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

Vincent Bouvet
AUTEUR DE LA THÈSE / AUTHOR OF THESIS

Ph.D. (Chemistry)
GRADE / DEGREE

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

Novel *O*-linked Antifreeze and *C*-linked Antifreeze Glycoprotein Analogues: Tailoring Recrystallization Inhibition and Thermal Hystersis Activities

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

André Beauchemin

William Willmore

Nathalie Goto (absent)

Mario Monteiro

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

NOVEL AFGP8 ANALOGUES: SYNTHESIS AND BIOLOGICAL EVALUATION OF POTENT RECRYSTALLIZATION INHIBITORS

BY

Vincent R. Bouvet

B.Sc (Honours in Chemistry) Université d'Orléans, France, 2000

M.Sc. (Honours in Organic Chemistry) Université d'Orléans, France 2001

Thesis submitted to the Faculty of Graduate & Postdoctoral Studies University of Ottawa in partial fulfillment of the requirements for the degree of Doctor of Philosophy

Ottawa-Carleton Chemistry Institute

Department of Chemistry

University of Ottawa

Ottawa, Ontario

Candidate

Vincent R. Bouvet

Supervisor

Professor Robert N. Ben

University of Ottawa

May, 2006-05-12

Vincent R. Bouvet

Your file Votre référence
ISBN: 978-0-494-25859-0
Our file Notre référence
ISBN: 978-0-494-25859-0

Accepted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in Chemistry
in the Graduate School of
Ottawa/Carleton University
2006

Robert N. Ben _____ September 11st, 2006
Advisor, Chemistry Department

André Beauchemin _____ September 11st, 2006
Ottawa Chemistry Department

Natalie K. Goto _____ September 11st, 2006
Ottawa Chemistry Department

William G. Willmore _____ September 11st, 2006
Carleton Chemistry Department

Mario A. Monteiro _____ September 11st, 2006
Outside Examiner, University of Guelph Chemistry Department

This work is dedicated to my parent and my family

Leuconoe, don't ask--it's forbidden to know what end the gods will give me or you.
Don't play with Babylonian fortune-telling either.
Better just deal with whatever comes your way.
Whether you'll see several more winters or whether the last one
be smart, drink your wine. Even as we speak, envious time is running away from us.
Seize the day (**Carpe diem**).

--"Hodes" I.XI Horace

ABSTRACT

This dissertation describes the study of native and synthetic antifreeze glycoproteins (AFGPs). Antifreeze glycoproteins (AFGPs) are biological antifreezes found in several species of Atlantic and Antarctic teleost fishes. These compounds have the ability to inhibit the growth of ice and protect organisms from cryoinjury and death. They exhibit two main physical properties associated with antifreeze activity. First, they possess the ability to create a non-colligative depression of the freezing point below that of the melting point; this phenomenon is referred to as *thermal hysteresis* (TH). Second, they possess the ability to inhibit the growth of ice, referred to as *recrystallization inhibition* (RI). Consequently the rational design and *de novo* synthesis of chemically and biologically stable AFGP analogues have become an attractive challenge. These properties have made AFGPs of great interest for medical, industrial and commercial applications.

AFGPs molecules have been thoroughly studied in order to understand their molecular mechanism of action. It is believed that conformation of these compounds plays an essential function during the ice adsorption process which is required for antifreeze activity. While several conformation studies of AFGP8 molecules have been published using different spectroscopic methods, an accepted conformation has failed to emerge and its role on antifreeze activity needs to be addressed. Chapter 3 examines the solution aggregation of AFGP8 using dynamic light scattering (DLS) and investigates if this phenomenon have any influence on the conformations of AFGP8 using circular dichroism (CD).

Although structure activity relationship (SAR) studies seem to be one of the most appropriate tools to understand the molecular mechanism of action of AFGPs, only few syntheses of AFGPs analogues, have been reported in the literature. Amongst these publications, even fewer have tested their analogues for thermal hysteresis activity and none have tested them for recrystallization inhibition. Recrystallization inhibition might be considered as the most relevant activity for cell storage at sub-zero temperatures, when dynamic ice shaping (DIS) and thermal hysteresis (TH) potentially damage cell membranes. Recently, it has been demonstrated that there is no correlation between these two activities. However, the required structural features of AFGPs for RI, TH and DIS activities are still to

be elucidated. Using *O*-linked glycopeptide analogues, the required structural feature for TH and RI activity are investigated in Chapter 4.

Although AFPG8 displays attractive RI activity, it also exhibits undesired properties for cryopreservation. For instance, it favours the formation of needle shape ice crystals below the TH gap. Furthermore, AFGPs molecules are not truly stable under basic, acid or enzymatic conditions due to their exocyclic oxygen. Consequently, chapter 5 describes the synthesis and study of stable *C*-linked analogues in order to obtain better cryoprotectants with potent RI activity and no DIS activity.

ACKNOWLEDGEMENTS

First of all I am very grateful to my supervisor, Dr. Robert N. Ben for his support, patience and guidance during these years. Thanks to him, I became a true chemist and learnt to enjoy synthetic challenges as well as mind puzzling mechanisms. Dear Rob, these past 5 years in SUNY Binghamton and University of Ottawa gave me an irreplaceable experience and for that I am forever thankful. I would also like to thank my dissertation defense committee, Dr. André Beauchemin, Dr. Natalie Goto, Mario Monteiro and Dr. William Willmore for taking the time to read my dissertation and providing me with suggestions for its improvement.

