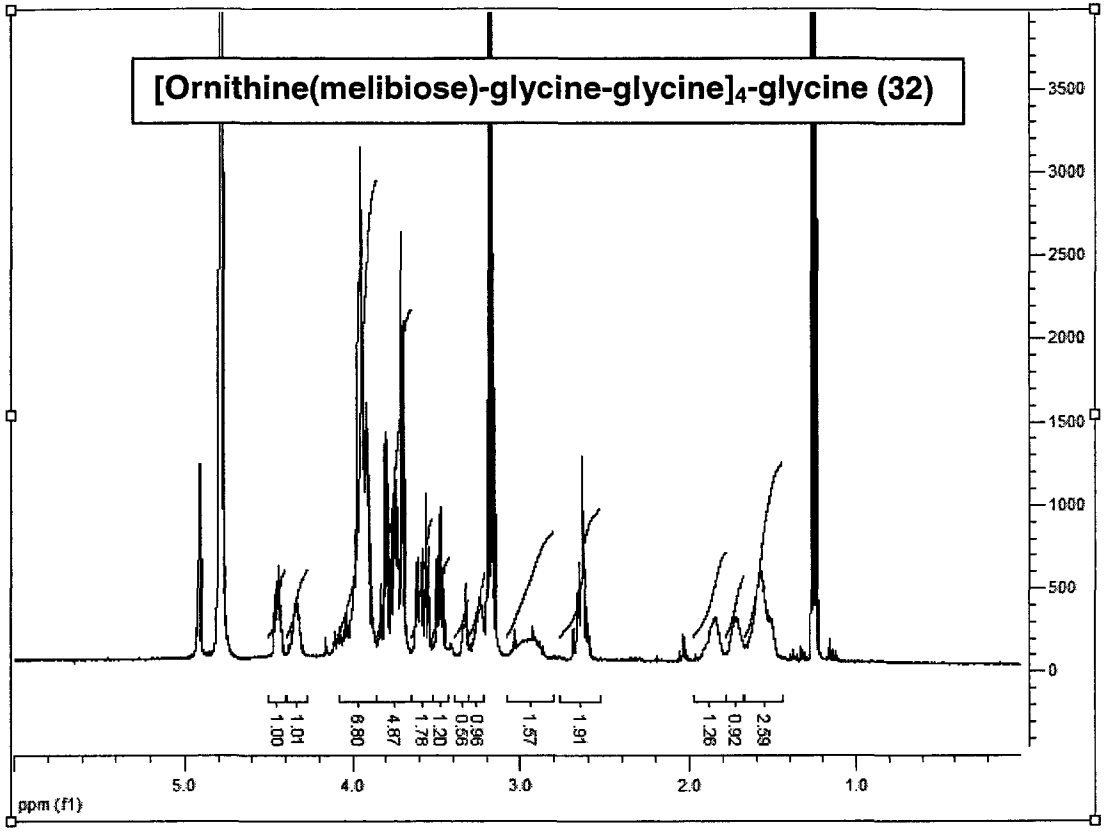

ppm (f1)
200
150
100
50
0
0
1000
2000
3000
4000
5000
6000
30
FmocHN
OH
O
NH
O
AcO
AcO
AcO
O
AcO
AcO
AcO
AcO
OAc
O
170.946
170.730
170.326
170.244
169.804
169.771
169.612
169.485
156.321
143.646
143.708
141.262
127.708
127.059
125.112
119.972
96.346
71.160
69.411
69.211
69.064
68.186
68.107
67.514
67.031
66.511
66.296
61.739
53.467
47.127
39.063
34.817
29.664
25.267
20.736
20.718
20.696

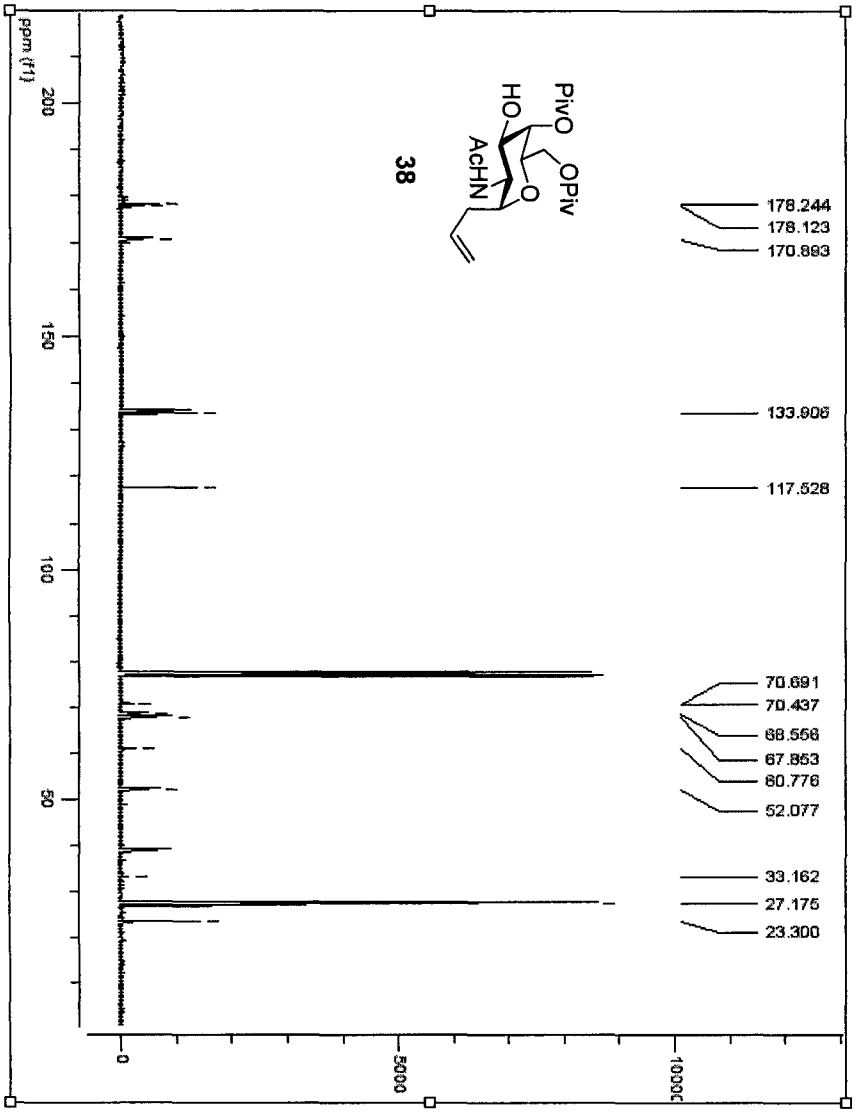

AcO
OAc
AcO
AcO
O
AcO
AcO
AcO
O
O
NH
FmocHN
OH
O
30
80000
70000
60000
50000
40000
30000
20000
10000
0
-1000
2.00
2.04
2.83
2.83
0.86
0.92
1.04
3.15
3.17
1.06
3.02
2.36
2.16
1.87
1.46
2.78
1.09
1.04
22.45
3.10
10.0
5.0
ppm (f1)

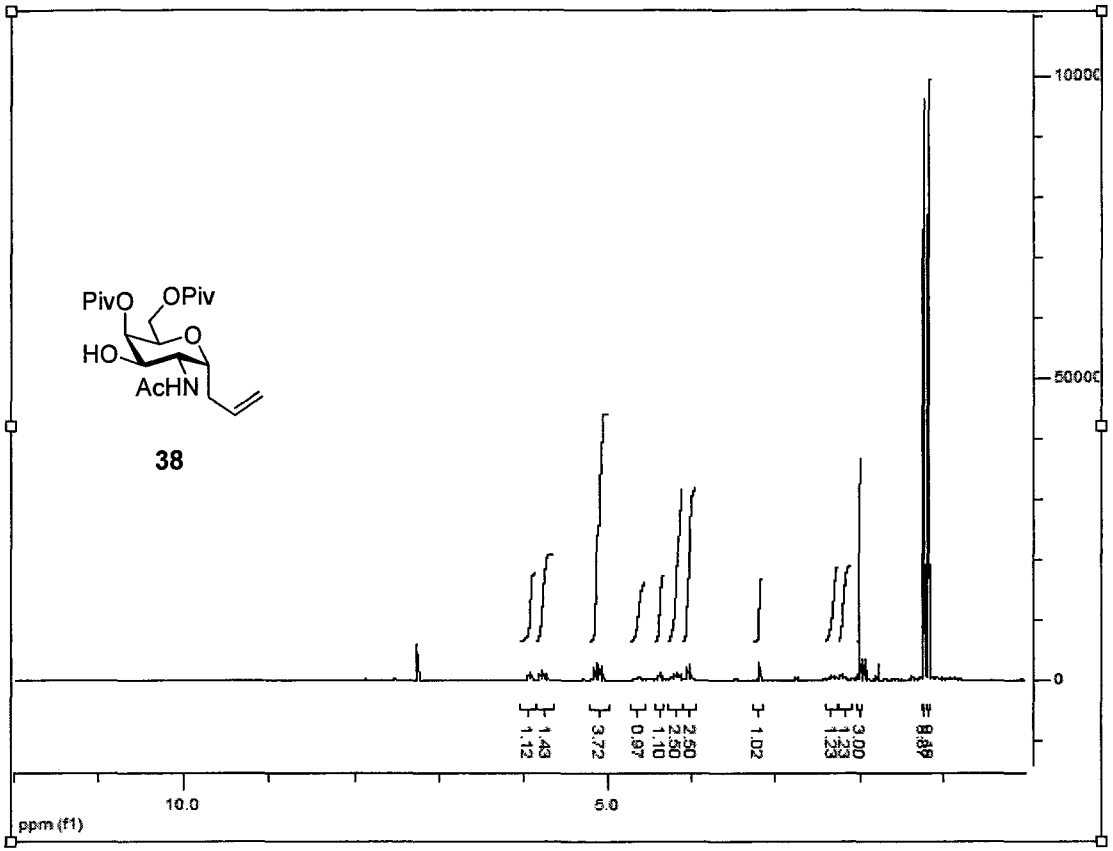

171.187
171.052
170.656
170.313
162.214
161.761
119.196
115.313
111.429
96.687
71.065
71.047
70.603
70.327
70.080
69.057
68.486
68.104
66.591
66.192
61.842
42.941
39.266
38.955
33.876
28.701
25.642
19.845
19.731
19.609
[Ornithine(peracetylated melibiose)-glycine-glycine]4-glycine
25000
20000
15000
10000
5000
200
150
100
50
0
ppm (t1)

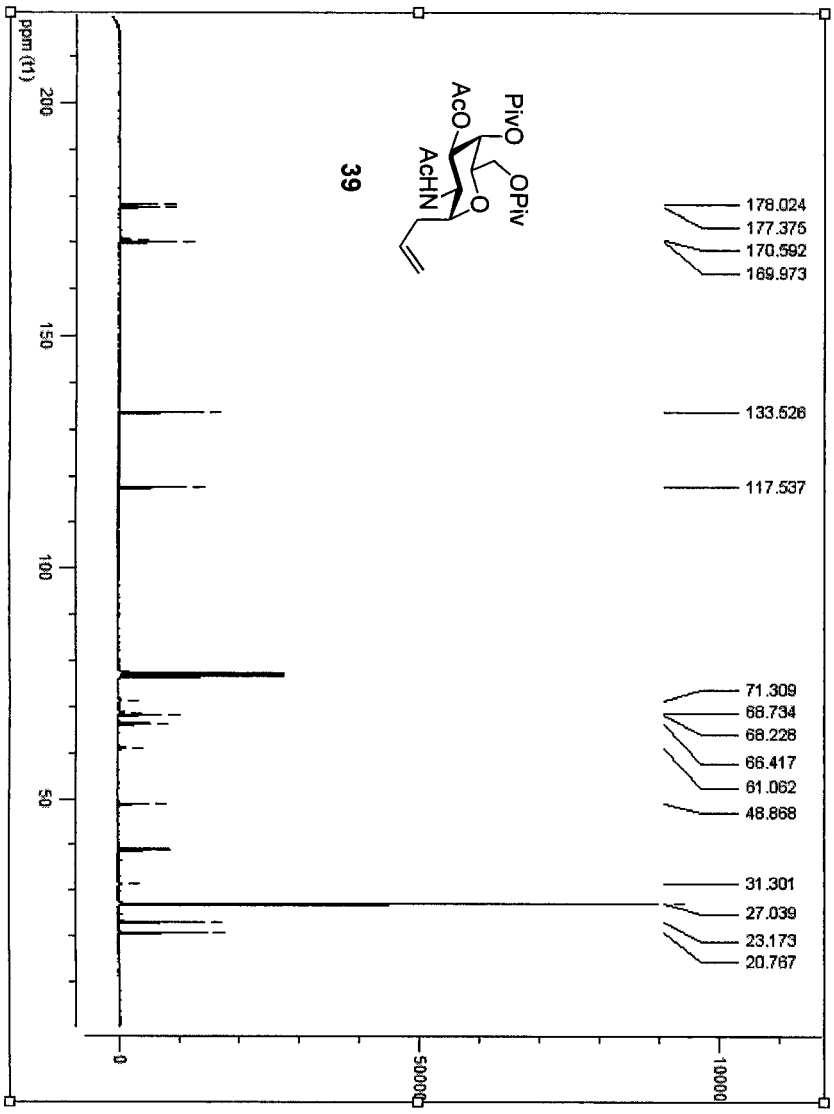

[Ornithine(peracetylated melibiose)-glycine-glycine]4-glycine
30000
20000
10000
0
1.00
1.72
3.52
0.80
1.94
8.38
1.31
1.00
1.74
15.21
1.14
0.96
2.32
5.0
4.0
3.0
2.0
1.0
ppm (t1)

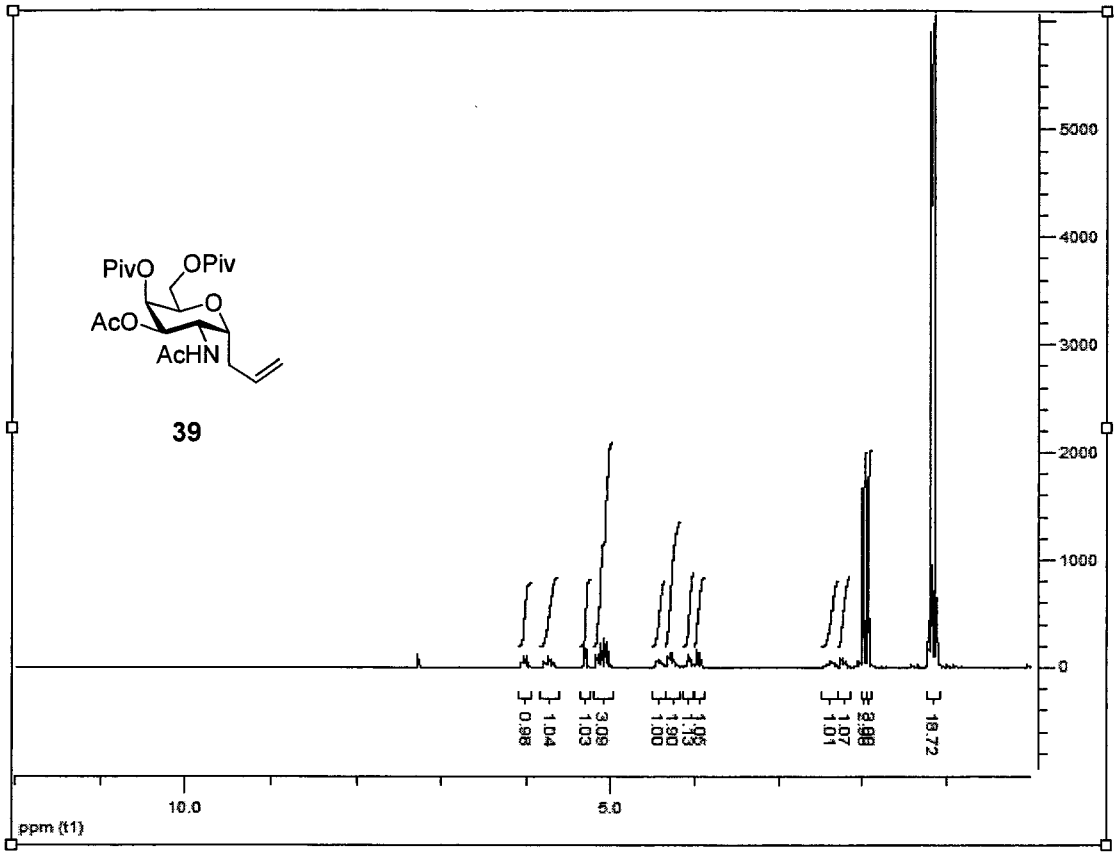

173.031
99.857
97.986
73.588
73.091
71.845
70.770
70.182
69.568
69.393
69.042
68.301
65.824
60.927
53.487
42.165
38.610
32.329
24.634
[Ornithine(melibiose)-glycine-glycine]4-glycine (32)
6000
5000
4000
3000
2000
1000
0
200
150
100
50
0
ppm (f1)

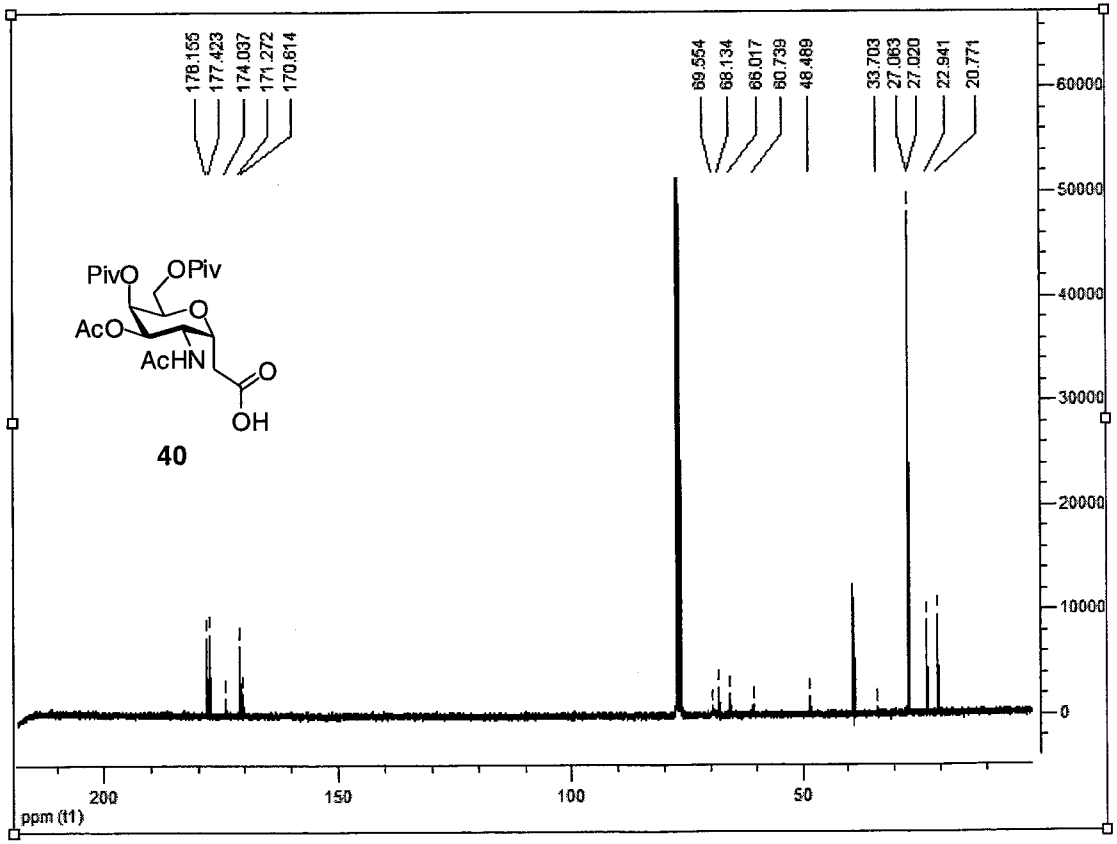

[Ornithine(melibiose)-glycine-glycine]4-glycine (32)
3500
3000
2500
2000
1500
1000
500
0
1.00
1.01
6.80
4.87
1.78
1.20
0.56
0.96
1.57
1.91
1.28
0.92
2.59
5.0
4.0
3.0
2.0
1.0
ppm (f1)

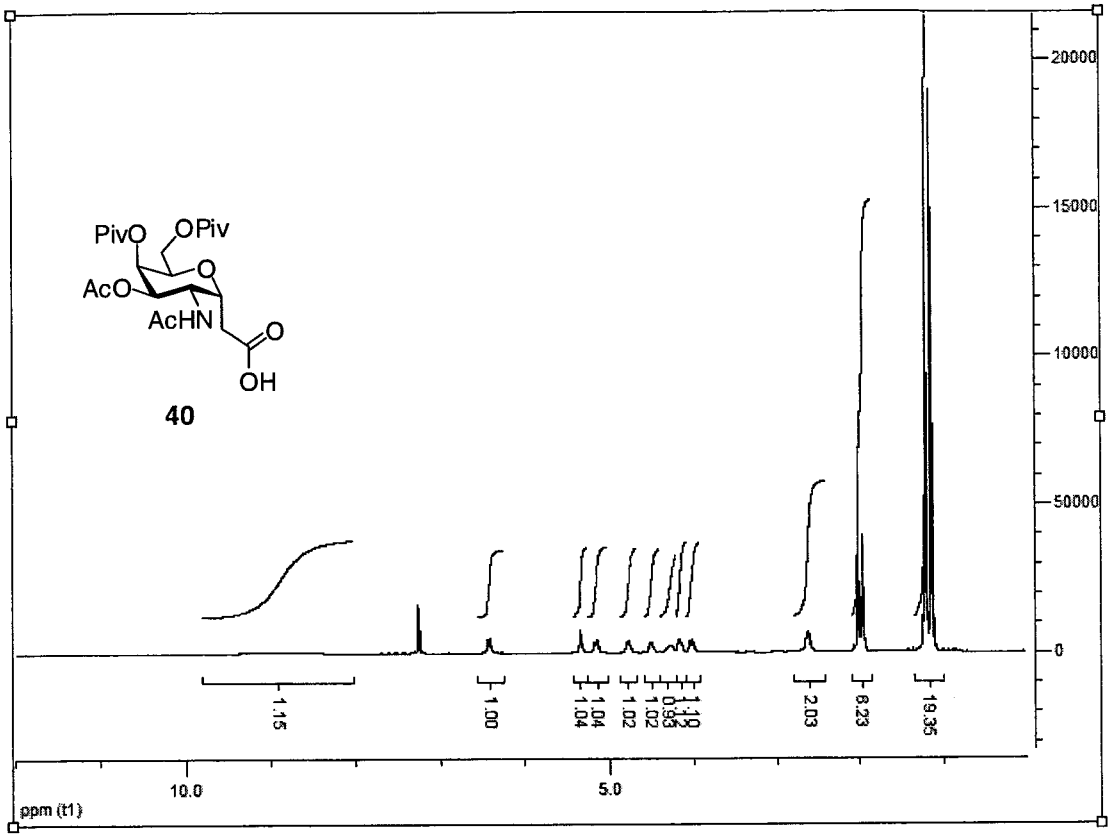

PivO
OPiv
HO
AcHN
O
38
178.244
178.123
170.893
133.908
117.528
70.691
70.437
68.556
67.853
60.776
52.077
33.162
27.175
23.300
ppm (f1)
200
150
100
50
0
5000
10000

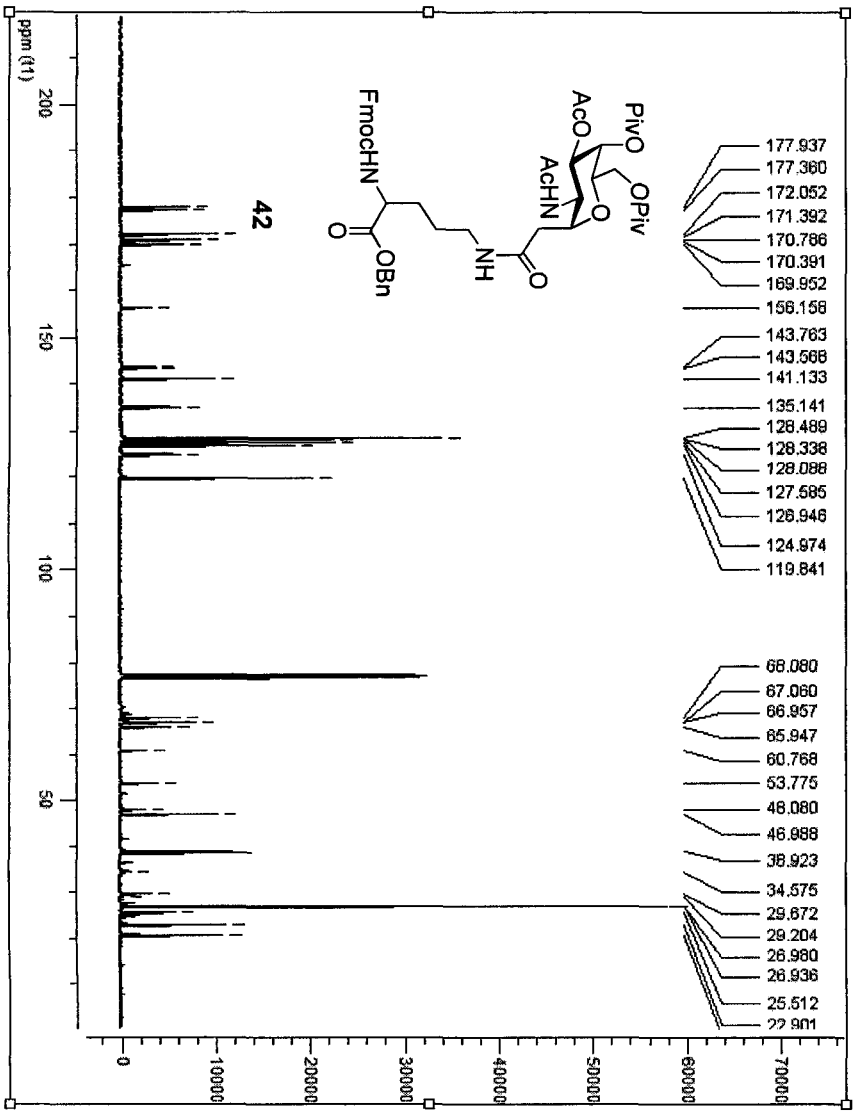

PivO
OPiv
O
HO
AcHN
38
10000
50000
0
1.12
1.43
3.72
0.97
1.10
2.50
2.50
1.02
1.23
1.23
3.00
9.89
10.0
5.0
ppm (f1)

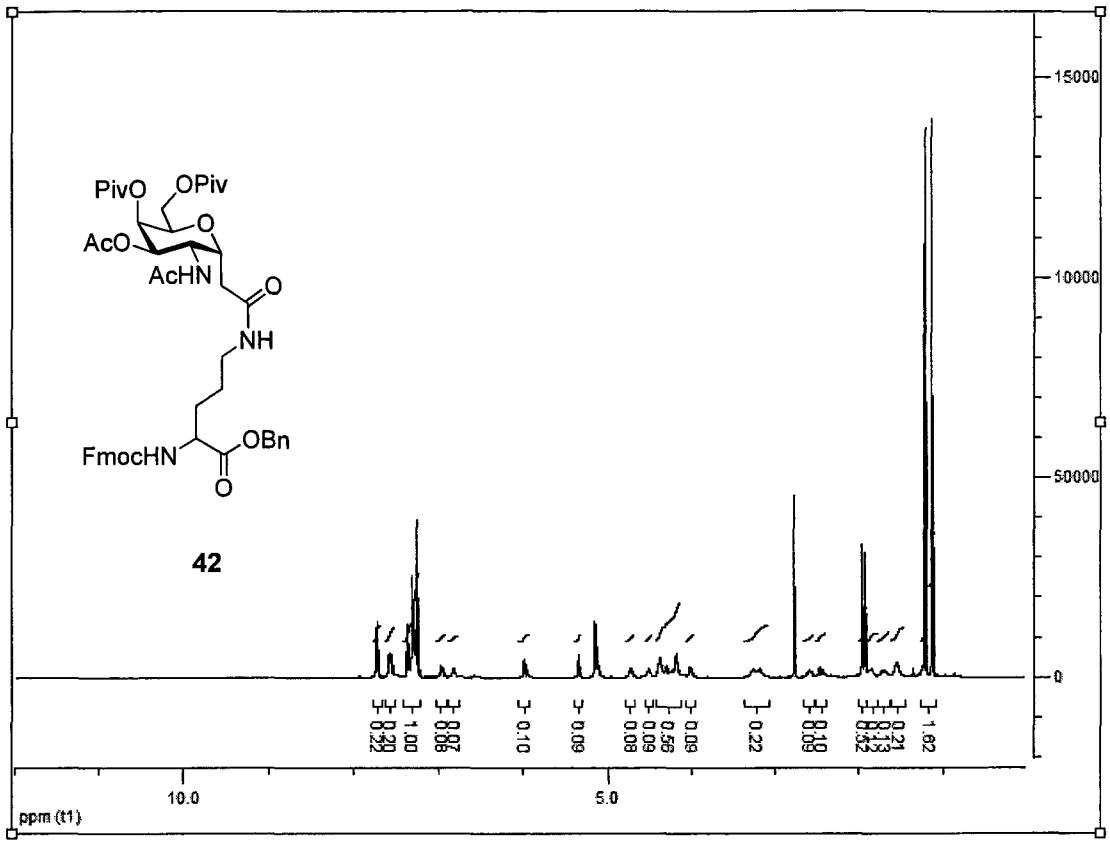

ppm (f1)
200
150
100
50
0
50000
100000
AcO
PivO
OPiv
AcHN
O
39
178.024
177.375
170.592
169.973
133.526
117.537
71.309
68.734
68.228
66.417
61.062
48.868
31.301
27.039
23.173
20.767


PivO
OPiv
O
AcO
AcHN
39
5000
4000
3000
2000
1000
0
0.98
1.04
3.09
1.03
1.00
1.07
1.01
18.72
10.0
5.0
ppm (t1)

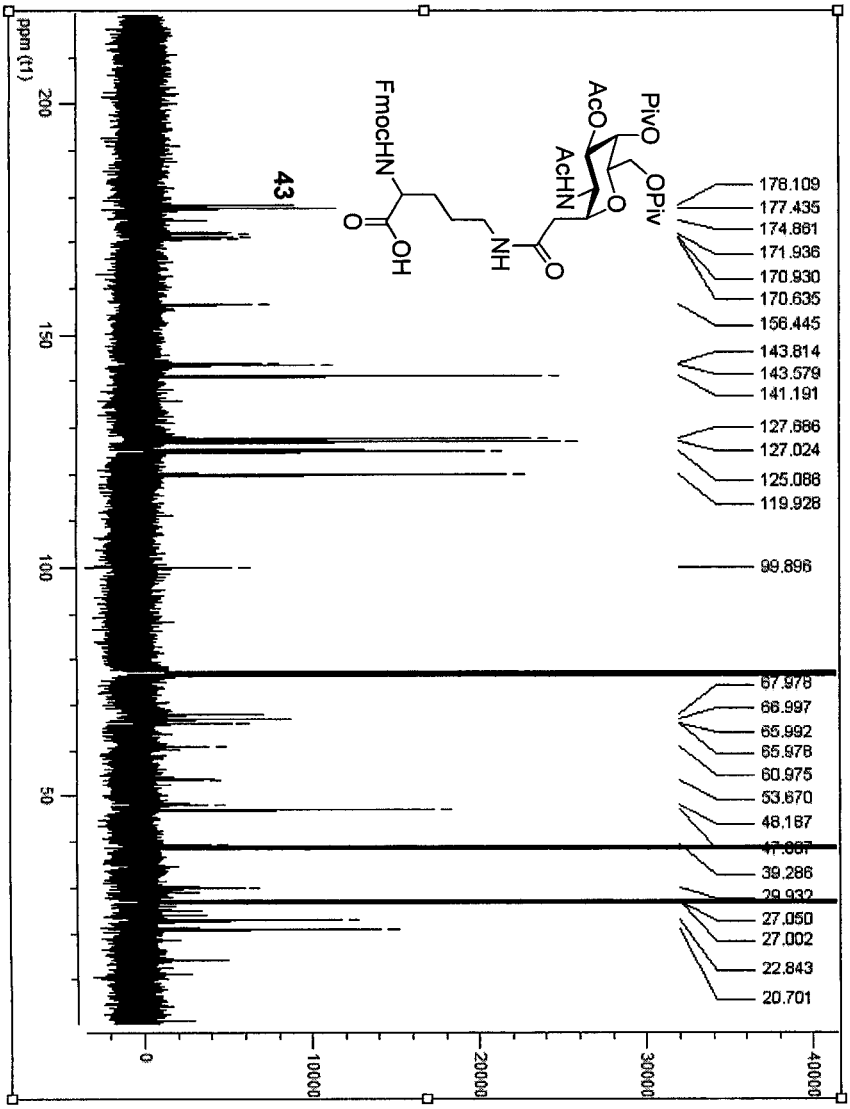

178.155
177.423
174.037
171.272
170.614
69.554
68.134
66.017
60.739
48.489
33.703
27.063
27.020
22.941
20.771
PivO
OPiv
AcO
AcHN
O
OH
40
60000
50000
40000
30000
20000
10000
0
200
150
100
50
ppm (t1)

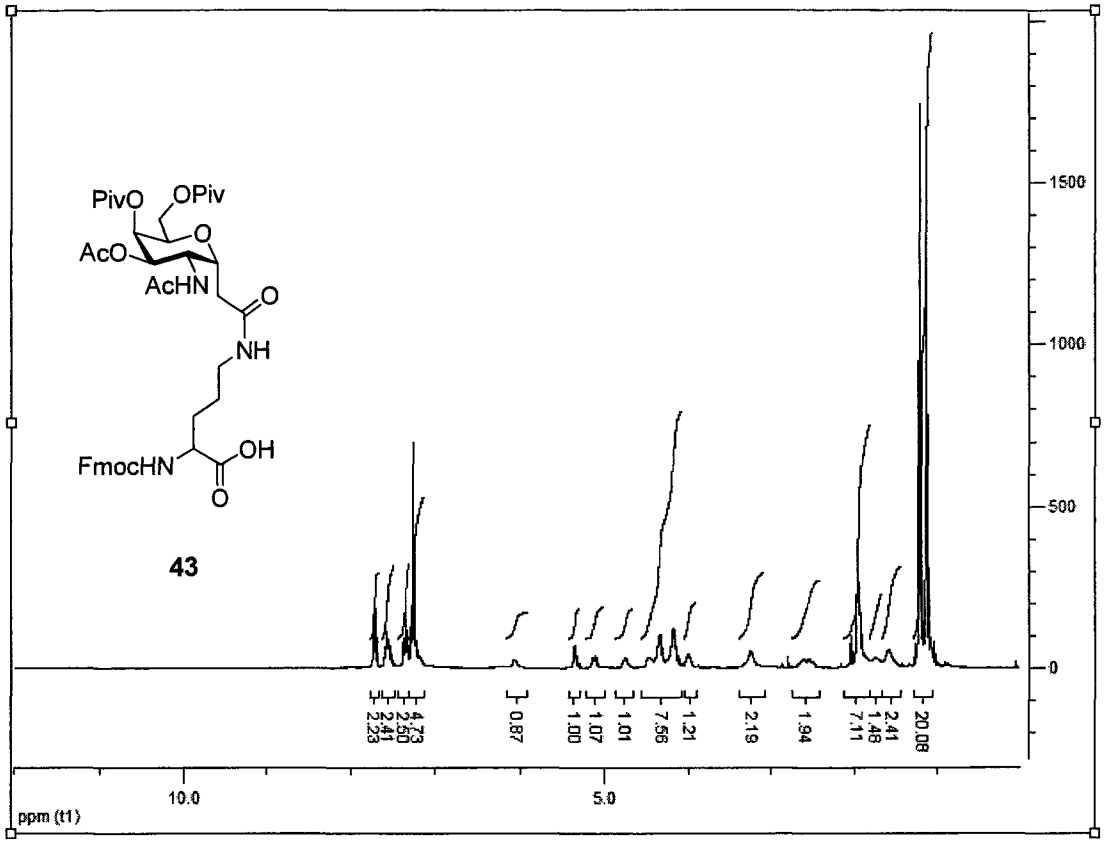

PivO
OPiv
O
AcO
AcHN
O
OH
40
20000
15000
10000
50000
0
1.15
1.00
1.04
1.04
1.02
1.02
2.03
6.23
19.35
10.0
5.0
ppm (t1)

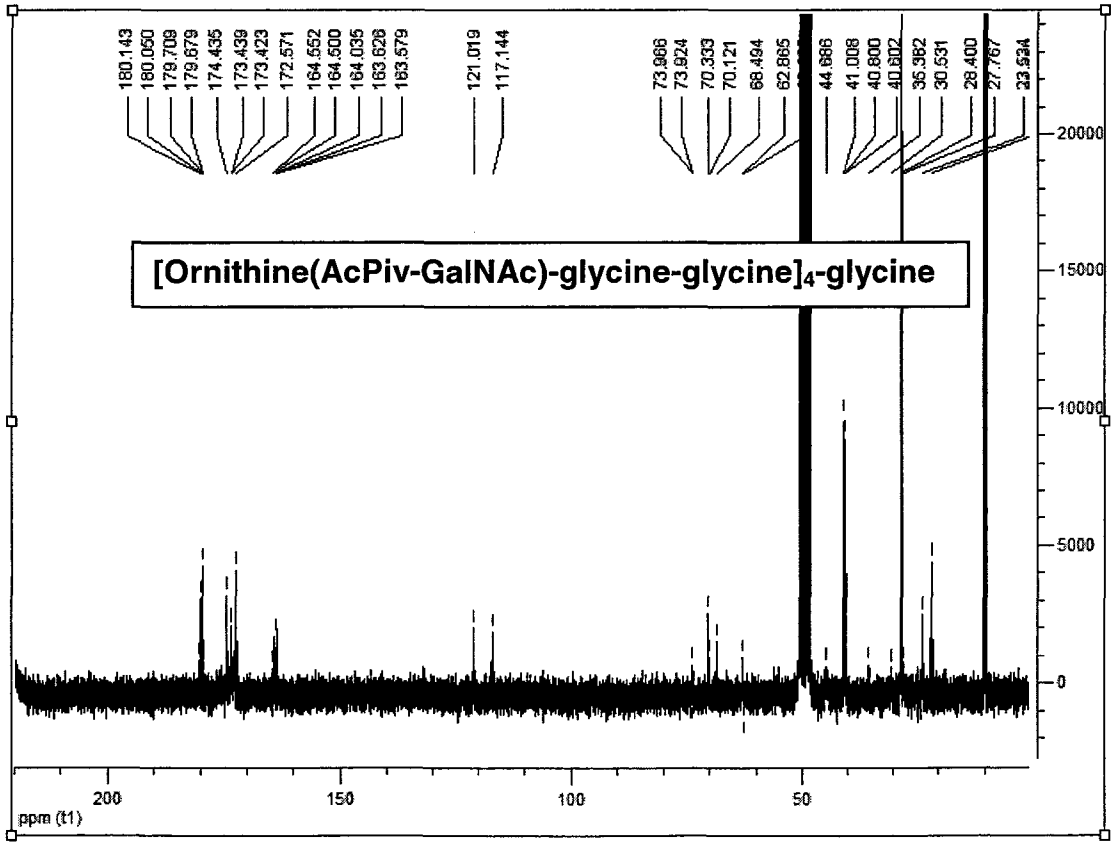

42
FmocHN
OBn
NH
AcHN
AcO
PivO
OPiv
177.937
177.360
172.052
171.392
170.786
170.391
169.952
156.156
143.763
143.566
141.133
135.141
128.489
128.338
128.088
127.585
126.946
124.974
119.841
68.080
67.060
66.957
65.947
60.768
53.775
48.080
46.988
38.923
34.575
29.672
29.204
26.980
26.936
25.512
22.901
ppm (t1)
200
150
100
50
0
10000
20000
30000
40000
50000
60000
70000

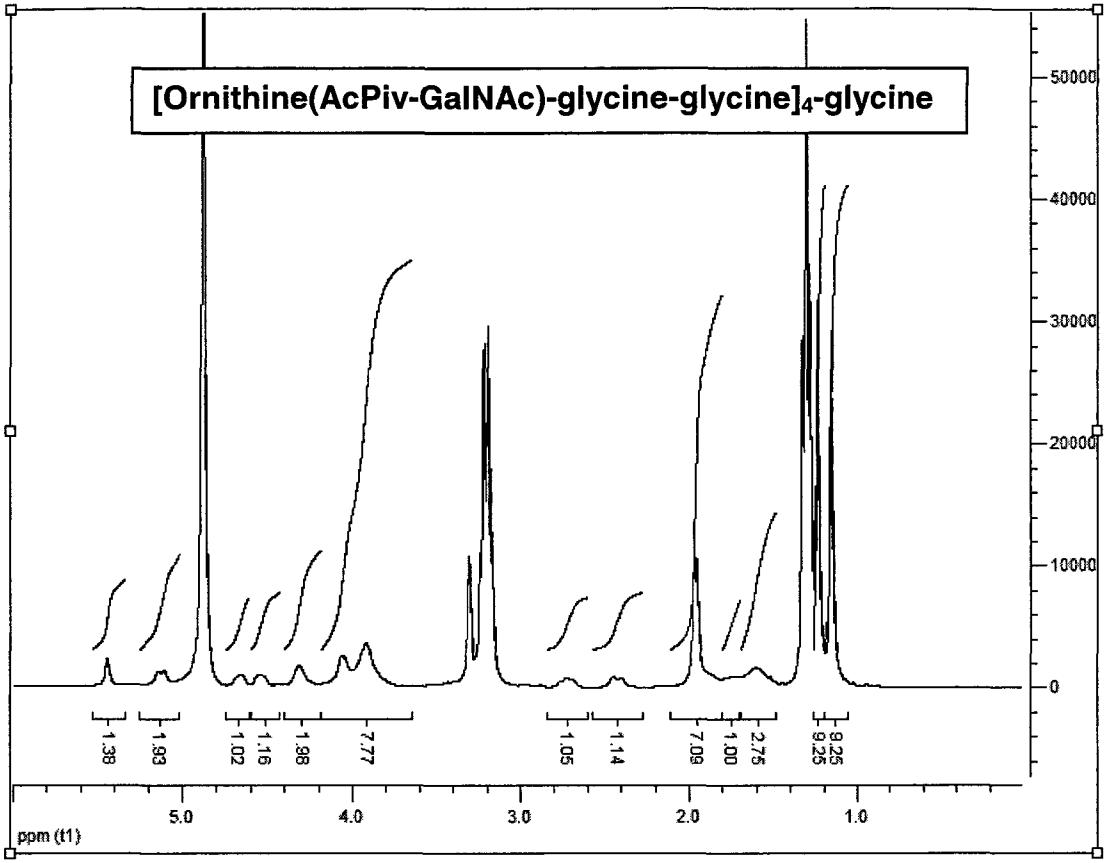

PivO
OPiv
AcO
AcHN
O
NH
FmocHN
OBn
O
15000
10000
50000
0
10.0
5.0
ppm (t1)