Next I would especially like to thank:

- Anastasia (Stacy) Murphy for her friendship, many scientific discussions and her help with the synthesis of some *O*-linked and *C*-linked analogues. Stacy thanks to you these long nights in the lab were more pleasant.
- Jennifer Chaytor for her patient helps with the editing of this document. Your help has been really appreciated and I wish you all the best for your Ph.D.
- The "Ben-Fallis Crew" (Frank, Pawel, Roger, Mathieu and Mike) for these endless mechanism discussions and for the enjoyable Friday night "safety meeting" to recuperate from the hard week and prepare for the week-end.
- The "French Crew" (Sara, Gwenaella and Christophe). These pleasant Friday night dinners and movie week-ends will never be forgotten. I wish you all the best for your career and hope our road will cross again.
- The "PUG (Productive undergrads) Crew" (Ferenst, Gianni, Jessica, Michael, Dana, Nicole and Alex) for your enthusiasm and hard work. Your productivity helped us all.

I am also thankful to all the present and past members of the Ben lab: Adewale, Owino, Madhu, Gene, Suhuai, Liz, Indura and John for all the memories. I would also like to express all my appreciation to the numerous undergrads who worked with me at Binghamton and Ottawa.

Most of all, I would like to thank my parent for their permanent encouragement and support during all these years. Without your help and your love nothing would have been possible.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

Chapter 2

Chapter 3

Chapter 4

Chapter 5

Future work

LIST OF SCHEMES

Chapter 1

Chapter 4

Chapter 5

Future work

LIST OF TABLES

Chapter 1

Chapter 3

Chapter 4

Chapter 5

LIST OF GRAPH

Chapter 1

LIST OF PICTURES

Chapter 1

LIST OF ABBREVIATIONS

Ac	Acetyl
AFP(s)	Antifreeze protein(s)
AFGP(s)	Antifreeze glycoprotein(s)
Ala	Alanie
Asn	Asparagine
b	broad
Bn	Benzyl
Boc	*tert*-Butyloxycabonyl
Bz	Benzoyl
CD	Circular dichroism
CDI	1,1-Carbonyldiimidazole
CI	Chemical Ionization
CH_2Cl_2	Dichloromethane
d	doublet
DCC	Dicyclohexylcarbodiimide
DCE	Dichloroethane
DCM	Dichloromethane
DLS	Dynamic light scattering
dd	doublet of doublets
DIPEA	Diisopropylethylamine
DIS	Dynamic Ice-Shaping
DMAP	4-Dimethylamino pyridine
DMF	*N*, *N*-Dimethyl formamide
dt	doublet of triplets
ESI	Electrospray Ionisation
EtOAc	Ethyl acetate
Et_2O	Diethyl ether
Fmoc	9-Fluorenylmethoxycarbonyl
Gal	Galactose
GalNAc	*N*-Acetyl-galactosamine
Glc	Glucose
Gly	Glycine
HBTU	2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	*N*-Hydroxybenzotriazole
Hz	Hertz
$IDCClO_4$	Iodonium dicollidine perchlorate
kDa	kiloDalton
Lys	Lysine
m	multiplet
M+	Parent molecular ion
MALDI	Matrix-assisted laser desorption ionization
Man	Mannose
MeCN	Acetonitrile
MLGS	Mean largest grain size

mOsmol	milliOsmolars
MS	Mass Spectrometry
ms	Molecular sieves
MTT	Methylthiazolydiphenyl-tetrazolium bromide
NIS	*N*-Iodosuccinimide
Orn	Ornitine
PBS	Phosphate-buffered saline
Pfp	Pentafluorophenyl
PPII	Polyproline type II
Ppm	parts per million
Pro	Proline
RI	Recrystallization inhibition
s	singlet
SEM	Standard error of the mean
Ser	Serine
SPPS	Solid-phase peptide synthesis
SAR	Structure-activity relationship
t	triplet
TBDMS	*tert*-Butyldimethylsilyl
TFA	Trifluoroacetic acid
TH	Thermal Hysteresis
Thr	Threonine
TMG	Tetramethylguanidine
TMSOTf	Trimethylsilyltrifluoromethane sulphonate

Chapter 1

INTRODUCTION

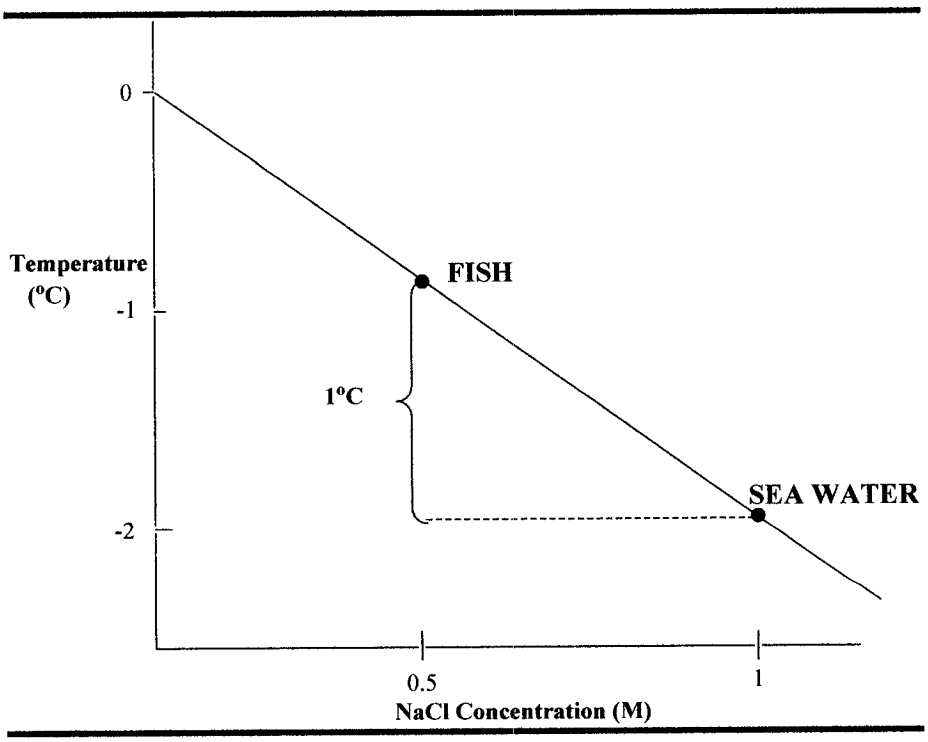
Bouvet, Vincent, Ben, Robert N.* *Cell Biochemistry and Biophysics* **2003**, 39, 133-144.