42

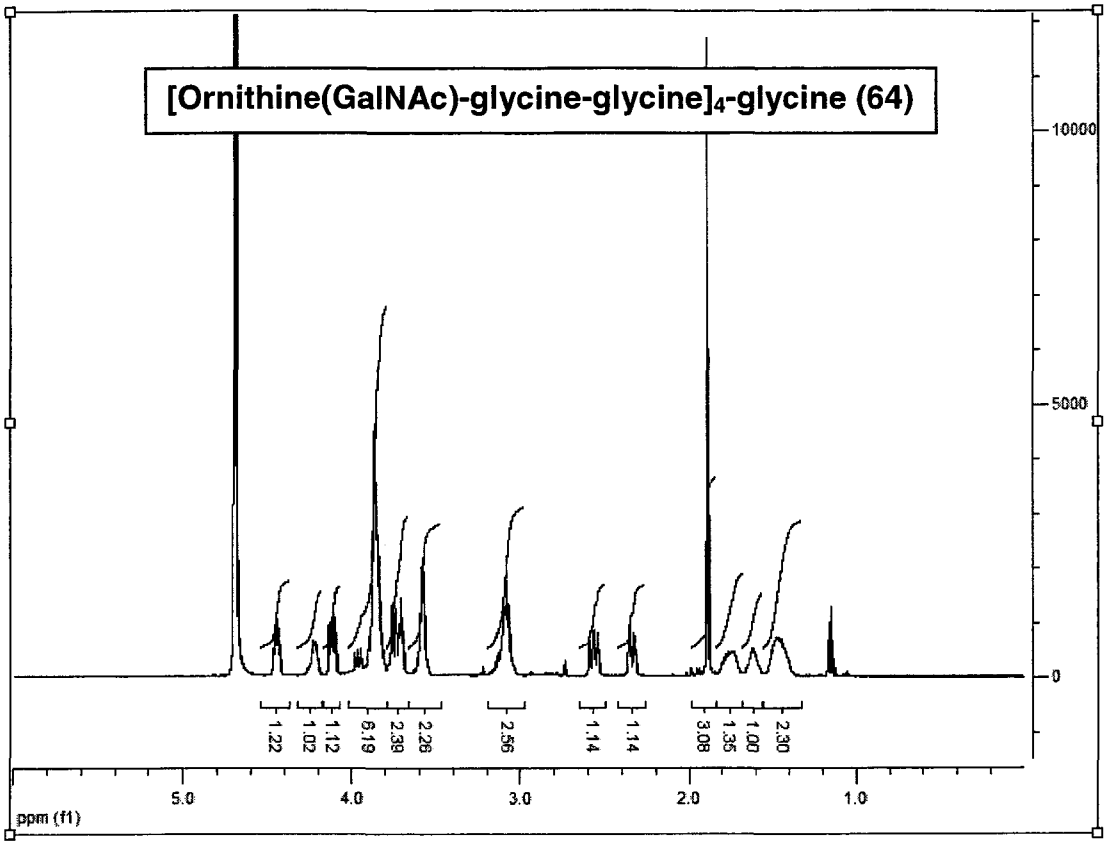

43
FmocHN
O
OH
NH
O
AcHN
AcO
PivO
OPiv
O
178.109
177.435
174.861
171.936
170.930
170.635
156.445
143.814
143.579
141.191
127.886
127.024
125.086
119.928
99.896
67.978
66.997
65.992
65.976
60.975
53.670
48.187
47.867
39.286
28.932
27.050
27.002
22.843
20.701
ppm (f1)
200
150
100
50
0
10000
20000
30000
40000

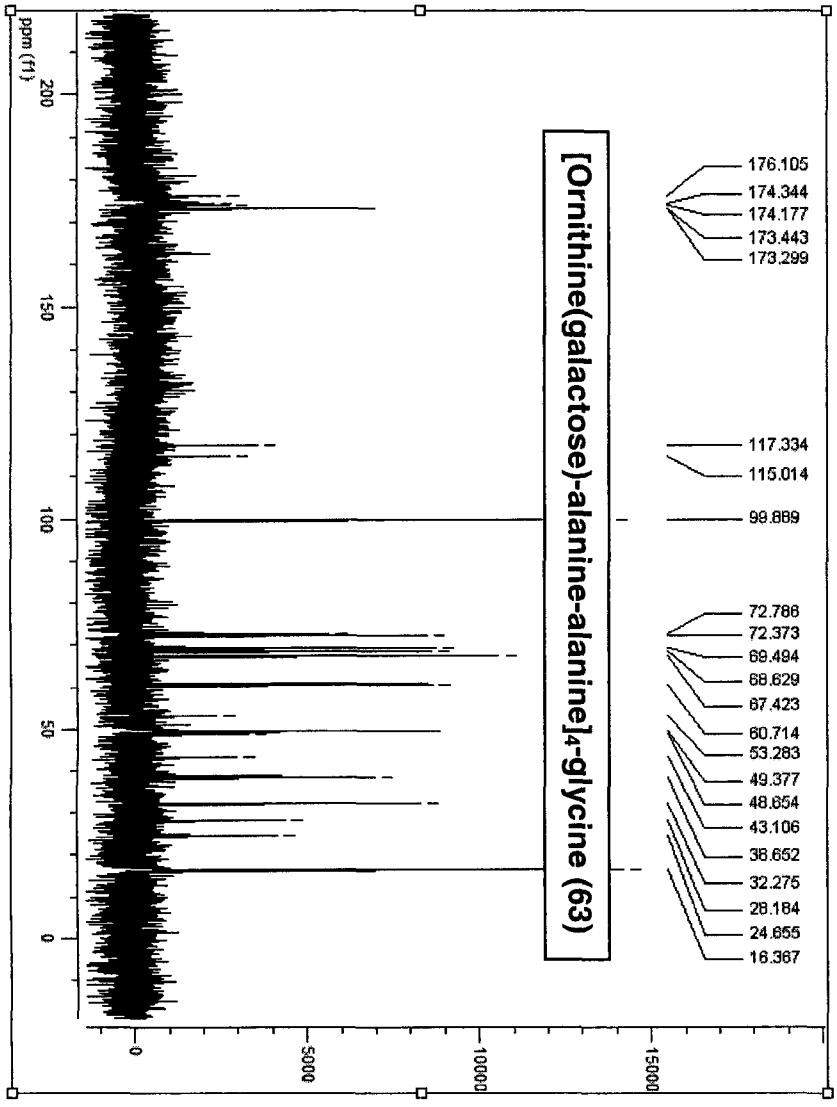

PivO
OPiv
AcO
AcHN
O
NH
FmocHN
OH
O
43
1500
1000
500
0
2.23
2.41
2.50
4.73
0.87
1.00
1.07
1.01
7.56
1.21
2.19
1.94
7.11
1.48
2.41
20.08
10.0
5.0
ppm (t1)

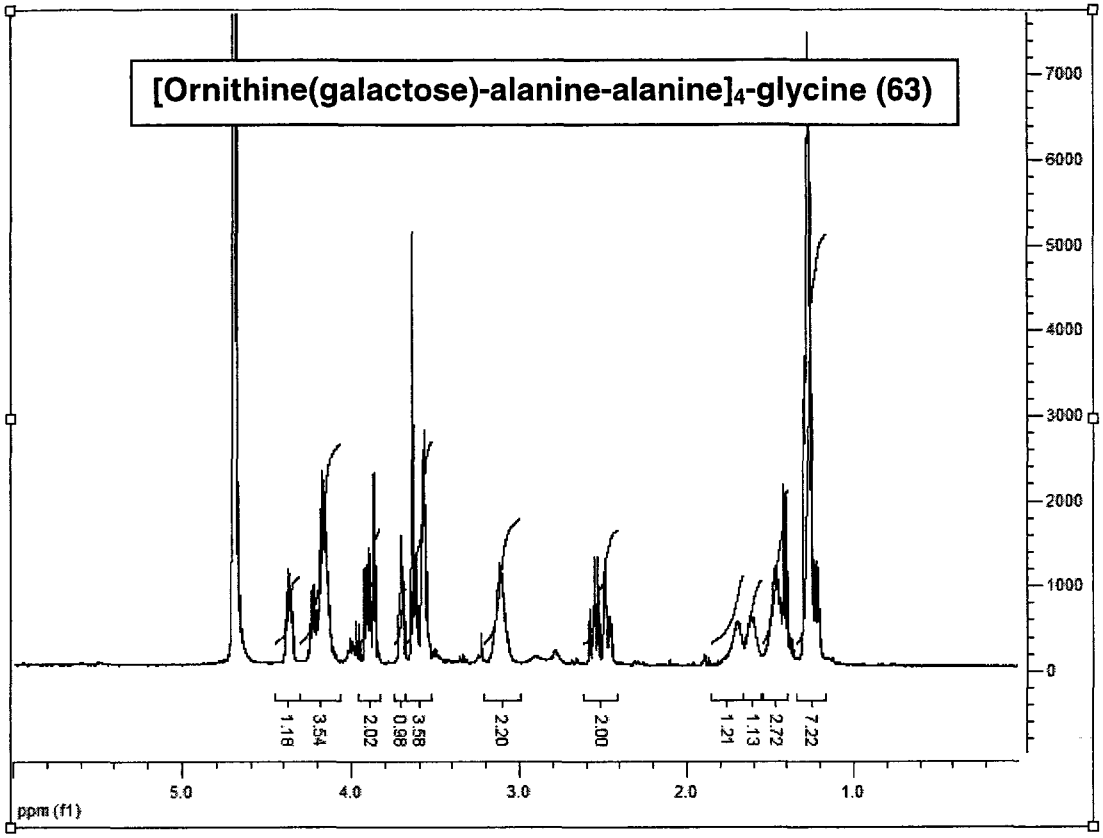

180.143
180.050
179.709
179.679
174.435
173.439
173.423
172.571
164.552
164.500
164.035
163.628
163.579
121.019
117.144
73.966
73.924
70.333
70.121
68.494
62.865
44.686
41.008
40.800
40.602
35.382
30.531
28.400
27.767
23.934
[Ornithine(AcPiv-GalNAc)-glycine-glycine]4-glycine
20000
15000
10000
5000
0
200
150
100
50
ppm (t1)

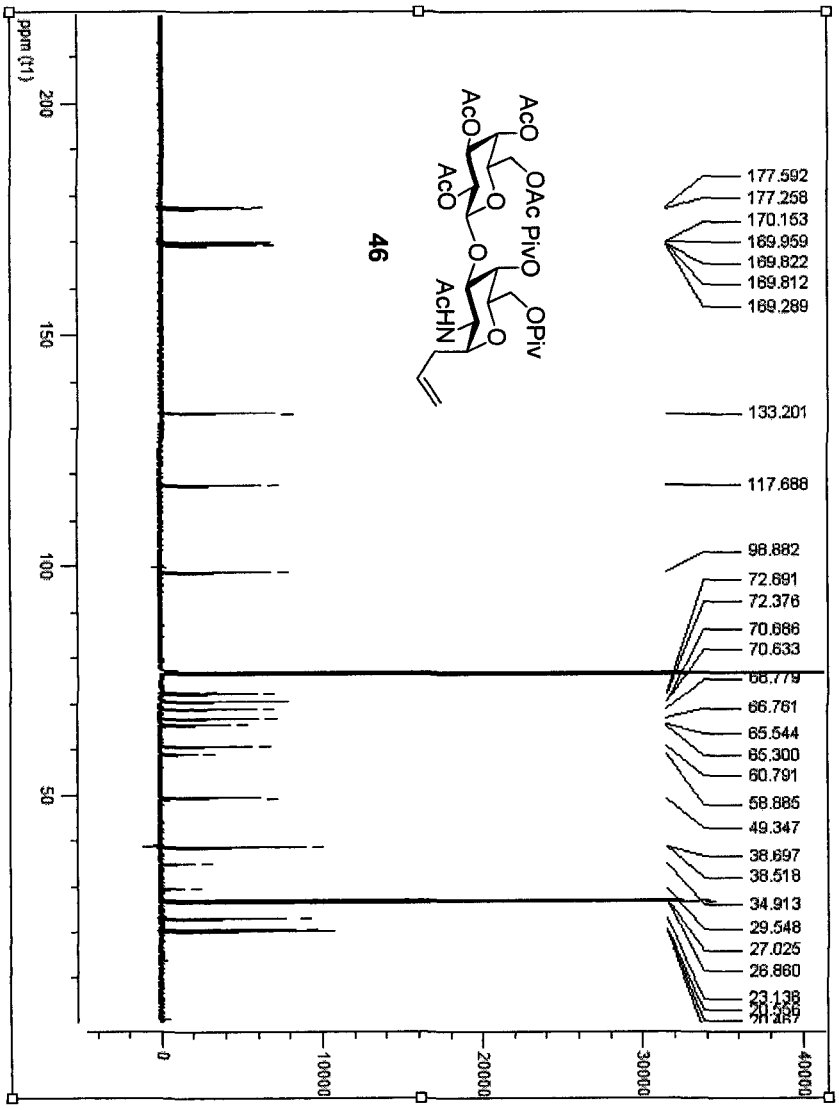

[Ornithine(AcPiv-GalNAc)-glycine-glycine]4-glycine
50000
40000
30000
20000
10000
0
1.38
1.93
1.02
1.16
1.98
7.77
1.05
1.14
7.09
1.00
2.75
9.25
8.25
5.0
4.0
3.0
2.0
1.0
ppm (t1)

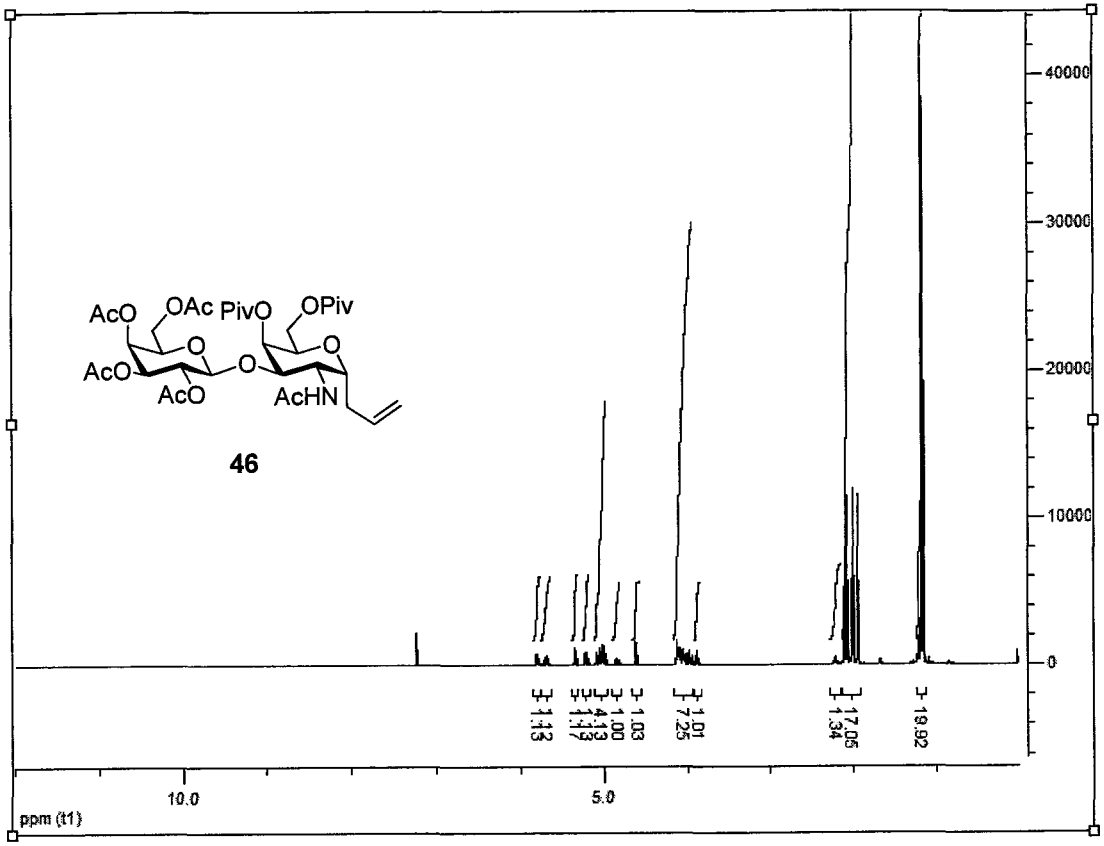
[Ornithine(GalNAc)-glycine-glycine]4-glycine (64)
ppm (f1)
200
150
100
50
0
5000
0
174.484
172.861
172.667
171.807
171.485
170.193
162.730
99.861
72.373
70.886
68.014
67.050
60.788
53.579
52.734
49.176
46.521
42.419
42.219
38.650
32.955
27.981
24.714
21.766

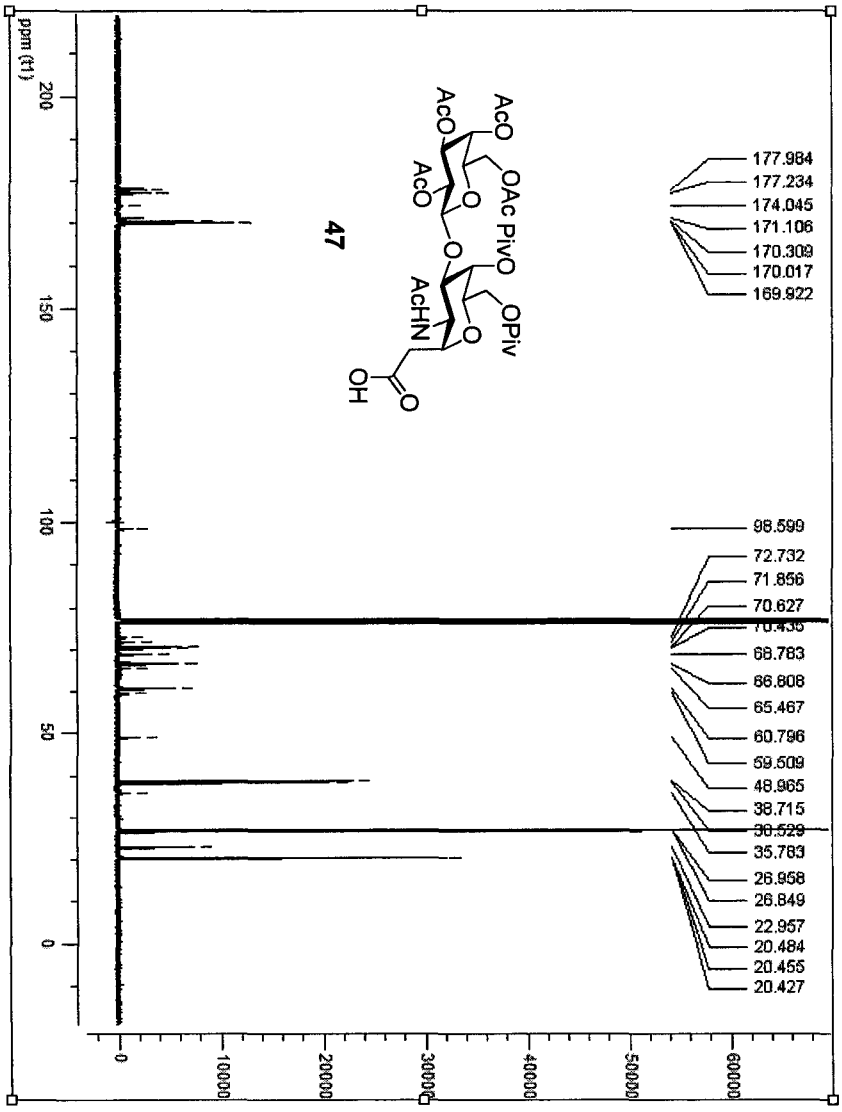

[Ornithine(GalNAc)-glycine-glycine]4-glycine (64)
10000
5000
0
1.22
1.02
1.12
6.19
2.39
2.26
2.56
1.14
1.14
3.08
1.35
1.00
2.30
5.0
4.0
3.0
2.0
1.0
ppm (f1)

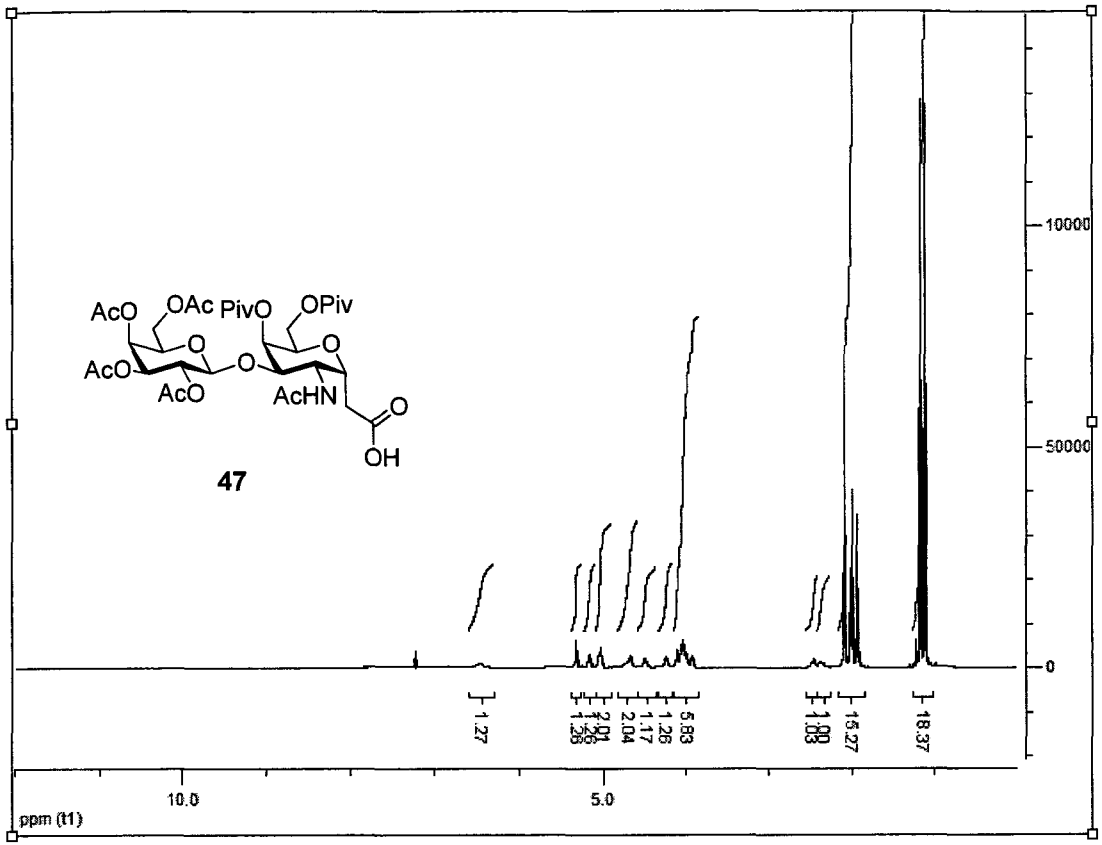

[Ornithine(galactose)-alanine-alanine]4-glycine (63)
176.105
174.344
174.177
173.443
173.299
117.334
115.014
99.889
72.786
72.373
69.494
68.629
67.423
60.714
53.283
49.377
48.654
43.106
38.652
32.275
28.184
24.655
16.367
ppm (f1)
200
150
100
50
0
0
5000
10000
15000