1 INTRODUCTION

Millions of years of evolution have allowed organisms to adapt to most of the biotopes present on earth. Adaptations that permit organisms to survive subzero environments[1-4] are amongst the most impressive. Such adaptations are often accomplished through the biosynthesis of molecules that have the ability to reduce the freezing point of their resident tissues thus confirming protection from cryoinjuries, cell damage and death.[5]

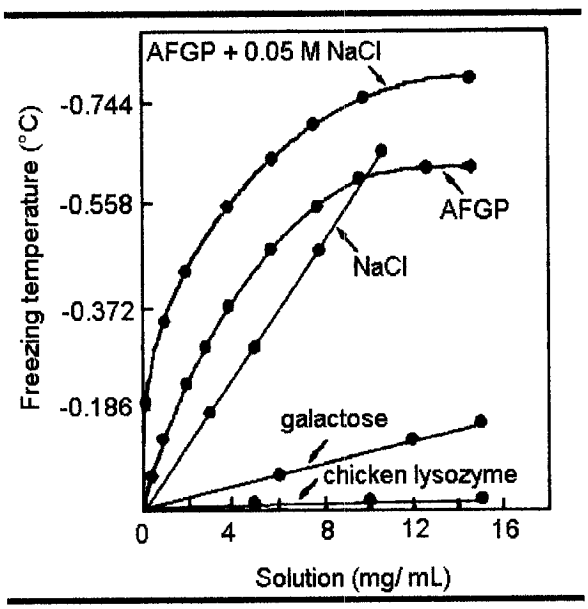
1.1 Biological Antifreezes

The freezing temperature of the extracellular media is approximately -0.86°C due to colligatively active substances, while the approximated temperature of polar Antarctic sea water is -1.86°C.[6] Hence there is a 1°C critical difference between the temperature of the fish and that of water. This difference is enough to kill any species unadapted to the Antarctic or Arctic water temperatures (Graph 1).



Graph 1: Freezing point dependence on NaCl concentration: *this 1°C different in freezing temperature is critical to the survival of the organisms*

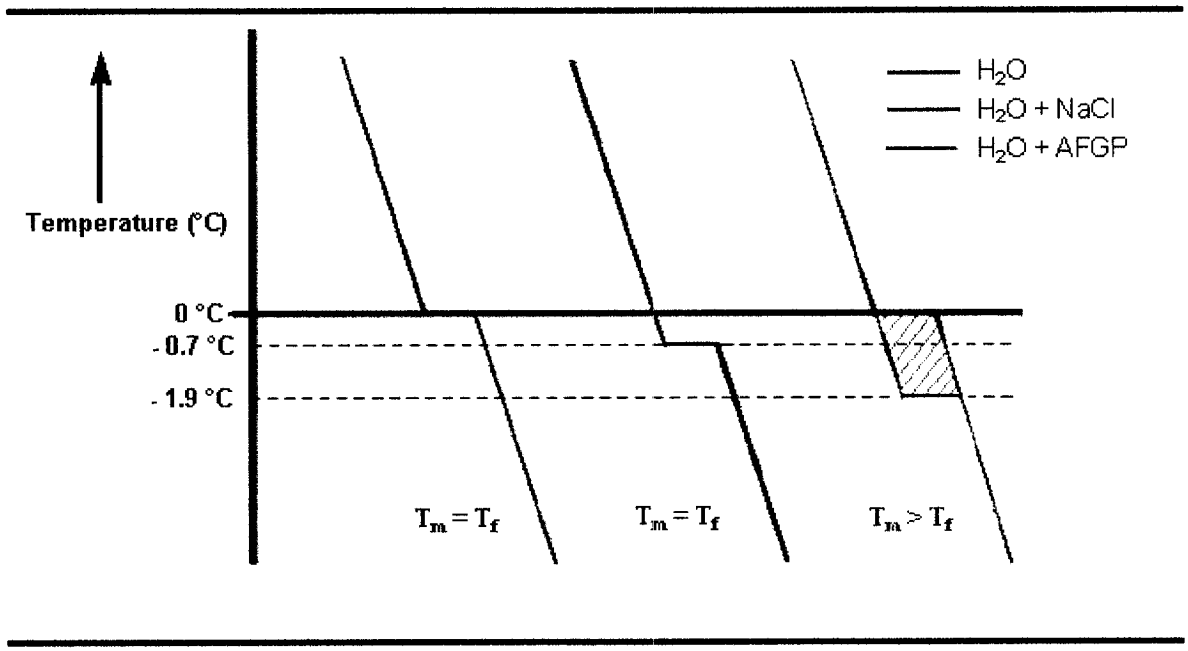
In order for Arctic and Antarctic teleost fish to survive this environmental challenge, biological antifreezes [antifreeze proteins (AFPs) or antifreeze glycoproteins (AFGPs)] are present in small quantities in the blood serum and enable the fish to survive in cold oceanic temperatures. These molecules possess the ability to depress the freezing point of the blood serum below the melting point in a non-colligative manner. The potency of these molecules predicted by the colligative effect is 500 times less than the potency that is actually observed.[7] Graph 2 represents the freezing point depression of biological antifreezes (AFGPs). In this graph chicken lysozyme and galactose are used to approximate the colligative effect of AFGP. Given that AFGPs are glycoproteins, their colligative effect on the freezing point depression should be approximated by measuring the colligative effect of a protein (chicken lysozyme) and that of a glycoside (galactose). As observed in Graph 2, AFGPs exhibit a much higher activity than expected based upon colligative effects.