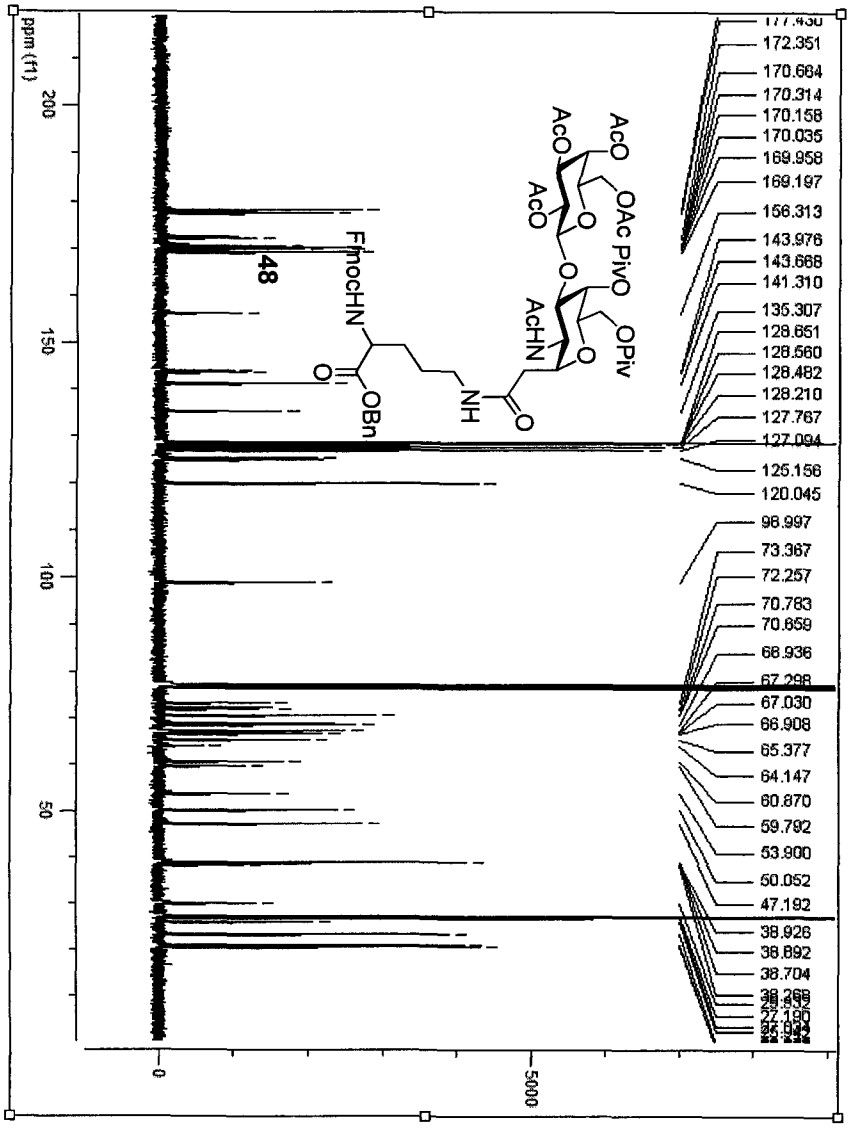

[Ornithine(galactose)-alanine-alanine]4-glycine (63)
7000
6000
5000
4000
3000
2000
1000
0
1.18
3.54
2.02
0.98
3.58
2.20
2.00
1.21
1.13
2.72
7.22
5.0
4.0
3.0
2.0
1.0
ppm (f1)

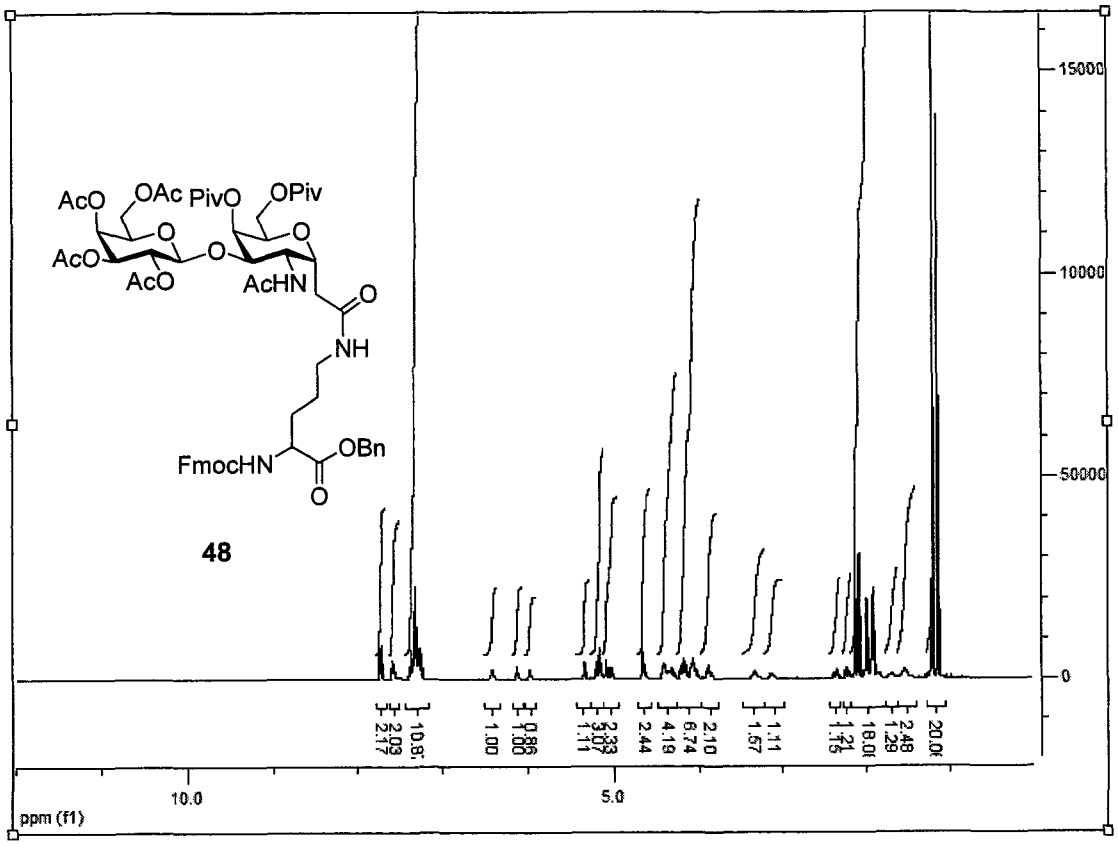

AcO
OAc
AcO
AcO
O
O
PivO
OPiv
AcHN
O
46
177.592
177.258
170.153
169.959
169.822
169.812
169.289
133.201
117.688
98.882
72.691
72.376
70.686
70.633
66.761
65.544
65.300
60.791
58.885
49.347
38.697
38.518
34.913
29.548
27.025
26.860
ppm (t1)
200
150
100
50
0
10000
20000
30000
40000

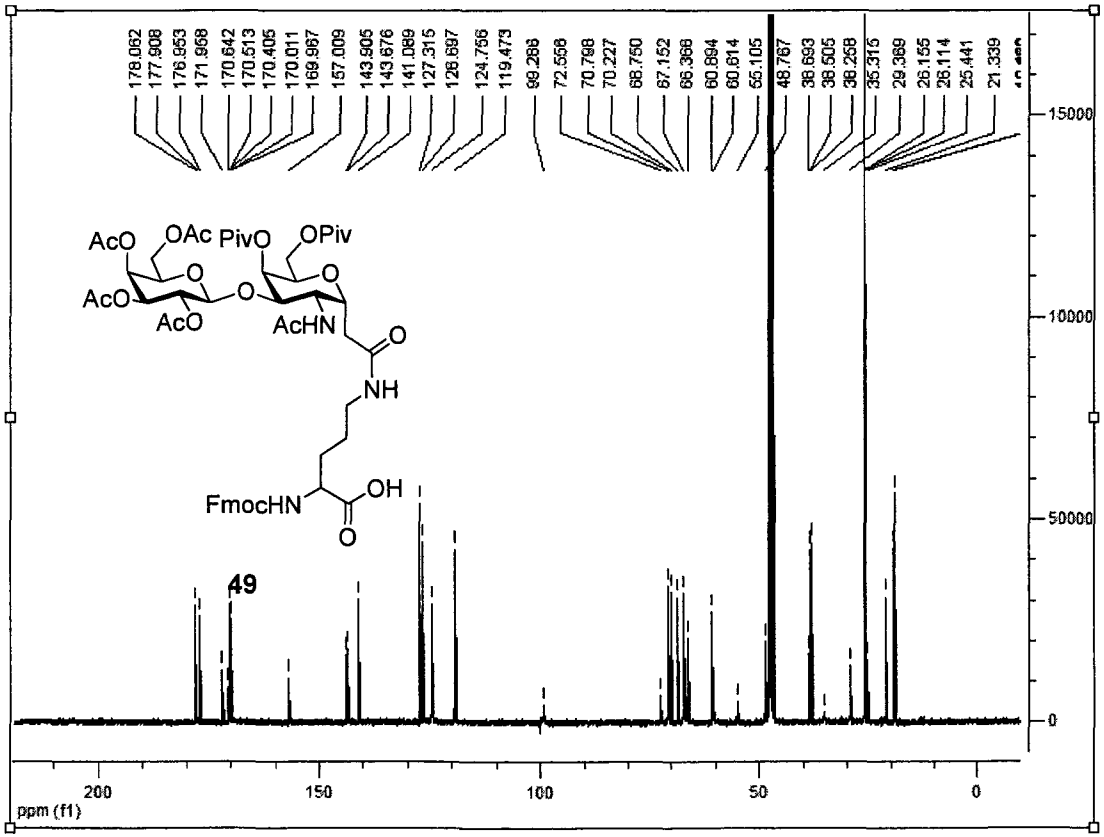

AcO
OAc
PivO
OPiv
AcO
AcO
AcHN
46
40000
30000
20000
10000
0
1.13
1.13
1.17
1.17
4.13
1.00
1.03
7.25
1.01
17.05
1.34
19.92
10.0
5.0
ppm (t1)

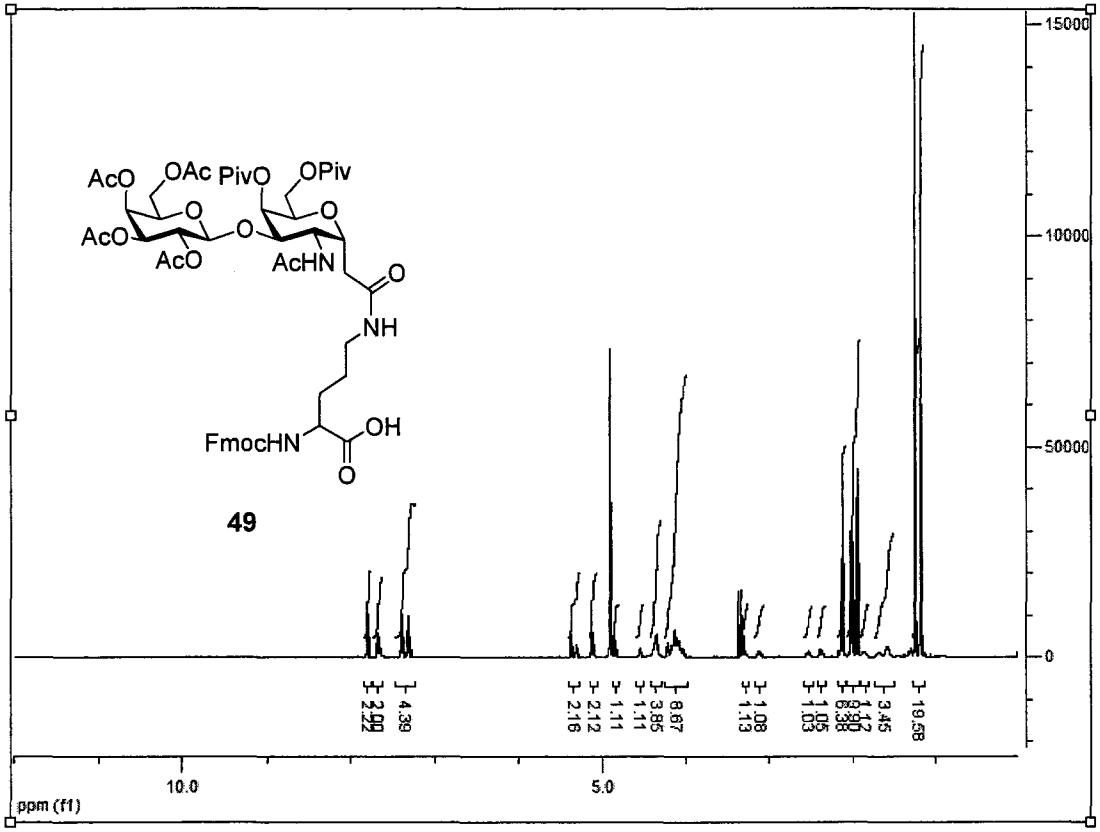

47
AcO
AcO
AcO
OAc
PivO
OPiv
AcHN
OH
O
177.984
177.234
174.045
171.106
170.309
170.017
169.922
98.599
72.732
71.856
70.627
68.783
66.808
65.467
60.796
59.509
48.965
38.715
35.783
26.958
26.849
22.957
20.484
20.455
20.427
ppm (t1)
200
150
100
50
0
0
10000
20000
30000
40000
50000
60000

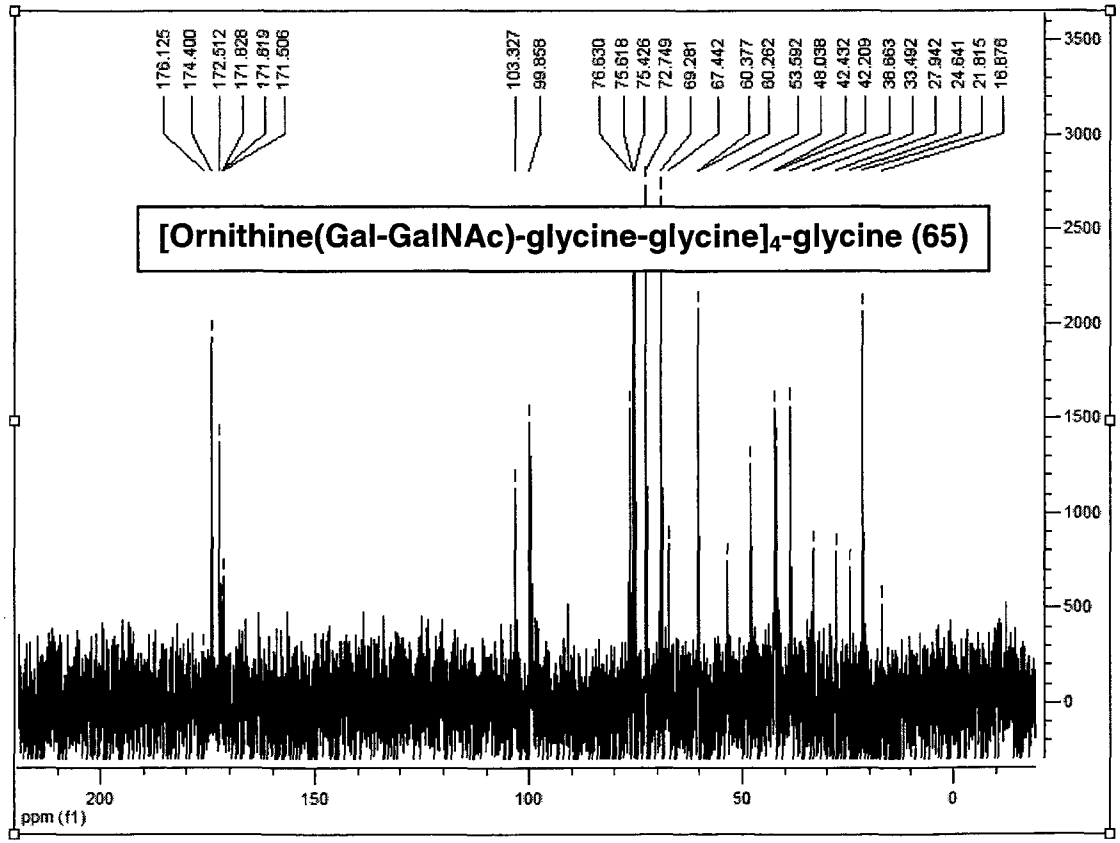

AcO
OAc
PivO
OPiv
AcO
AcO
AcHN
O
OH
47
10000
50000
0
1.27
1.26
2.04
1.17
1.26
5.83
15.27
18.37
10.0
5.0
ppm (t1)

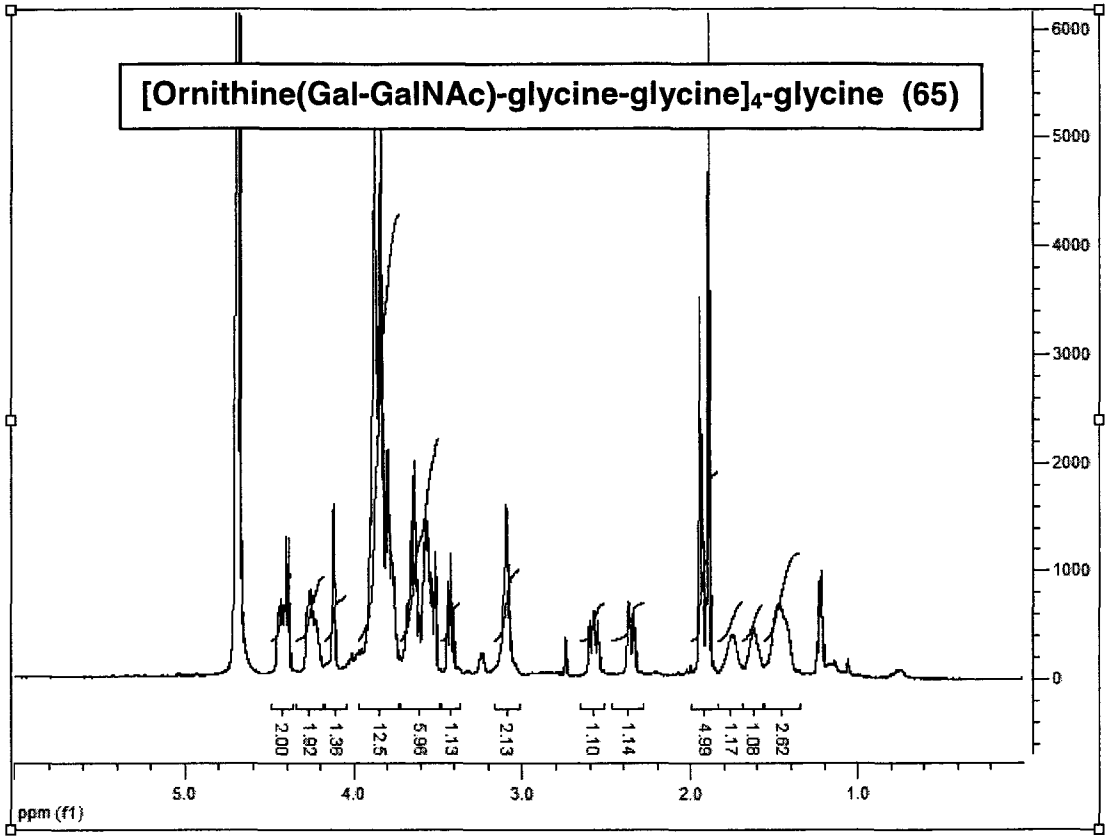

48
FmocHN
OBn
NH
AcHN
AcO
AcO
AcO
OAc
PivO
OPiv
177.430
172.351
170.664
170.314
170.158
170.035
169.958
169.197
156.313
143.976
143.668
141.310
135.307
128.651
128.560
128.482
128.210
127.767
127.094
125.156
120.045
98.997
73.367
72.257
70.783
70.659
66.936
67.298
67.030
66.908
65.377
64.147
60.870
59.792
53.900
50.052
47.192
38.926
38.892
38.704
ppm (f1)
200
150
100
50
0
5000