Graph 2: Diagram illustrating the changes in freezing point of solutions that contain colligatively acting substances and AFGPs[7]

This non-colligative effect is unique to biological antifreezes and is known as thermal hysteresis (TH) (Graph 3).[8] The size of the TH gap, which is the difference between the freezing and melting points, is dependent on the type of AF(G)Ps. Biological antifreezes

found in insects are more potent and possess a TH gap of 6.5°C,[9] whereas antifreezes isolated from fish are not as potent and can only depress the freezing point by 2°C.



Graph 3: Qualitative Diagram of Thermal Hysteresis

Although Schölander and coworkers first reported and investigated antifreeze activity in marine fish,[10-13] it took 16 years for the molecule responsible for this process to be isolated and partially characterized. It was only between 1969 and 1971 that Devries and Wohlschlag published a series of articles,[7,14,15,16] where they described a glycoprotein isolated from the cold-adapted Antarctic fish *Trematomus bernacchil* and *Trematomus borchgrevinki* that depressed the freezing point temperature of the blood serum without increasing the osmotic pressure.[13,17,18,19] The isolated glycoprotein was only composed of alanine and threonine amino acids. At that point the exact structure and activity of the glycoprotein was not assessed; however, its mass was confirmed to be approximately 25 kDa with 20% of which being composed of D-galactosamine. Since then antifreeze glycoproteins (AFGPs)[20-24] and proteins (AFPs)[25-31] have been extensively studied. Today, eight antifreeze glycoproteins and five antifreeze proteins families have been identified and described.

1.2 Structure of Biological Antifreezes

There are two types of known biological antifreezes, antifreeze proteins (AFPs) and antifreezes glycoproteins (AFGPs). Even though antifreeze proteins (AFPs) share similar physical properties, their primary, secondary and tertiary structures drastically differ from one species to another. Conversely antifreeze glycoproteins possess very low structural diversity and mainly differ by their molecular weight.

1.2.1 AFGP [15,16,32]

The first structurally characterized AFGP was isolated from fish blood serum and it was found to consist of glycopeptide repeats. The peptide backbone was composed of alanine-alanine-threonine (Ala-Ala-Thr) repeating tripeptide units with minor sequence variation, while the glycoside moiety was a disaccharide unit β-D-galactosyl-(1,3)-D-*N*-acetylgalactosamine) linked 1-3 to the threonine residue (Figure 1). AFGPs have been isolated and characterized into eight different fractions based on their molecular weight ranging from 32 kDa for AFGP1 to 2.6 kDa for AFGP8 (Table 1). This classification was made on the basis of the relative rates of electrophoretic migration for the different fractions.[33] In some species it has been observed that in the low molecular weight AFGP7-8 the glycosylated threonine can be substituted by arginine[34,35,36] and that the alanine residues can be substituted by proline.[37,38]

Table 1: Molecular Weight of the different fractions of AFGPs

AFGP1-2	33.7-28.8 kDa
AFGP3-5	21.5-10.5 kDa
AFGP6-7	5-3.8 kDa
AFGP8	**2.6 kDa**

Figure 1 shows the most common AFGP structures. Some novel AFGPs recently isolated from Antartic fish (*Pleuragramma antarcticum)* have been shown to contain *N*-

acetyl-glucosamine instead of *N*-acetylgalactosamine as well as amino acids such as Asn, Gln, Gly, Ala, Val, Thr and Leu.[39]

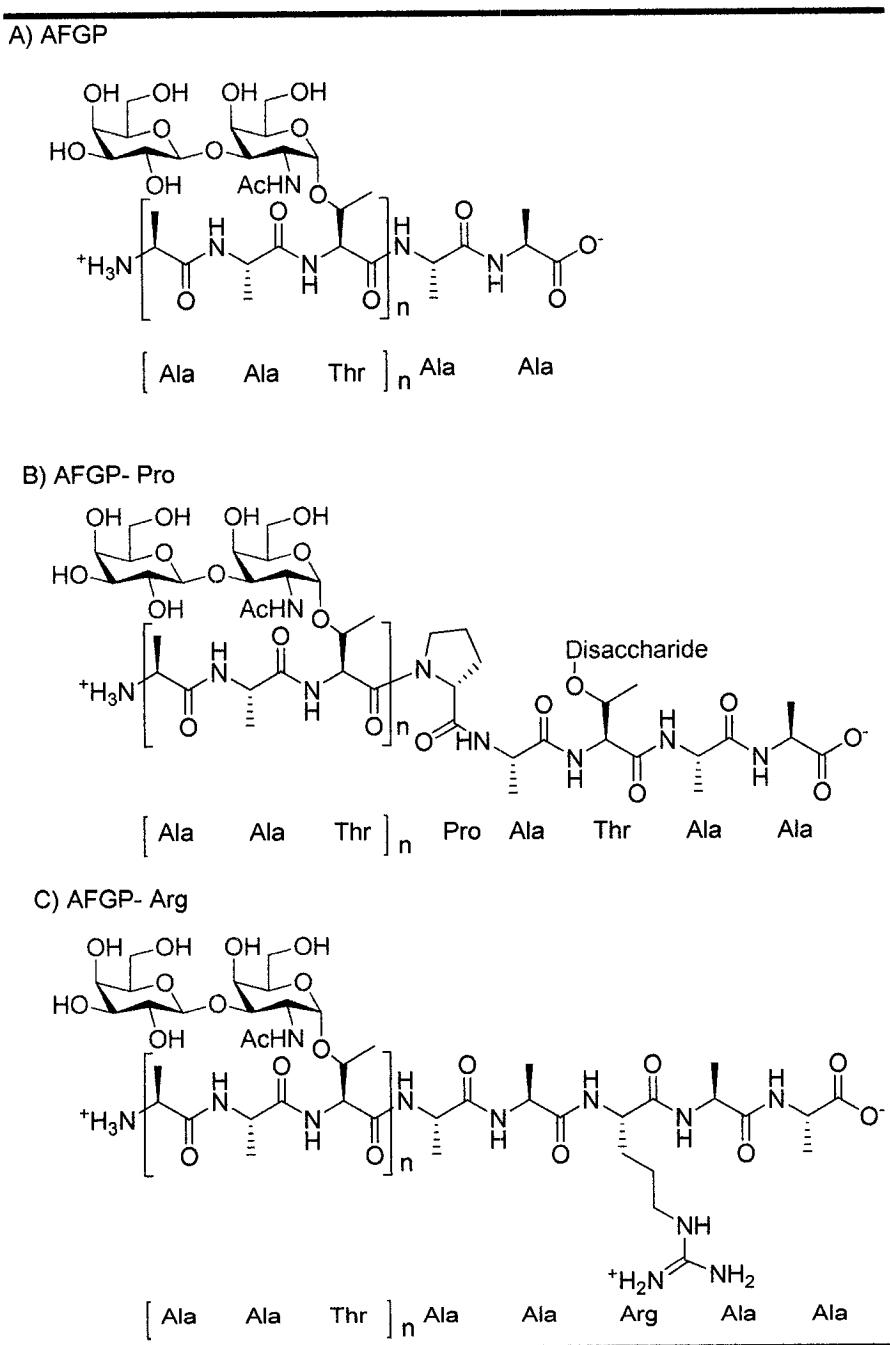


Figure 1: Examples of naturally occurring AFGP alterations[39]

A) -AFGP: most common structural motif with n = 4-50

*B) -AFGP-Pro in which Pro replaces Ala**

C) -AFGP-Arg in which Arg replaces Thr.*

**AFGP-Pro and AFGP-Arg constitute less than 5% of the naturally occurring glycoproteins. It is unknown whether the arginine is glycosylated.*

It has also been shown that AFGP6 from the rock cod *Gadus ogac* is comprised of 14 different isoforms (different peptide sequences).[40] All of these 14 characterized AFGPs have been isolated and range in size from 6026 kDa to 9784 kDa, thus the abbreviations AFGPx (x = 1-8) do not refer to a single compound, but a mixture of glycopeptides in a designated approximate mass range.

Recently Hsiao *et al.*[41] have succeeded in sequencing the genetic code of an AFGP from an Antartic nototheniid. Sequence comparisons of the genes encoding AFGPs have shown that the regions encoding the glycoproteins have evolved by repeated duplication and rearrangement of small fragments of other genes. The gene that the fragments were transcribed from varies between the species of fish, with no similarity in coding function. This suggests that different AFGPs were formed in independent evolutionary events, thus explaining the variations in structure observed in characterized AFGPs. AFGPs of various lengths most likely arise from micro-adaptations, where the activity that is being selected for is present in a small piece of DNA, which can be repeatedly duplicated to progressively yield longer and more potent biological antifreezes (*i.e.* AFGP1).[42]

1.2.2 AFP

There are five main subtypes of AFPs (I-IV; and unrelated AFPs) with massive variability at all structural levels. They range in size from 3 to 24 kDa and are usually found in the extra-cellular space. Some of the secondary and tertiary structural motifs are illustrated in Table 2. The diversity associated with their narrow distribution has led to the hypothesis that AFPs have evolved recently and independently, probably due to selective pressure such as glaciations.

Table 2: Overview of antifreeze protein characteristics[42]

Protein	Species	Structural type	Protein homology	Ice binding site
Fish type I AFP	righteye flounders, sculpins	single α-helix, many residues with sequence repeat	undetected	Thr and associated with it correspond to Ca^{2+} binding site residues on and flanking surface
Fish type II	sea raven, rainbow smelt, Atlantic herring	globular; smelt and herring are Ca^{2+} dependent	galactose-binding; C-type lectins in herring	
Fish type III	eel pouts such as ocean pout (*Macrozoarces americanus*)	globular with one flat flattened surface	undetected	
Fish type IV	longhorn sculpin (*Myoxocephalus octodecimspinosis*)	antiparallel helix bundle	low-density lipoprotein receptor-binding domain of apolipoprotein E3	unknown
Unrelated AFP				
Chitinase AFP	winter rye	endochitinase	endochitinase	unknown
Glucanase AFP	winter rye	endoglucanase	endoglucanase	unknown
Thaumatin AFP	winter rye	thaumatin-like	thaumatin	unknown
Budworm AFP	spruce budworm (*Choristoneura fumiferana*)	Thr- and Cys-rich (non-repeating)	undetected	unknown
Beetle AFP	*Dendroides canadensis*, mealworm beetle	Thr- and Cys-rich with repeating structure	undetected	unknown
Snow Flea AFP	*Hypogastrura harveyi*	unknown	unknown	unknown

1.2.2.1 Type I AFP

The existence of these proteins was first reported by Duman and Devries,[43,44] and Hew.[42] The most studied protein in this family has been the AFP extracted from the winter flounder *Pseudopleuronectus americanus*. The secondary structure (Figure 2) of this protein was determined using circular dichroism (CD) and it was shown to be mainly α-helical. When the temperature is decreased to 0°C, the percentage of secondary structure that exists as an α-helix increased to 85%. X-ray diffraction data also shows that the residues implicated in ice binding are localized on one face, while the non-polar residues (such as alanine) are mainly on the other face.