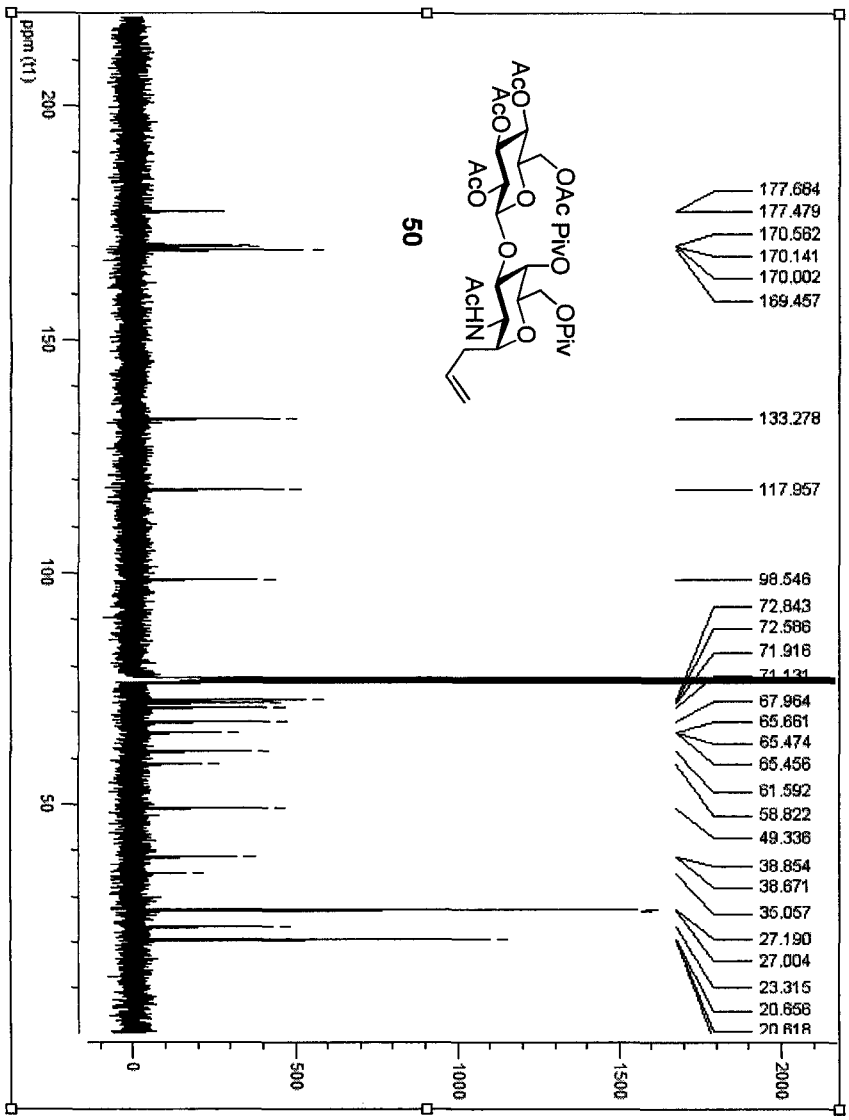

AcO
OAc
PivO
OPiv
AcO
AcO
AcHN
O
NH
FmocHN
OBn
O
48
15000
10000
50000
0
10.0
5.0
ppm (f1)

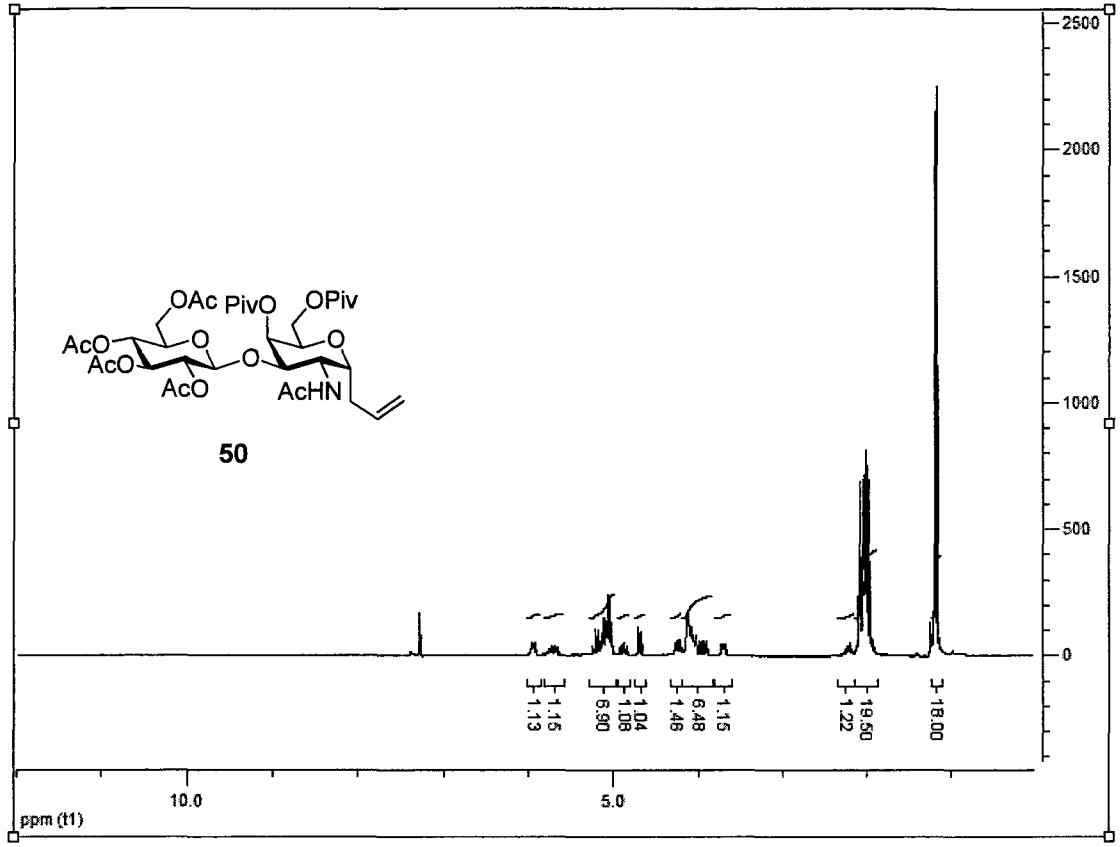

178.062
177.908
176.953
171.968
170.642
170.513
170.405
170.011
169.967
157.009
143.905
143.676
141.089
127.315
126.697
124.756
119.473
99.266
72.558
70.798
70.227
68.750
67.152
66.368
60.894
60.614
55.105
48.767
38.693
38.505
38.258
35.316
29.369
26.155
26.114
25.441
21.339
AcO
OAc
PivO
OPiv
AcO
AcO
AcHN
NH
FmocHN
OH
49
15000
10000
50000
0
200
150
100
50
0
ppm (f1)

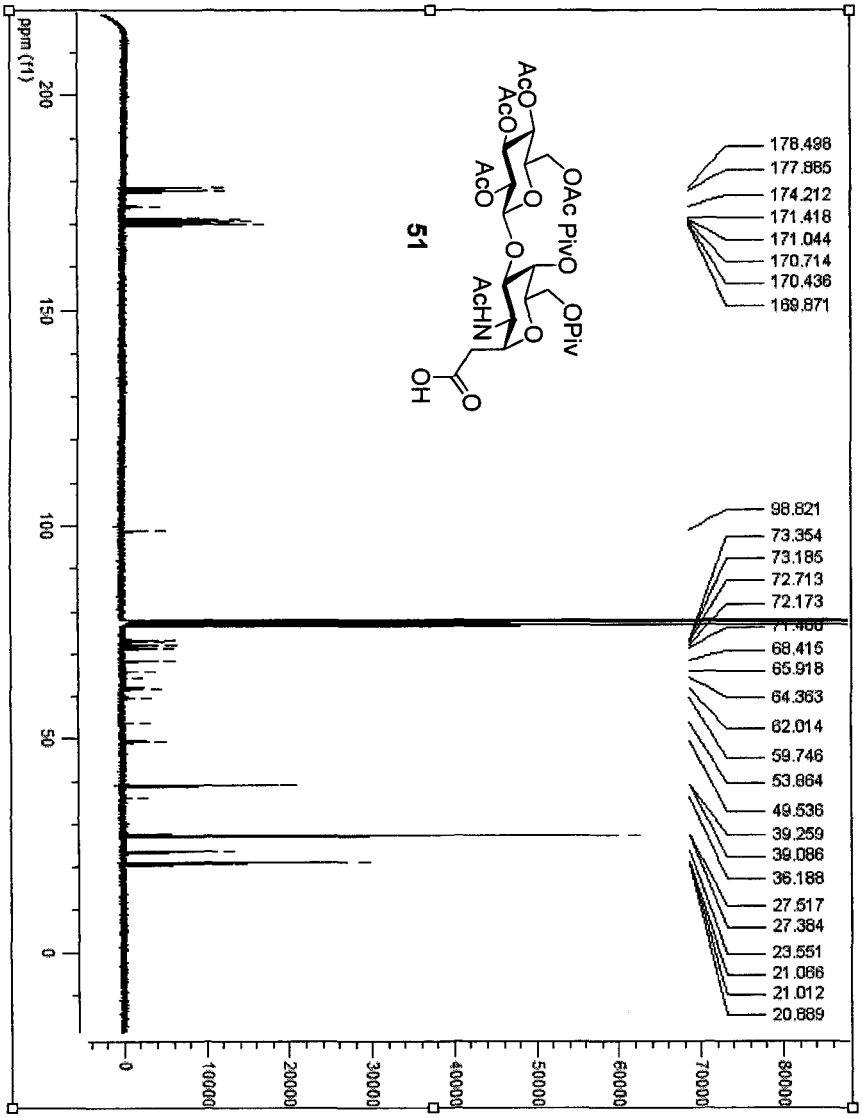

AcO
OAc
PivO
OPiv
AcO
AcO
AcHN
NH
FmocHN
OH
49
15000
10000
50000
0
4.39
2.16
2.12
1.11
1.11
8.67
1.08
1.13
1.05
1.03
3.45
19.58
10.0
5.0
ppm (f1)

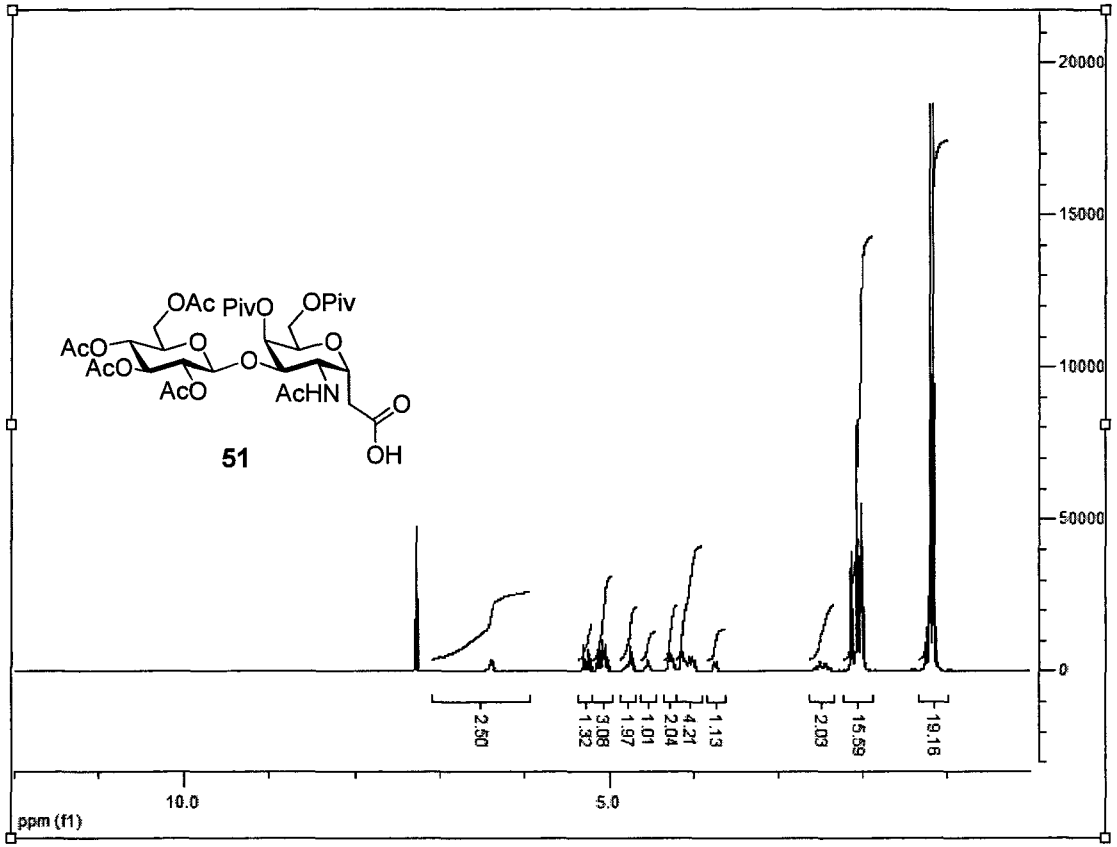

176.125
174.400
172.512
171.826
171.819
171.506
103.327
99.858
76.630
75.618
75.426
72.749
69.281
67.442
60.377
60.262
53.592
48.038
42.432
42.209
38.863
33.492
27.942
24.641
21.815
16.878
[Ornithine(Gal-GalNAc)-glycine-glycine]4-glycine (65)
3500
3000
2500
2000
1500
1000
500
0
200
150
100
50
0
ppm (f1)

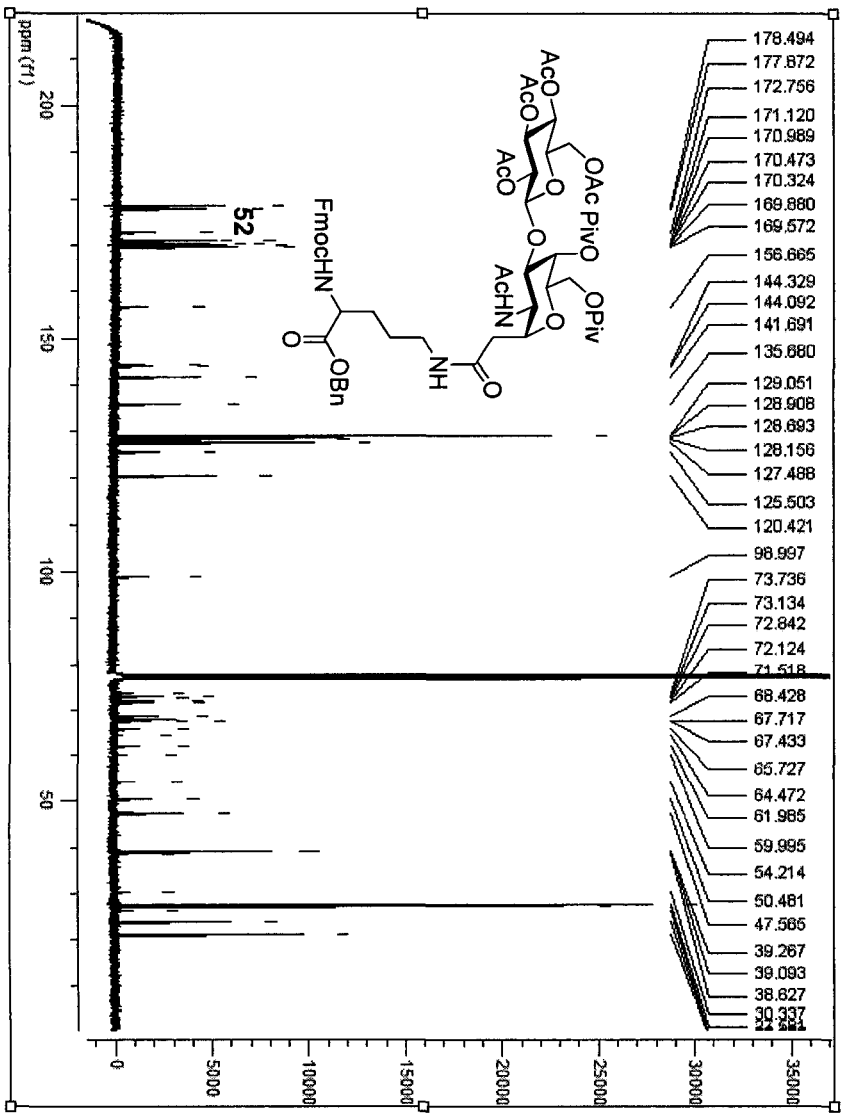

[Ornithine(Gal-GalNAc)-glycine-glycine]4-glycine (65)
6000
5000
4000
3000
2000
1000
0
2.00
1.92
1.36
12.5
5.96
1.13
2.13
1.10
1.14
4.99
1.17
1.08
2.62
5.0
4.0
3.0
2.0
1.0
ppm (f1)

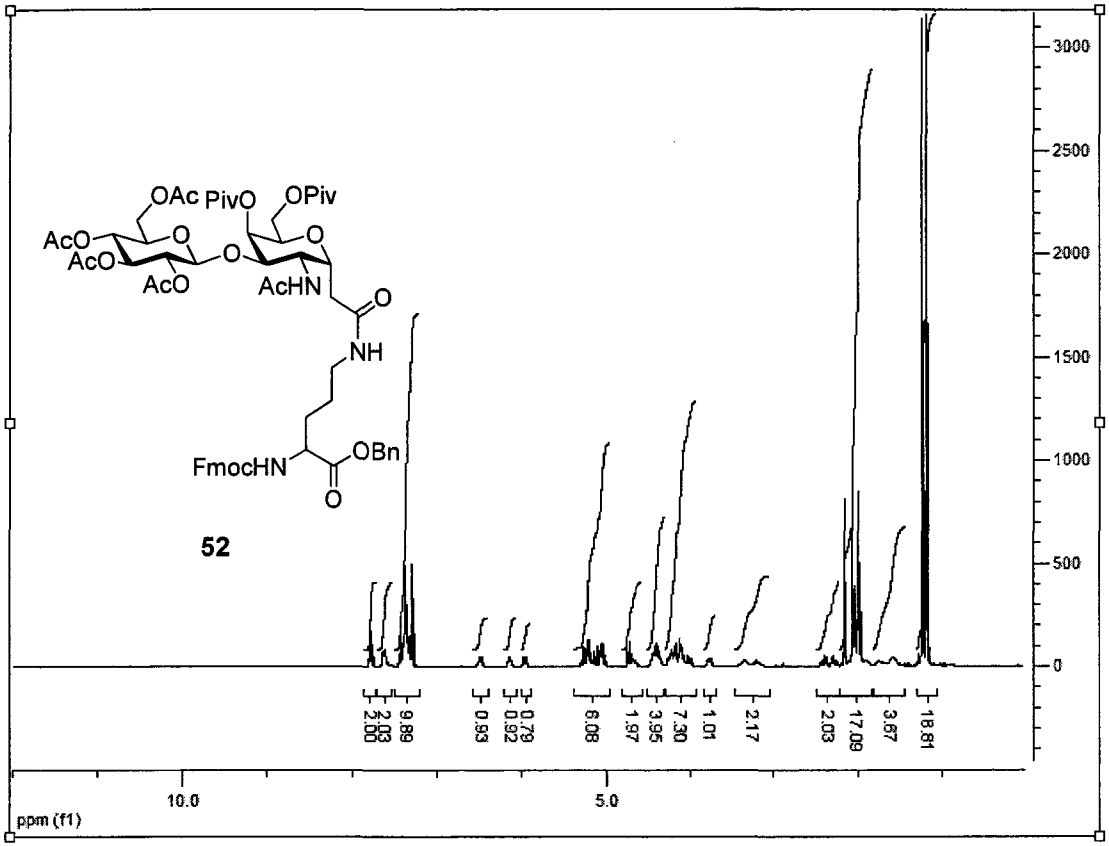

ppm (t1)
200
150
100
50
0
500
1000
1500
2000
AcO
AcO
AcO
OAc
PivO
OPiv
AcHN
50
177.684
177.479
170.562
170.141
170.002
169.457
133.278
117.957
98.546
72.843
72.586
71.916
67.964
65.661
65.474
65.456
61.592
58.822
49.336
38.854
38.671
36.057
27.190
27.004
23.315
20.656
20.618

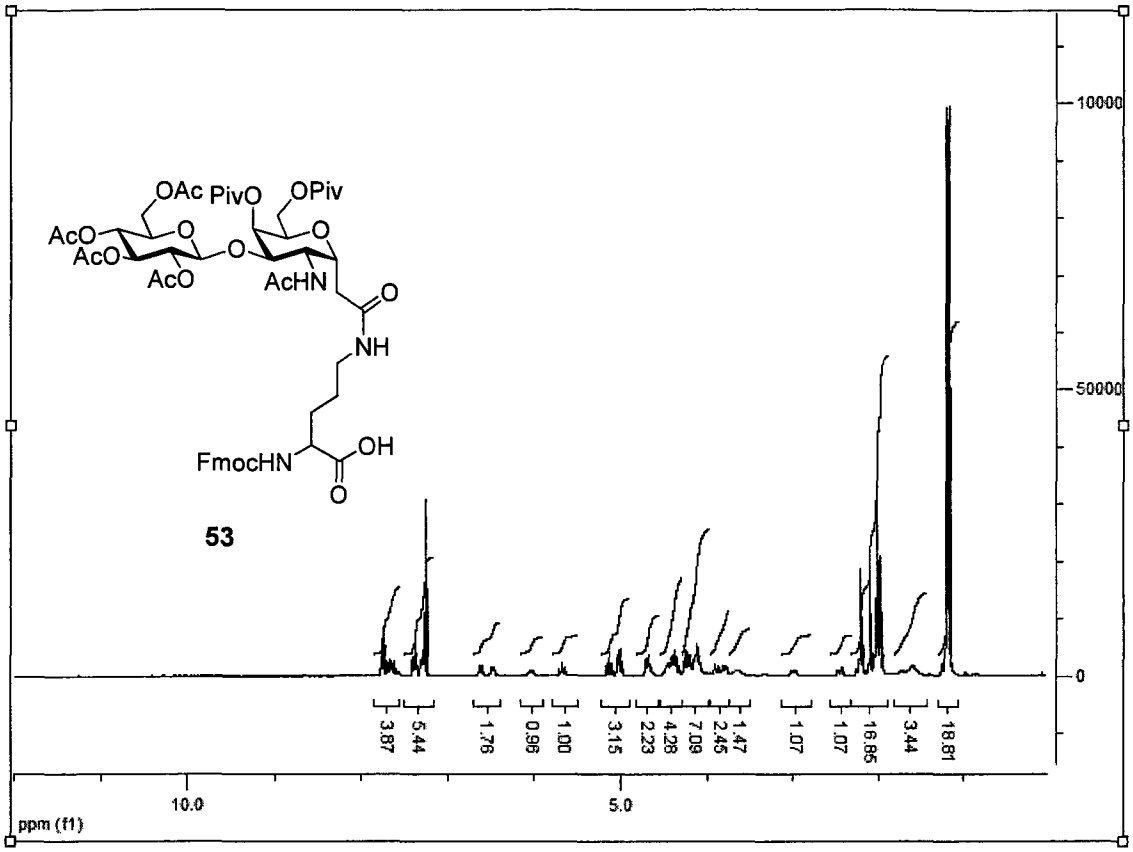

OAc
PivO
OPiv
AcO
AcO
AcO
AcHN
50
2500
2000
1500
1000
500
0
1.13
1.15
6.90
1.08
1.04
1.46
6.48
1.15
1.22
19.50
18.00
10.0
5.0
ppm (t1)

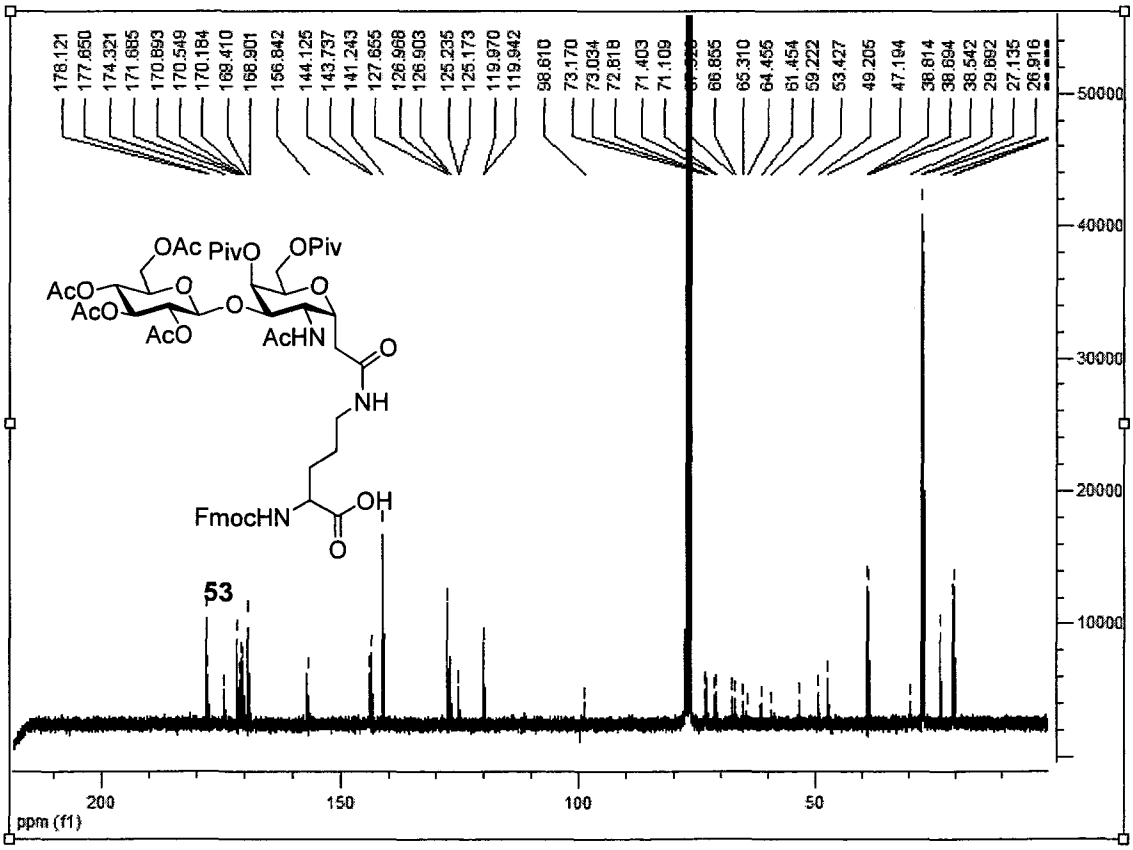

ppm (t1)
200
150
100
50
0
AcO
AcO
AcO
OAc
PivO
OPiv
AcHN
OH
O
51
178.498
177.885
174.212
171.418
171.044
170.714
170.436
169.871
98.821
73.354
73.185
72.713
72.173
68.415
65.918
64.363
62.014
59.746
53.864
49.536
39.259
39.086
36.188
27.517
27.384
23.551
21.066
21.012
20.869
0
10000
20000
30000
40000
50000
60000
70000
80000

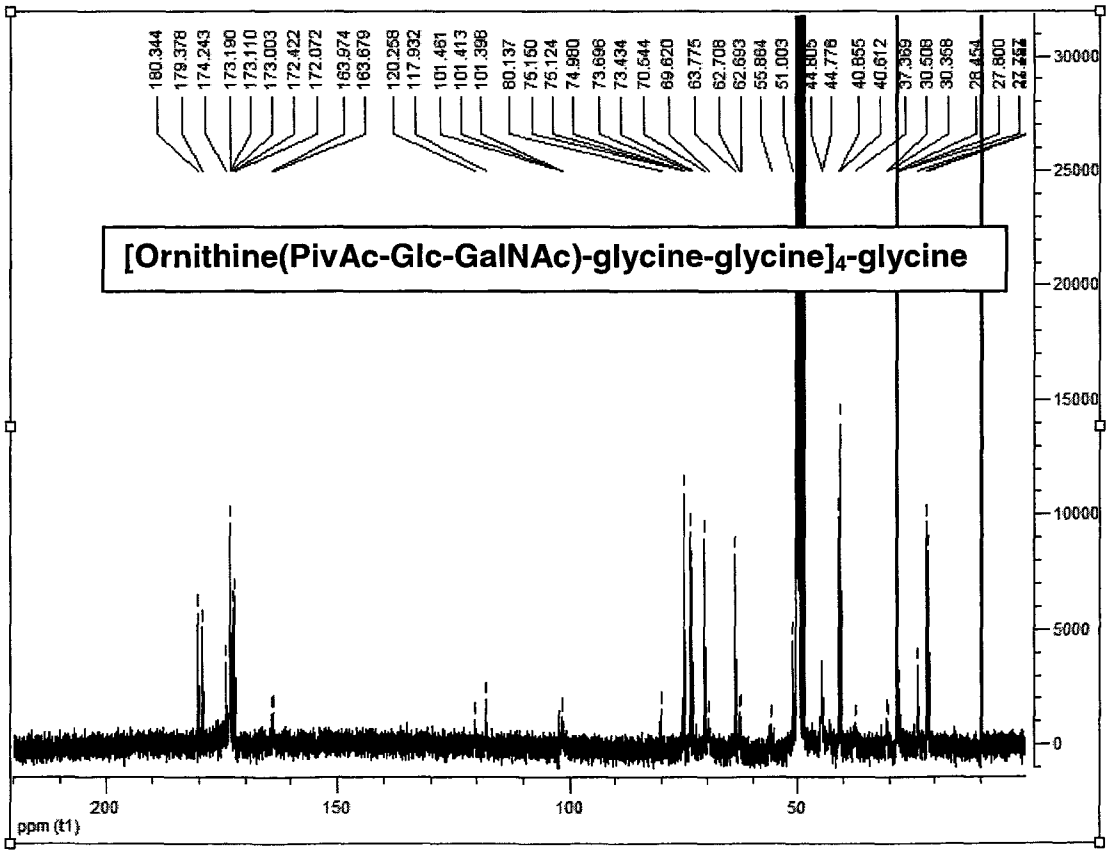

OAc
PivO
OPiv
AcO
AcO
AcO
AcHN
O
OH
51
20000
15000
10000
50000
0
2.50
3.08
1.32
1.97
1.01
2.04
4.21
1.13
2.03
15.59
19.16
10.0
5.0
ppm (f1)

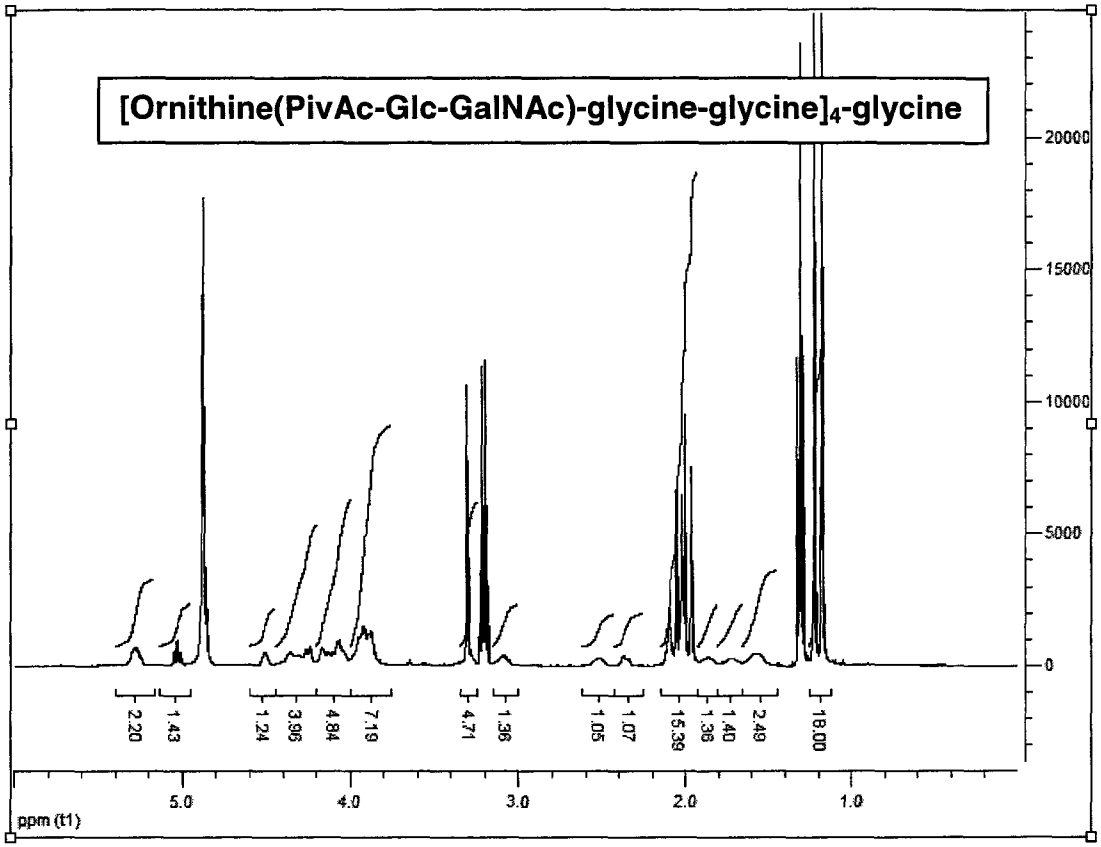

52
FmocHN
OBn
NH
AcHN
AcO
AcO
AcO
OAc
PivO
OPiv
178.494
177.872
172.756
171.120
170.989
170.473
170.324
169.860
169.572
156.665
144.329
144.092
141.691
135.680
129.051
128.908
128.693
128.156
127.488
125.503
120.421
98.997
73.736
73.134
72.842
72.124
68.428
67.717
67.433
65.727
64.472
61.985
59.995
54.214
50.481
47.565
39.267
39.093
38.827
30.337
ppm (f1)
200
150
100
50
0
5000
10000
15000
20000
25000
30000
35000

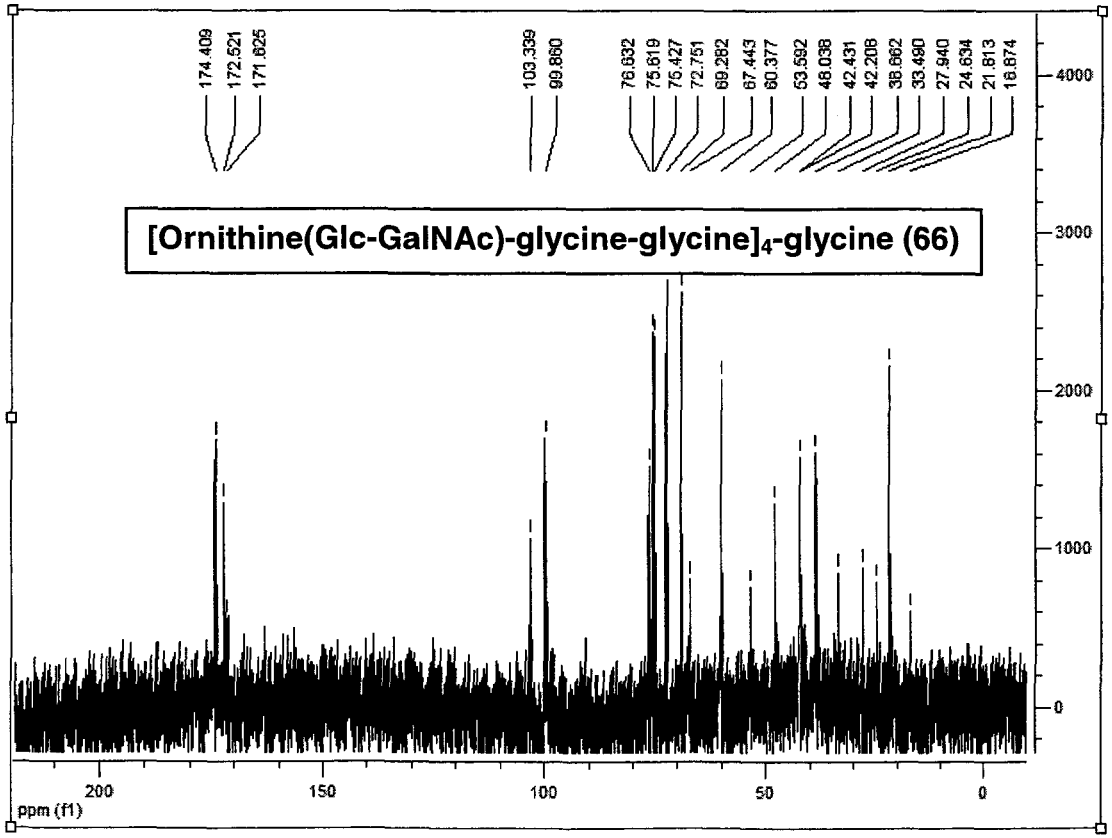

OAc
PivO
OPiv
AcO
AcO
AcO
O
O
O
AcHN
O
NH
FmocHN
OBn
O
52
3000
2500
2000
1500
1000
500
0
2.00
2.03
9.89
0.93
0.92
0.79
6.08
1.97
3.95
7.30
1.01
2.17
2.03
17.09
3.67
18.81
10.0
5.0
ppm (f1)

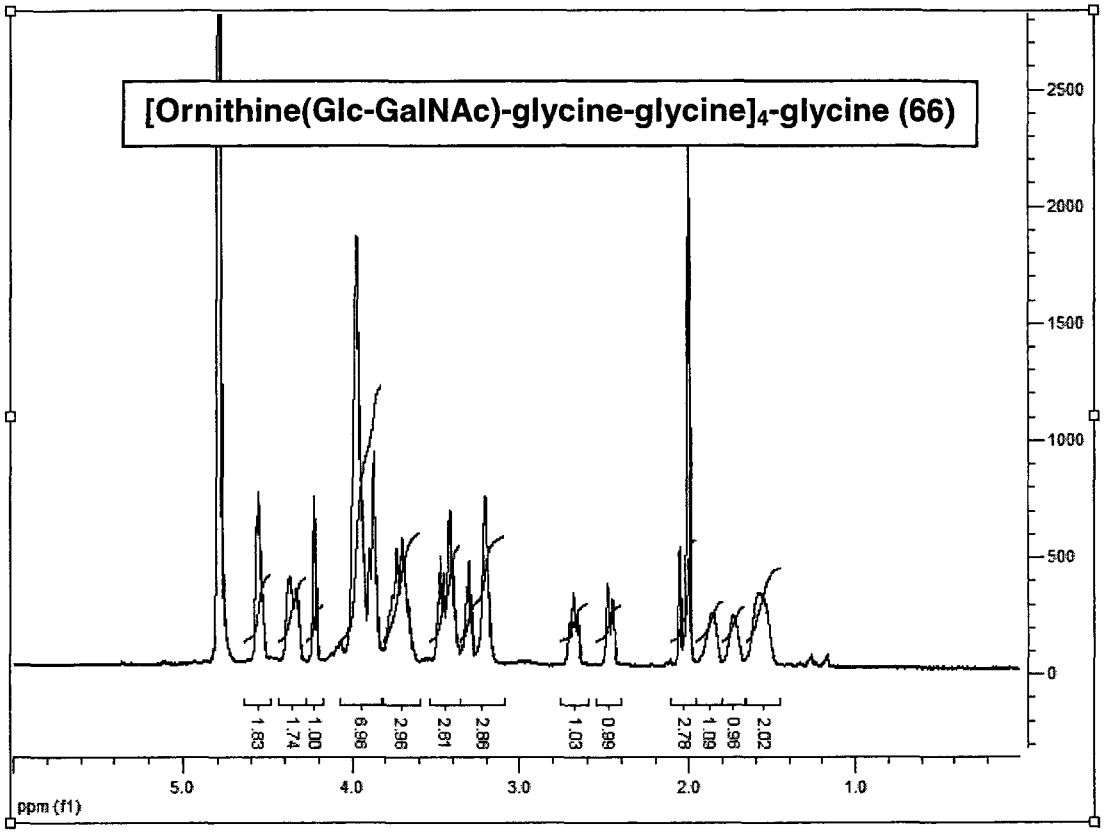

OAc
PivO
OPiv
AcO
AcO
AcO
AcHN
NH
FmocHN
OH
53
3.87
5.44
1.76
0.99
1.00
3.15
2.23
4.28
7.09
2.45
1.47
1.07
1.07
19.85
3.44
18.81
10000
50000
0
10.0
5.0
ppm (f1)

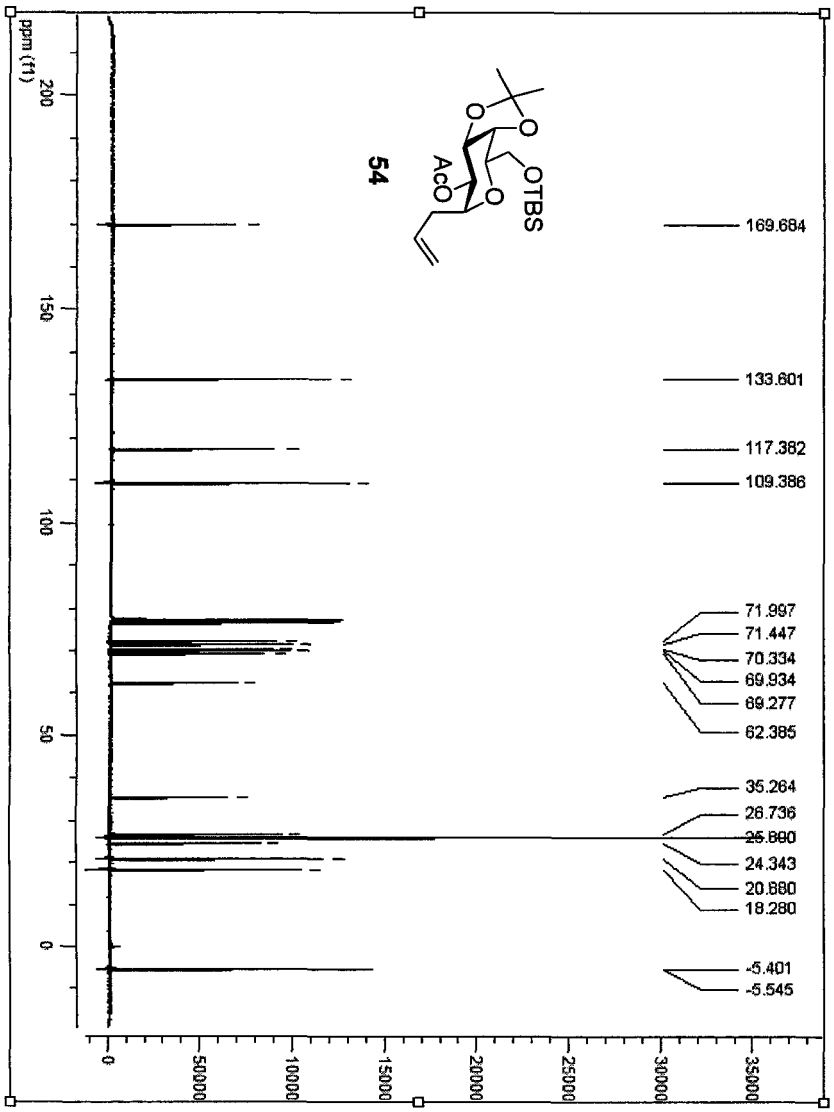

178.121
177.650
174.321
171.885
170.893
170.549
170.184
169.410
168.901
156.842
144.125
143.737
141.243
127.655
126.968
126.903
125.235
125.173
119.970
119.942
98.610
73.170
73.034
72.818
71.403
71.109
66.855
65.310
64.455
61.454
59.222
53.427
49.205
47.184
38.814
38.694
38.542
29.892
27.135
26.916
OAc
PivO
OPiv
AcO
AcO
AcO
AcHN
NH
FmocHN
OH
53
50000
40000
30000
20000
10000
200
150
100
50
ppm (f1)