1.2.2.2 Type II AFP

Type II AFPs[45,46] have been isolated from sea raven, insects, spruce budworm and smelt. This family of antifreezes (Figure 2) is relatively rich in cysteine (≈9 mol %) and alanine (≈14 mol %). Even though they exhibit a comparable Thermal Hysteresis activity

(TH), their efficiency by unit weight is significantly lower than for AFP type I. The secondary structure of this AFP mainly consists of random coils, where the disulfide bonds created by the cysteines expose the hydrophilic residues (threonine) to the aqueous environment.

1.2.2.3 Type III AFP

Little is known about type III AFPs.[47,48] Indeed, they have only been isolated from Arctic and Antarctic Ocean pout. These short proteins (62-66 amino acids) are not dominated by any particular amino acid and possess some unusual folding (Figure 2). They are composed of β-sheet structures that are orthogonally arranged resulting in a globular β-sandwich. Due to their high hydrophobicity factors, their adsorption process on ice is probably different from the previous AFPs.

1.2.2.4 Type IV AFP

Type IV AFPs[49,50] have only been isolated from longhorn sculpin (*Myoxocephalus octodecimsponosis*) and possess a rich alanine and glutamine content (Figure 2).[30] This protein contains four α-helical bundles and is approximately 12 kDa in molecular weight. The amino acid sequence and the structure are similar to the lipoprotein receptor-binding domain of apolipoproteins.

Type	**Classification**	**Size(kDa)**	**Repeat**
Fish			
I	Ala rich α helix	3-5	11 aa(-3 turn of helix)
II	C-type lectin fold of mixed α, β and loop structure	14-24	None
III	Globular protein contains short β strands	7	None
IV	Helix-bundle	12	22 aa
Insects	*Unrelated AFPs*		
Tm and *Dc*	Right handed β helix	8-9	12-13 aa (containing Thr-Cys-Thr)
Cf	Left handed β helix	9-12	15 aa (containing Thr-Xaa-Thr)

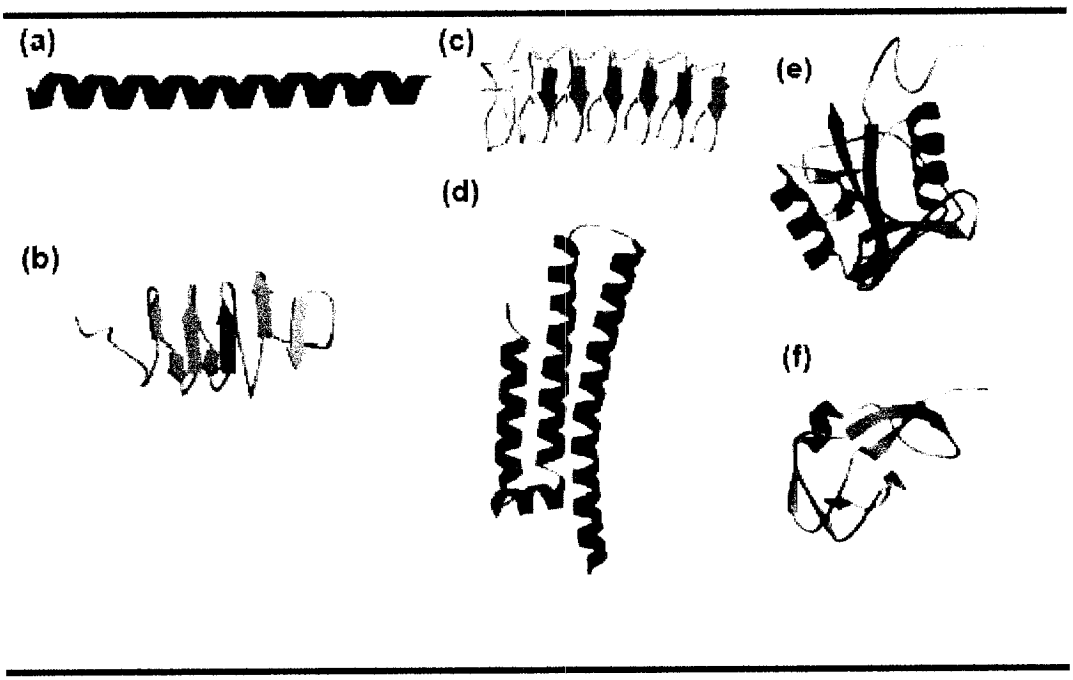


Figure 2: Representative structures of four types of fish (a,b,e-f) and insect (c,d) antifreeze proteins[30]
(a) Type I AFP from winter flounder. (b) Spruce budworm Type II AFP. (c) Tenebrio molitorTypre III AFP. (d) Type IV AFP modeled as α helix bundle. (e,f) Non-repetitive AFPs. (e) Type II AFP from sea raven (2AFP). (f) Type III AFP from ocean pout. Structures were generated using Molscript

1.2.2.5 Other AFPs

Recently other types of AFPs, which are not members of any of the previous groups, have been isolated. For instance, two unrelated threonine-rich AFPs of moths[42] and beetles[30] as well as glycine-rich AFPs of snow fleas[9] have been discovered in the past 5 years (Figure 2). These recent discoveries indicate that AFPs may have evolved in arthropods after speciation, and it is anticipated that other unrelated AFPs will be found in insects, and other arthropods.

1.2.3 In-Vivo localization of AF(G)Ps

AF(G)Ps are usually isolated from the blood plasma or serum; however, they are also present in other tissues. Biological antifreezes are produced in the liver of the vertebrate over the winter, and are circulated to all fluid compartments except the brain.[51,52] It has also been speculated that the synthesis of AF(G)Ps does also occur in other tissues. This theory

has been recently confirmed by Valerio,[53,54] who has isolated AF(G)Ps from epithelium cells, demonstrating that they also had the ability to synthesize these molecules.

More specifically AFGP6, 7 and 8 are usually secreted into the intestinal fluids and are present in ocular fluids at very low concentrations. It was thought that the localization and the activity of these antifreeze glycoproteins were limited to the extra-cellular media, however, recent studies have shown that this was not the case and that AF(G)Ps could be internalized.[42] The role of this internalization is not yet clear, but the importance of their intracellular properties could be significant for the survival at subzero temperatures.

1.3 AF(G)Ps interaction with ice and cell membrane

In order to fully understand the advantages of using biological antifreezes for cryopreservation, we will describe their different physical and biological properties as well as how they interact with the ice lattice. First, they possess the ability to depress the freezing point of water, while the observed melting point remains close to 0°C. The difference in temperature between the freezing point and the melting point is known as the thermal hysteresis (TH) gap. This property is non-colligative and a saturation effect is observed at higher concentrations.

Along with thermal hysteresis, a second property is observed. Biological antifreezes have the ability to affect the rate of the growth of the ice crystals, as well as their morphology, at a supercooled temperature. This phenomenon is called dynamic ice shaping (DIS) and results in the formation of needle-shaped ice crystals.

The final physical property associated with AFGP and ice interaction is recrystallization inhibition (RI) activity. RI activity is probably one of the most interesting properties of biological antifreezes. While crystals are formed, the antifreeze molecules have the ability to inhibit thermodynamic boundary migration that leads to the formation of larger ice crystals. This results in stabilization of small crystals, and limits cell damage during cooling and thawing processes in cryopreservation. In addition to all these physical properties, it has been demonstrated that AFGPs possess the ability to stabilize the cellular membrane during the thermotropic transition phase.

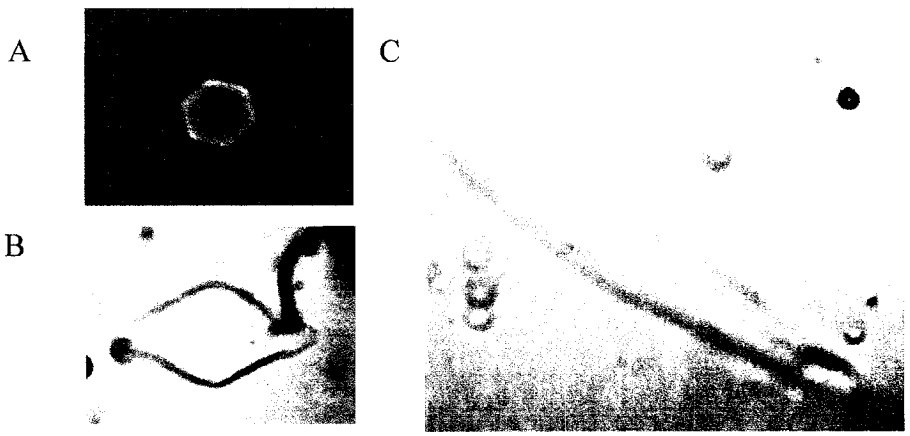
1.3.1 Thermal hysteresis (TH) and dynamic ice shaping (DIS)

On the macromolecular level, the most accepted theory of the mechanism of action of the antifreeze molecules is an irreversible adsorption inhibition process.[19] However, several other interesting theories have been explored. For instance, these molecules could be inhibitors of nucleation[55,56,57] and their adsorption onto the ice crystals could be reversible.[58]

Kuroda[58] was the first one to suggest that the adsorption might not be totally irreversible. Using Langmuir equation, he showed that if the adsorption inhibition was irreversible, the melting point should increase. This phenomenon was reported afterward by Knight and Devries; however, the range is slightly different than the calculations predicted. Along these lines, a fairly recent NMR study suggests that the adsorption would be reversible, but here again these results are highly debated.

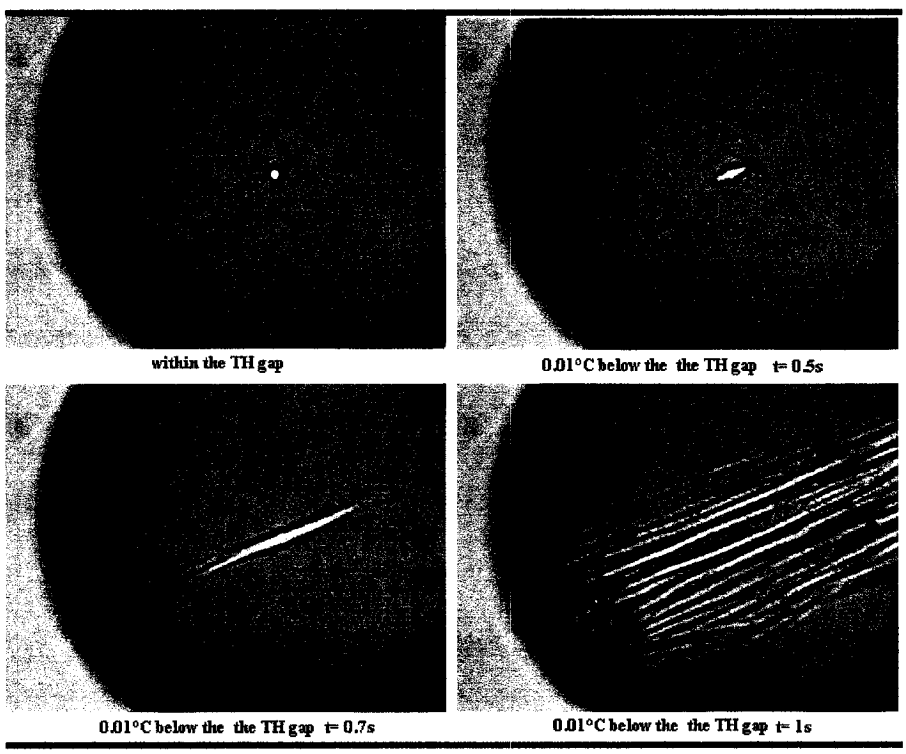
As the ice front grows, solutes are excluded from the ice lattice, therefore concentrating the intercrystal solution. Using etching technique,[59] antifreeze molecules have been observed to adsorb and integrate onto the ice crystal lattice. This adsorption[60-64] prevents ice growth within the TH gap.[65] However, Devries work demonstrated that ice crystals have the ability to resume growth outside of this TH gap. The morphology and the dynamics of the growth are entirely different from a pure ice-water system, and form spicules of ice that grow perpendicular to the basal plane (Picture 1).[66,67] These needle shaped ice crystals "burst" when they are formed (Picture 2). This ability to form needle shaped ice crystals[19] is detrimental for cryopreservation, but can be extremely useful for cryosurgery.