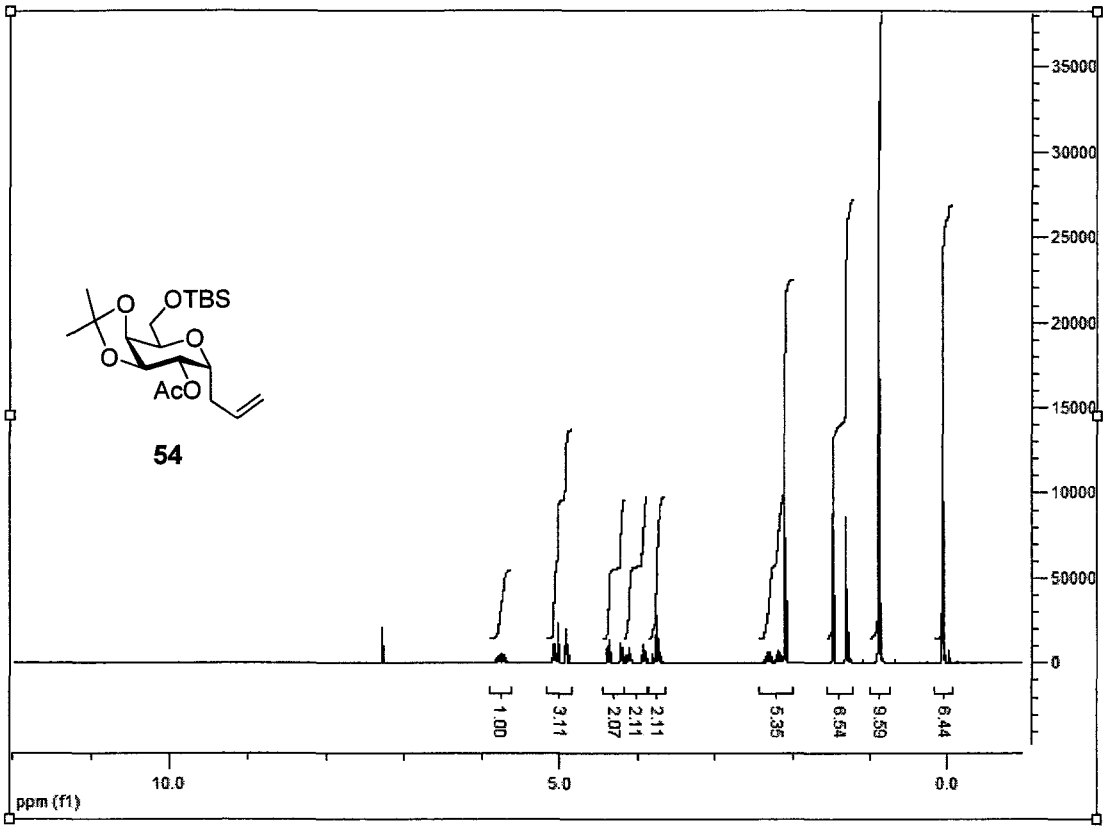

180.344
179.378
174.243
173.190
173.110
173.003
172.422
172.072
163.974
163.679
120.258
117.932
101.461
101.413
101.396
80.137
75.150
75.124
74.980
73.696
73.434
70.544
69.620
63.775
62.708
62.693
55.884
51.003
44.805
44.776
40.855
40.612
37.369
30.508
30.358
28.454
27.800
[Ornithine(PivAc-Glc-GalNAc)-glycine-glycine]4-glycine
30000
25000
20000
15000
10000
5000
0
200
150
100
50
ppm (t1)

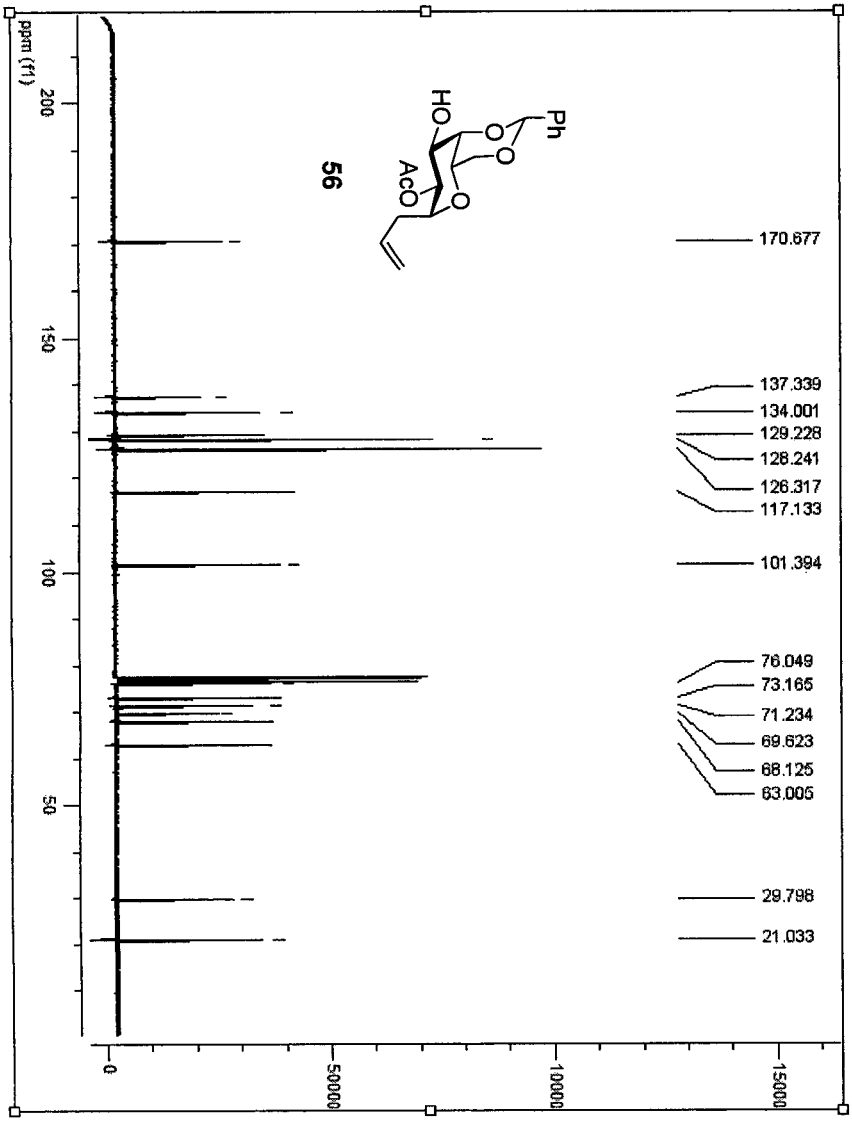

[Ornithine(PivAc-Glc-GalNAc)-glycine-glycine]4-glycine
20000
15000
10000
5000
0
2.20
1.43
1.24
3.96
4.84
7.19
4.71
1.36
1.05
1.07
15.39
1.36
1.40
2.49
18.00
5.0
4.0
3.0
2.0
1.0
ppm (t1)

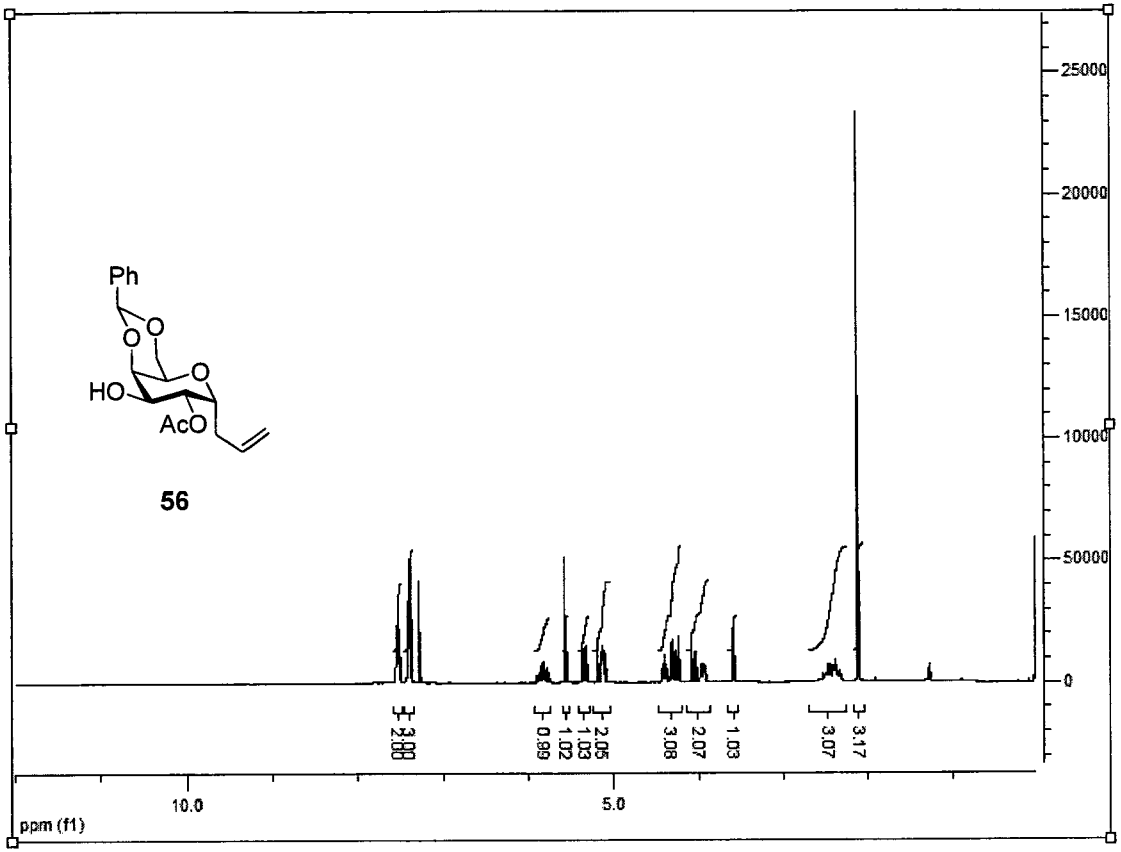

174.409
172.521
171.625
103.339
99.860
76.632
75.619
75.427
72.751
69.282
67.443
60.377
53.592
48.038
42.431
42.208
38.862
33.490
27.940
24.634
21.813
16.874
[Ornithine(Glc-GalNAc)-glycine-glycine]4-glycine (66)
4000
3000
2000
1000
0
200
150
100
50
0
ppm (f1)


[Ornithine(Glc-GalNAc)-glycine-glycine]4-glycine (66)
2500
2000
1500
1000
500
0
1.83
1.74
1.00
6.96
2.96
2.91
2.86
1.03
0.99
2.78
1.09
0.96
2.02
5.0
4.0
3.0
2.0
1.0
ppm (f1)

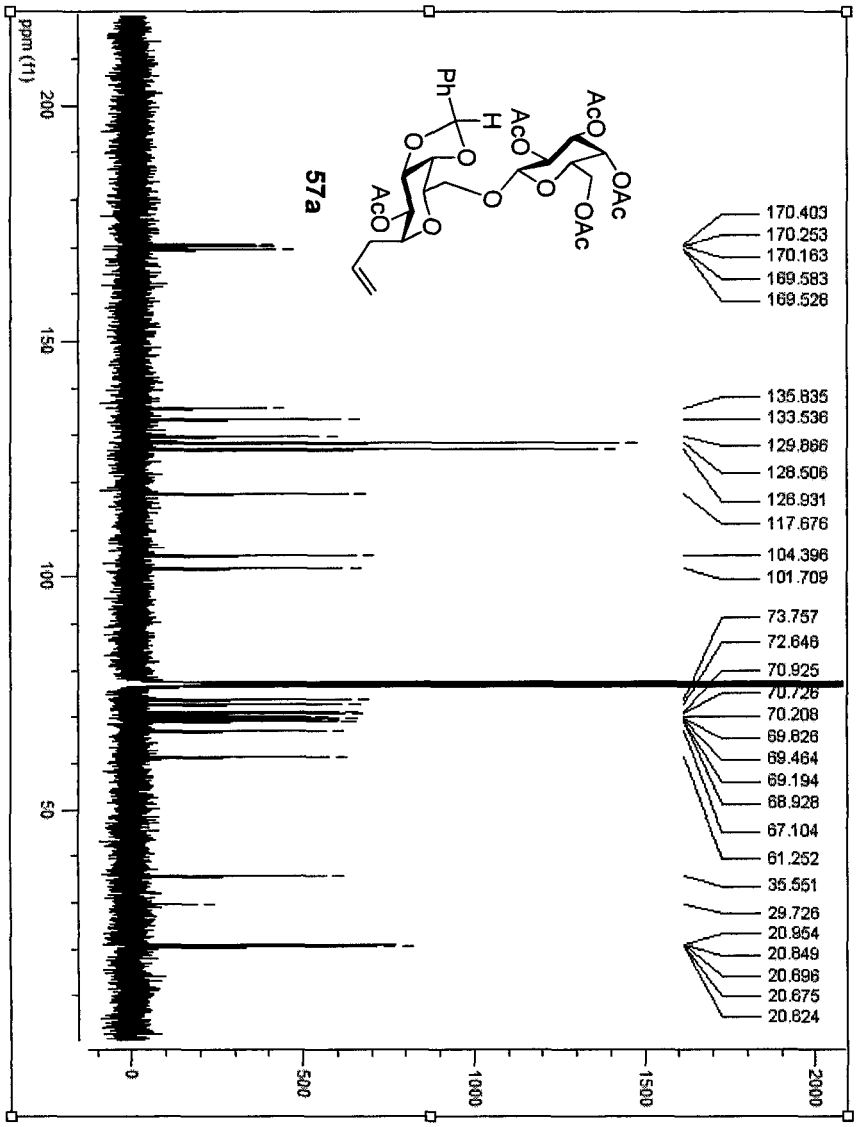

54
OTBS
AcO
O
O
O
O
ppm (f1)
200
150
100
50
0
169.684
133.601
117.382
109.386
71.997
71.447
70.334
69.934
69.277
62.385
35.264
26.736
25.880
24.343
20.880
18.280
-5.401
-5.545
0
50000
100000
150000
200000
250000
300000
350000

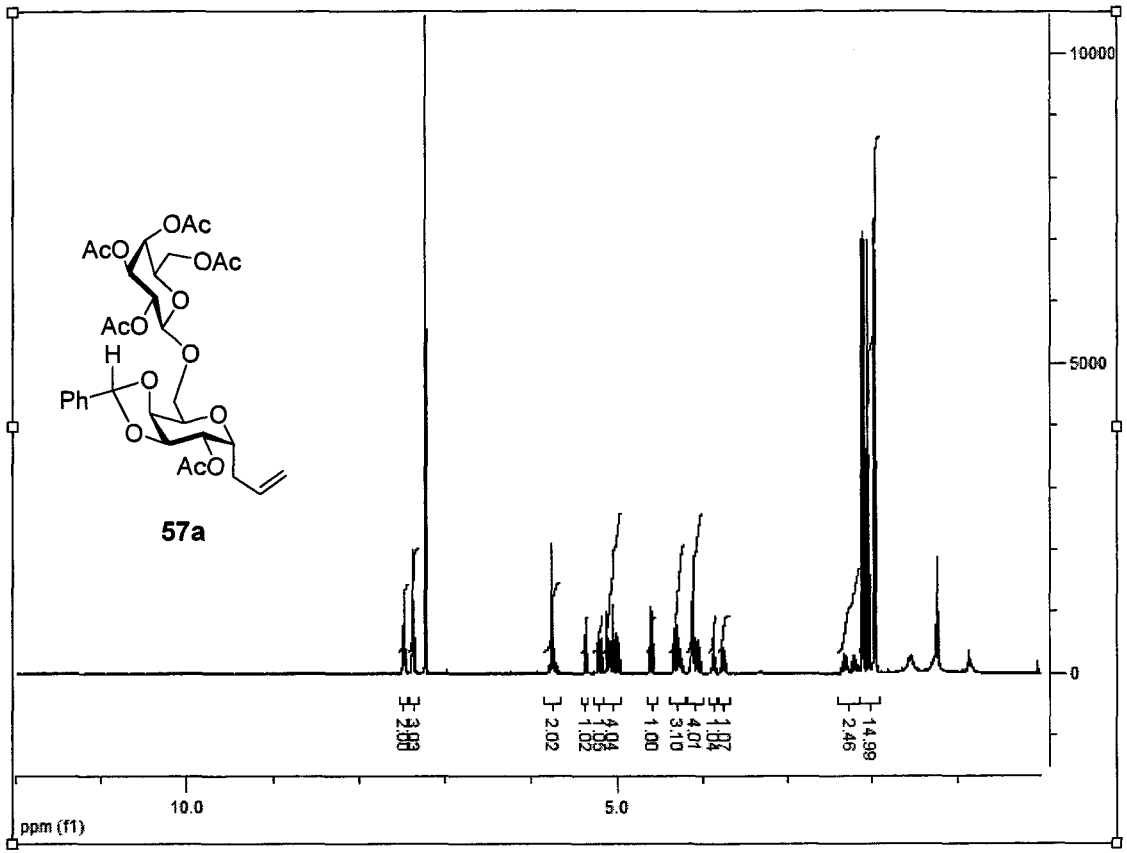

OTBS
O
O
O
AcO
54
35000
30000
25000
20000
15000
10000
50000
0
1.00
3.11
2.07
2.11
2.11
5.35
6.54
9.59
6.44
10.0
5.0
0.0
ppm (f1)


ppm (f1)
200
150
100
50
0
50000
100000
150000
HO
Ph
O
O
AcO
O
56
170.677
137.339
134.001
129.228
128.241
126.317
117.133
101.394
76.049
73.165
71.234
69.623
68.125
63.005
29.798
21.033

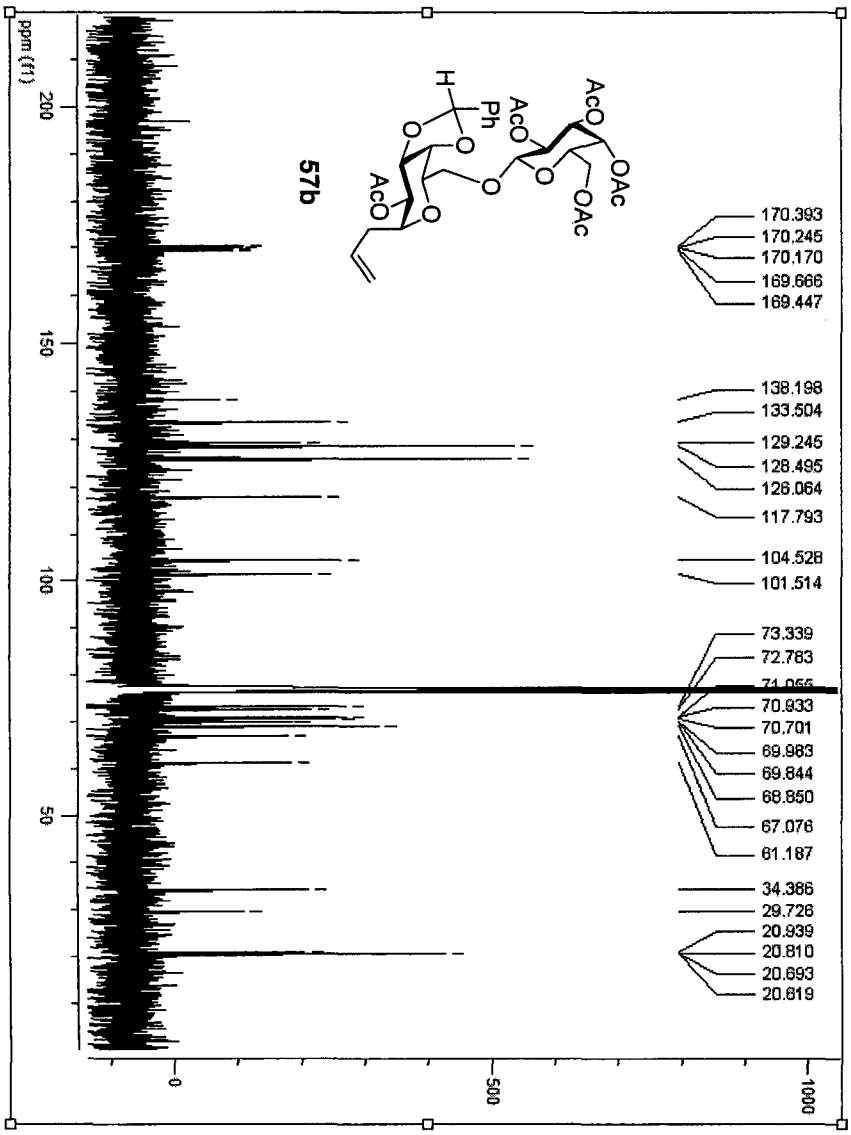

Ph
O
O
O
HO
AcO
25000
20000
15000
10000
50000
0
3.00
0.89
1.02
1.03
2.05
3.08
2.07
1.03
3.07
3.17
10.0
5.0
ppm (f1)

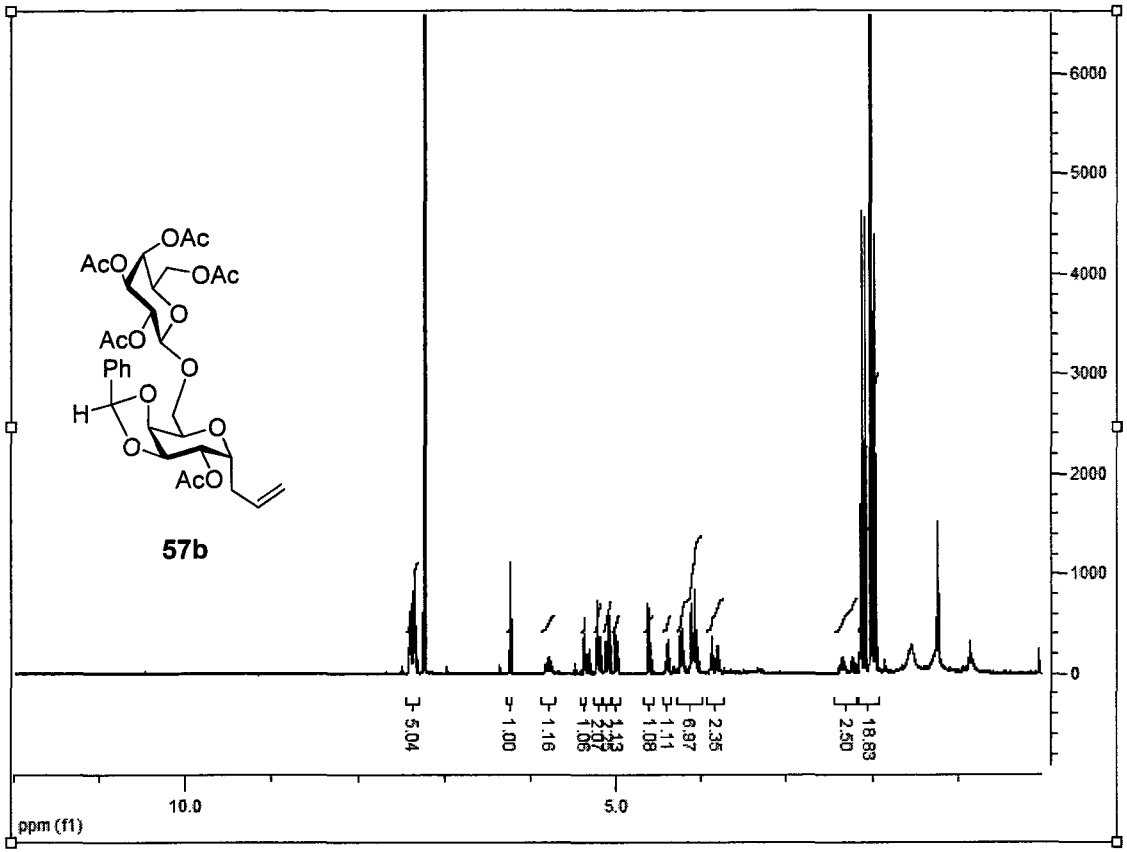
56


57a
Ph
H
AcO
AcO
AcO
OAc
OAc
170.403
170.253
170.163
169.583
169.528
135.835
133.536
129.866
128.506
126.931
117.676
104.396
101.709
73.757
72.646
70.925
70.726
70.208
69.826
69.464
69.194
68.928
67.104
61.252
35.551
29.726
20.954
20.849
20.696
20.675
20.624
ppm (f1)
200
150
100
50
0
500
1000
1500
2000

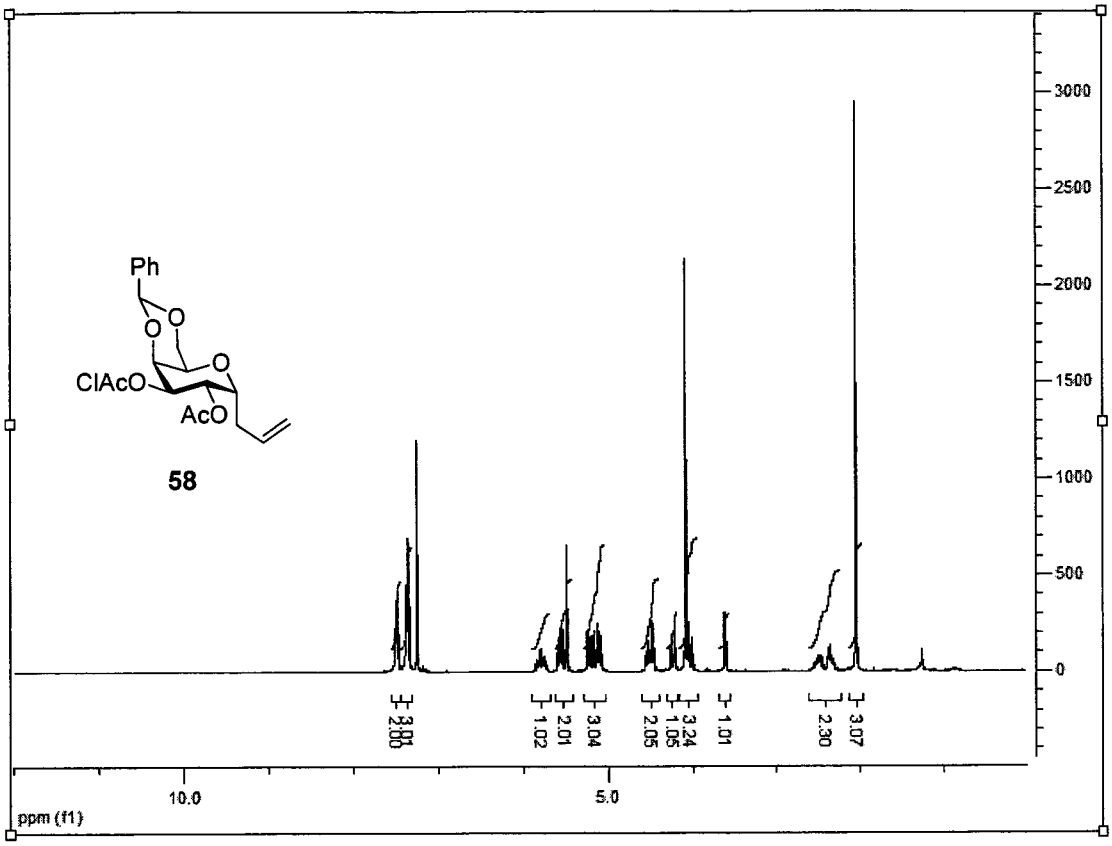

OAc
AcO
OAc
AcO
H
Ph
AcO
10000
5000
0
2.00
2.02
1.02
1.00
3.10
4.01
2.46
14.99
10.0
5.0
ppm (f1)

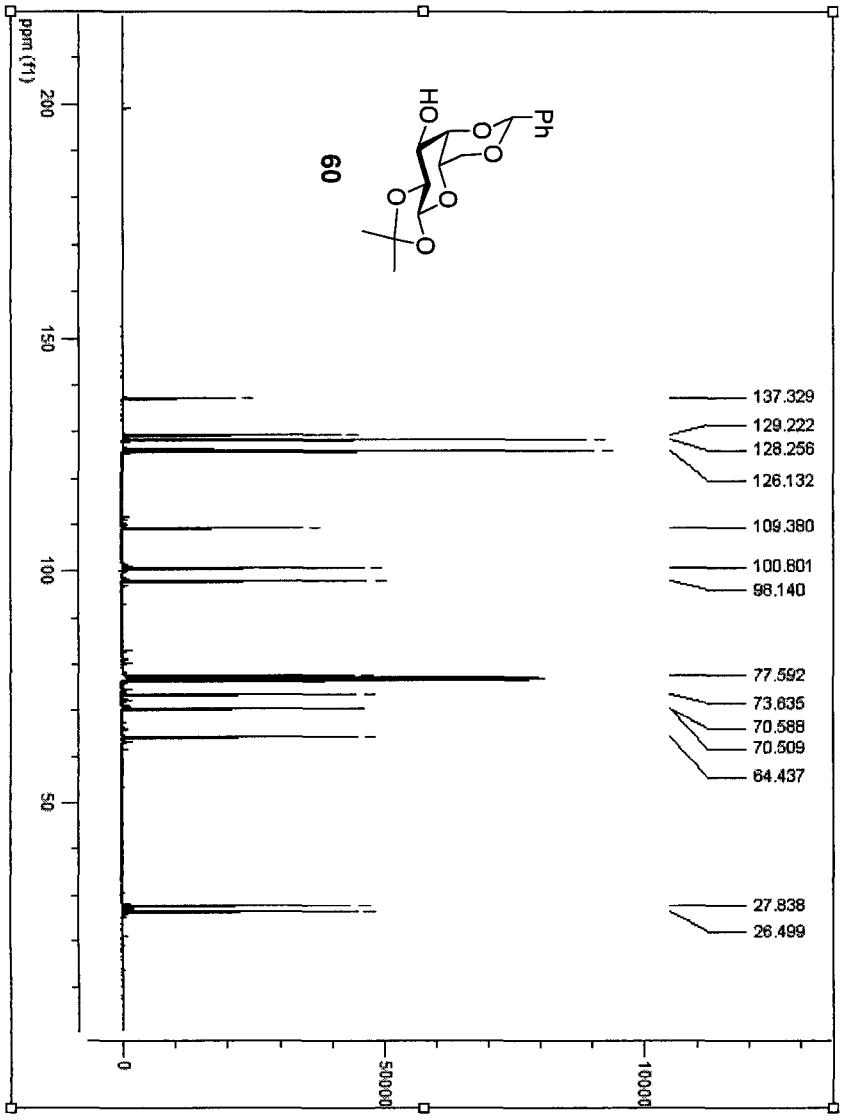
57a

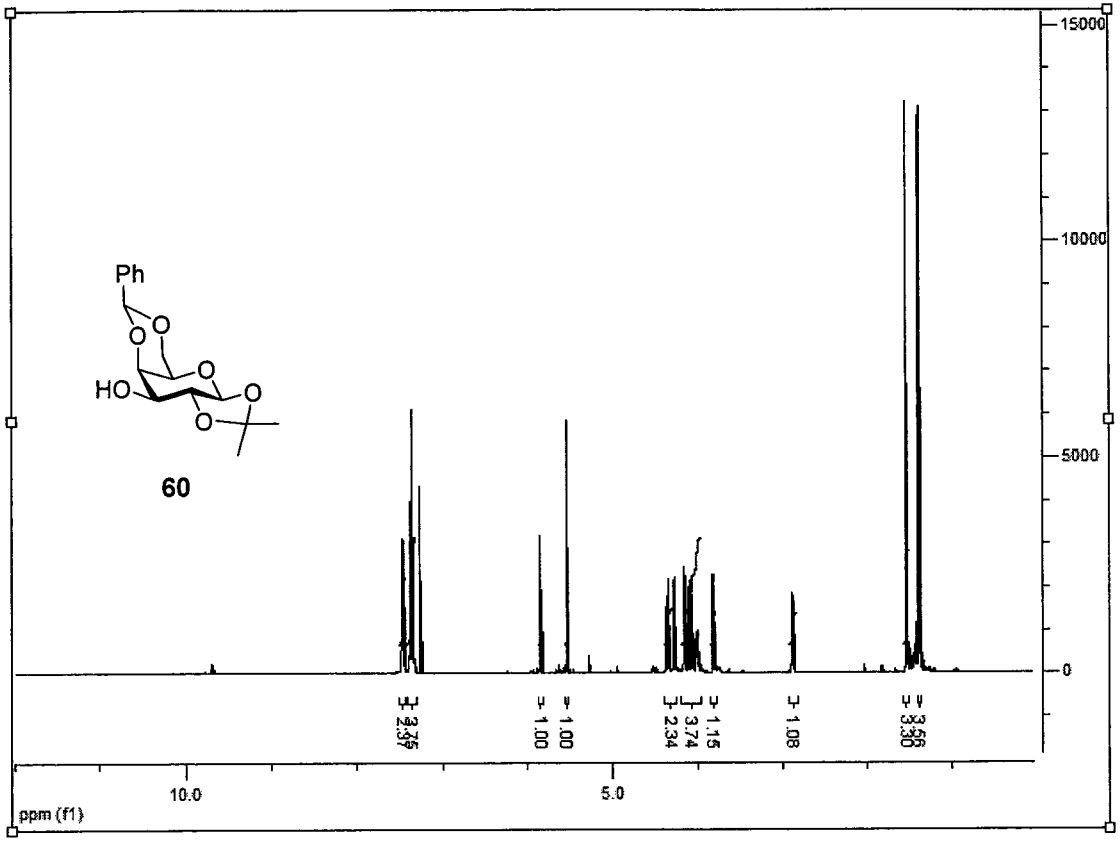

57b
ppm (f1)
200
150
100
50
0
500
1000
170.393
170.245
170.170
169.666
169.447
138.198
133.504
129.245
128.495
126.064
117.793
104.528
101.514
73.339
72.783
70.933
70.701
69.983
69.844
68.850
67.076
61.187
34.386
29.726
20.939
20.810
20.693
20.619

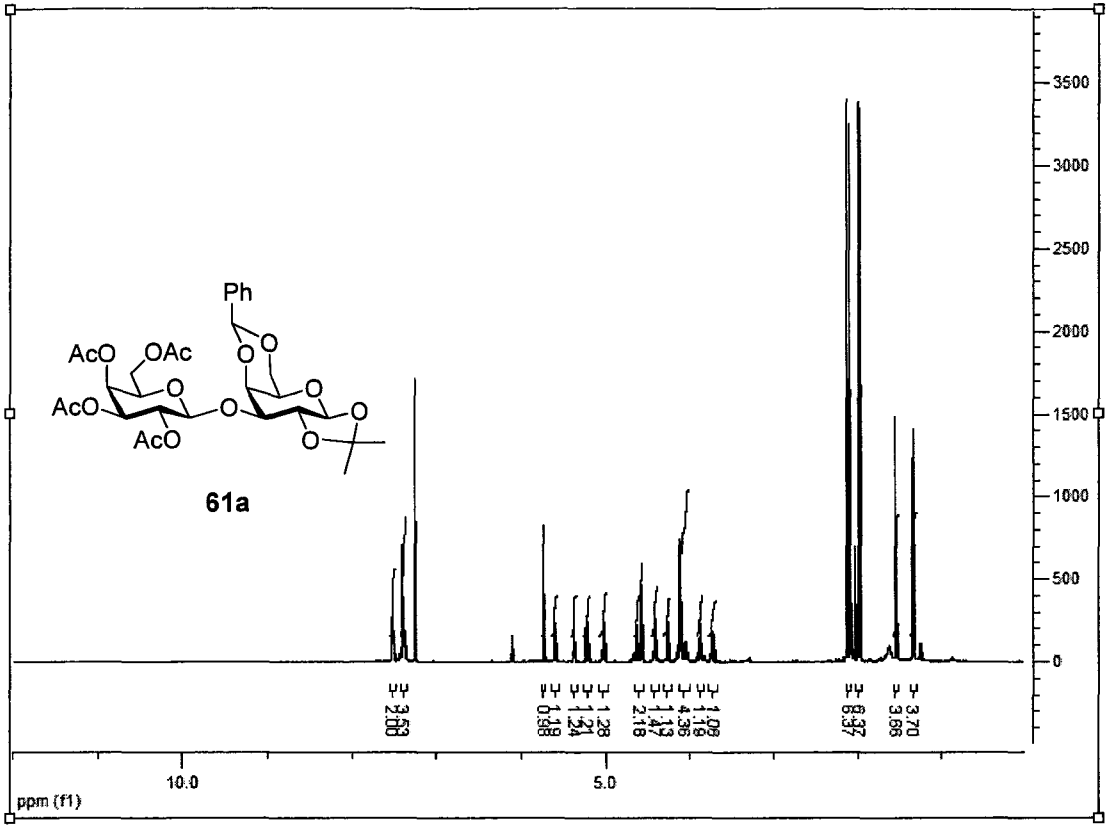

OAc
AcO
OAc
O
AcO
Ph
O
H
O
O
O
AcO
6000
5000
4000
3000
2000
1000
0
5.04
1.00
1.16
1.08
1.11
6.97
2.35
2.50
18.83
10.0
5.0
ppm (f1)

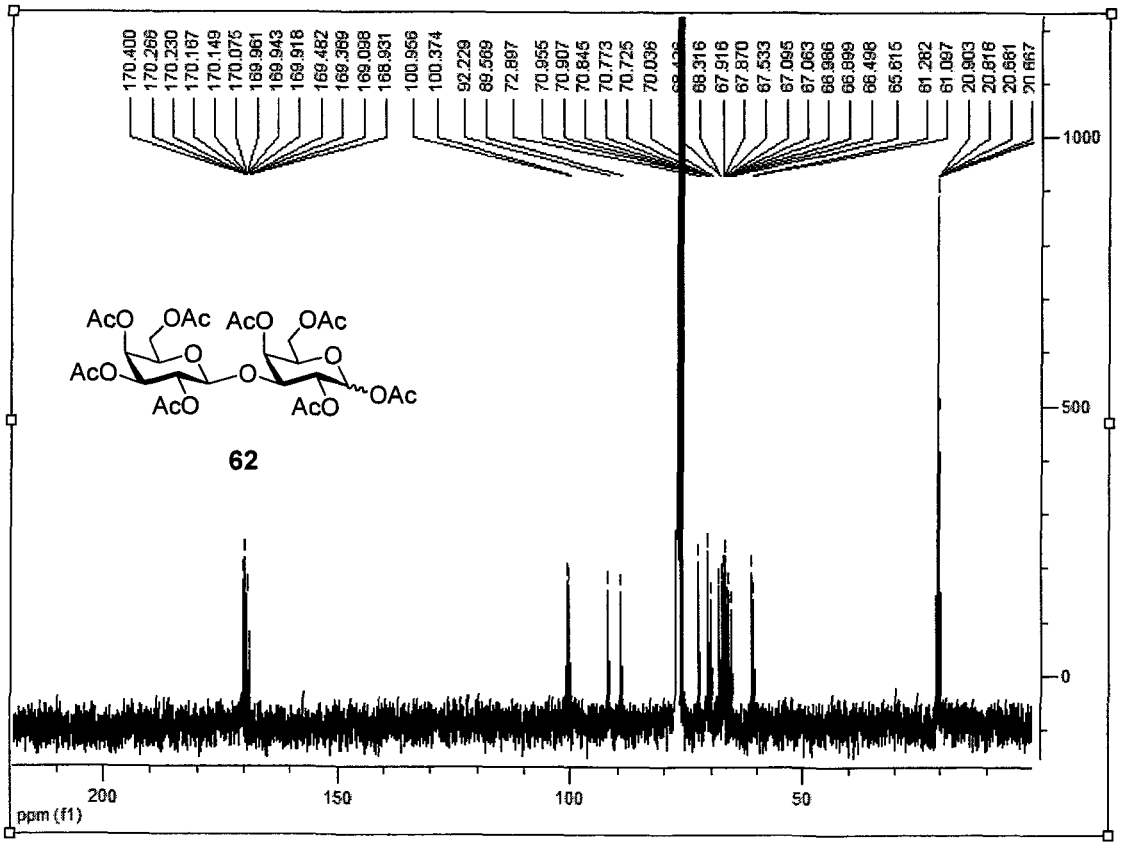
57b


Ph
O
O
O
ClAcO
AcO
58
3000
2500
2000
1500
1000
500
0
2.80
3.01
1.02
2.01
3.04
2.05
1.05
3.24
1.01
2.30
3.07
10.0
5.0
ppm (f1)

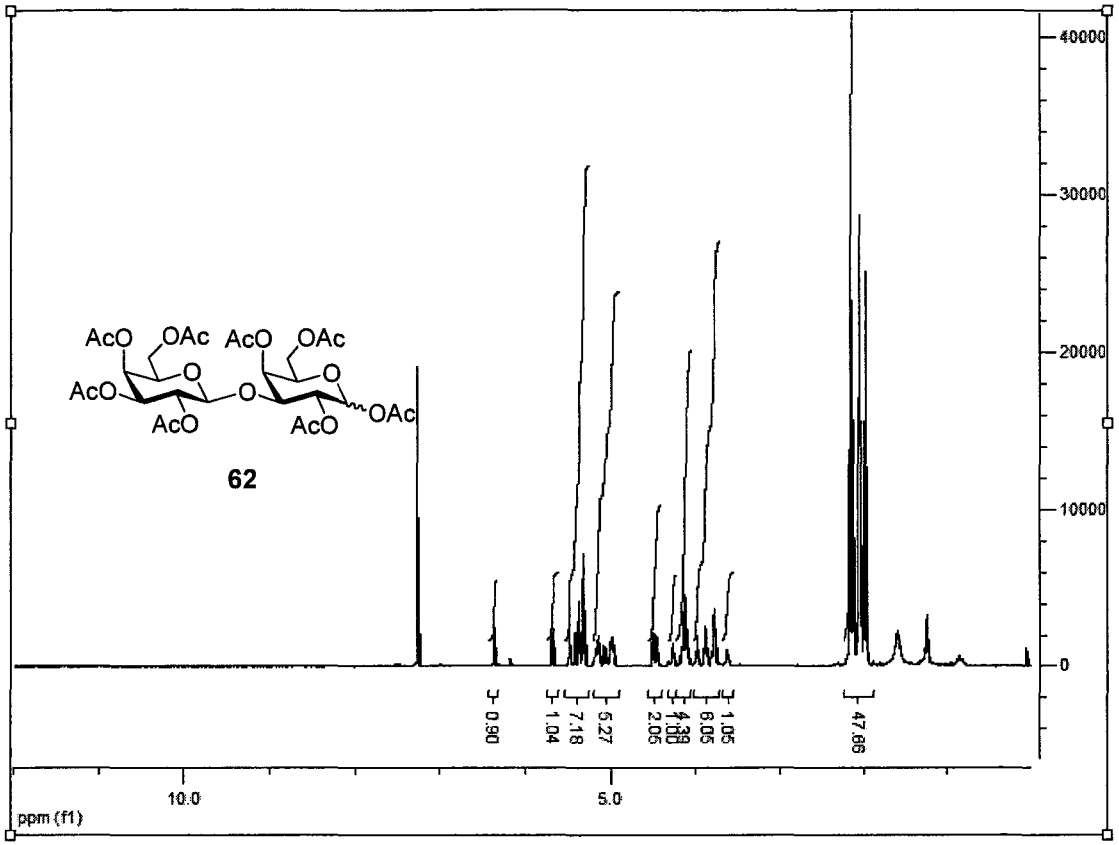

ppm (f1)
200
150
100
50
0
50000
10000
HO
Ph
O
O
O
O
O
60
137.329
129.222
128.256
126.132
109.380
100.801
98.140
77.592
73.635
70.588
70.509
64.437
27.838
26.499


Ph
O
O
O
HO
O
O
60
15000
10000
5000
0
3.75
2.39
1.00
1.00
1.15
3.74
2.34
1.08
3.56
3.30
10.0
5.0
ppm (f1)


Ph
AcO
OAc
AcO
AcO
O
O
O
O
O
O
O
61a
3500
3000
2500
2000
1500
1000
500
0
10.0
5.0
ppm (f1)


170.400
170.266
170.230
170.187
170.149
170.075
169.961
169.943
169.916
169.482
169.389
169.098
168.931
100.956
100.374
92.229
89.569
72.897
70.955
70.907
70.845
70.773
70.725
70.036
68.316
67.916
67.870
67.533
67.095
67.063
66.986
66.899
66.498
65.615
61.282
61.097
20.903
20.816
20.661
1000
500
0
200
150
100
50
ppm (f1)
AcO
OAc
AcO
OAc
AcO
O
O
AcO
AcO
AcO
OAc

62


AcO
OAc
AcO
OAc
AcO
O
O
AcO
OAc
AcO
AcO
62
40000
30000
20000
10000
0
0.90
1.04
7.18
5.27
2.05
6.05
1.05
47.66
10.0
5.0
ppm (f1)