It has been demonstrated that AF(G)Ps affect the direction of growth because of the adsorption-inhibition along certain prism faces. Knight *et al.*[66,67] showed that the adsorption of AF(G)Ps were preferentially on the a-axis of the ice crystal during the growth (outside the TH gap), resulting on the formation of a hexagonal bipyramidal ice crystal (Figure 3).



Picture 1: Dynamic Ice shaping of AFGP8

Picture A shows the dynamic ice shaping of AFGP8 within the TH gap. (top view of a hexagonal bipyramidal crystal).
Picture B is the same ice crystal when the temperature reaches the limit of the TH gap (hexagonal bipyramidal crystal).
Picture C is the same crystal outside the TH gap (needle shaped crystal).



Picture 2: Needle shape ice crystals formation in AFGP8 solution

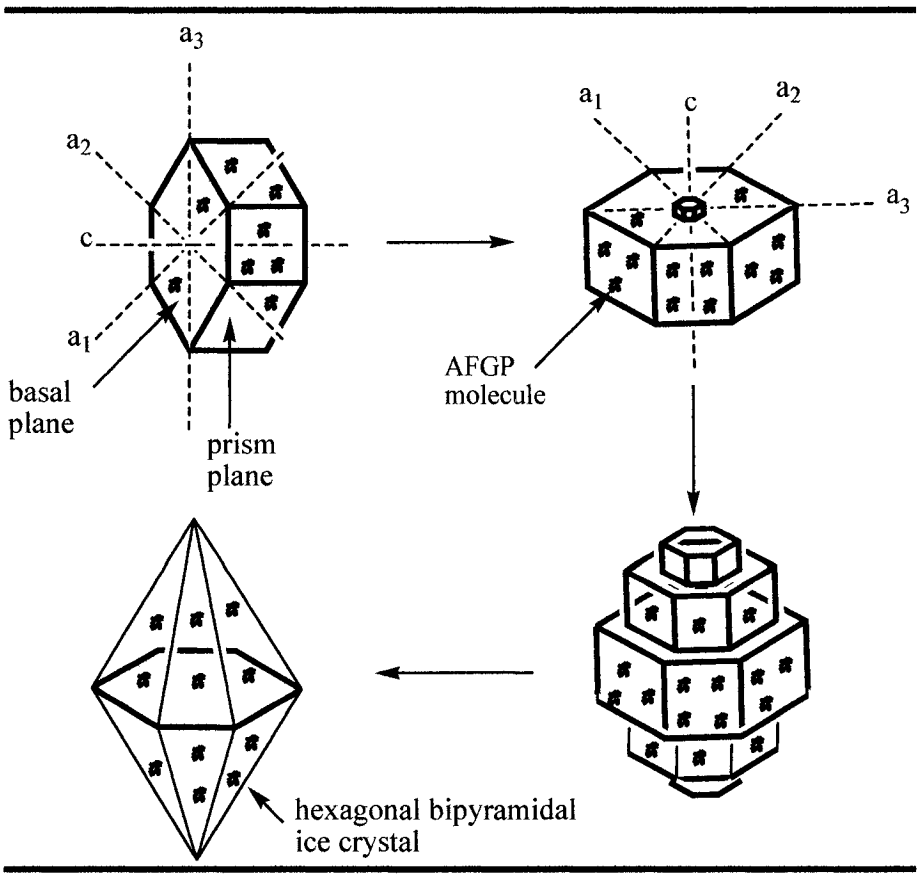


Figure 3: Dynamic ice shaping

Within the TH gap, the AF(G)Ps adsorb onto all planes of the ice crystal and stop the growth along all these planes, resulting in a hexagonal crystal. Below the TH gap the growth of the ice crystal is still inhibited along the a_1-a_2-a_3 axis where the adsorption is higher[68] (shown by ellipsometry measurement on AFGP7 and AFGP8), but not anymore along the c-axis; leading to the formation of a new prism whose growth will be more inhibited along the a-axis than the c-axis, ultimately leading to the formation of an hexagonal bipyramidal ice crystal.

Collectively, these experiments (i.e. ice etching, ice hemisphere) confirm that in order to inhibit the growth of ice crystals, AFGPs have to adsorb on the surface of growing ice crystals. However, understanding how the adsorption can depress the freezing point is still in progress. Raymond and Devries[19] were the first to mention that if a molecule of AF(G)P lodges onto the crystalline surface, the distortion created at that point will slow the growth of the ice crystal by creating a high radius of curvature where the addition of a new molecule

will be thermodynamically unfavourable. This results in a non-equilibrium freezing point depression known as the Kelvin Effect (Figure 4).[69] Several models have proposed to rationalize the ice growth inhibition.

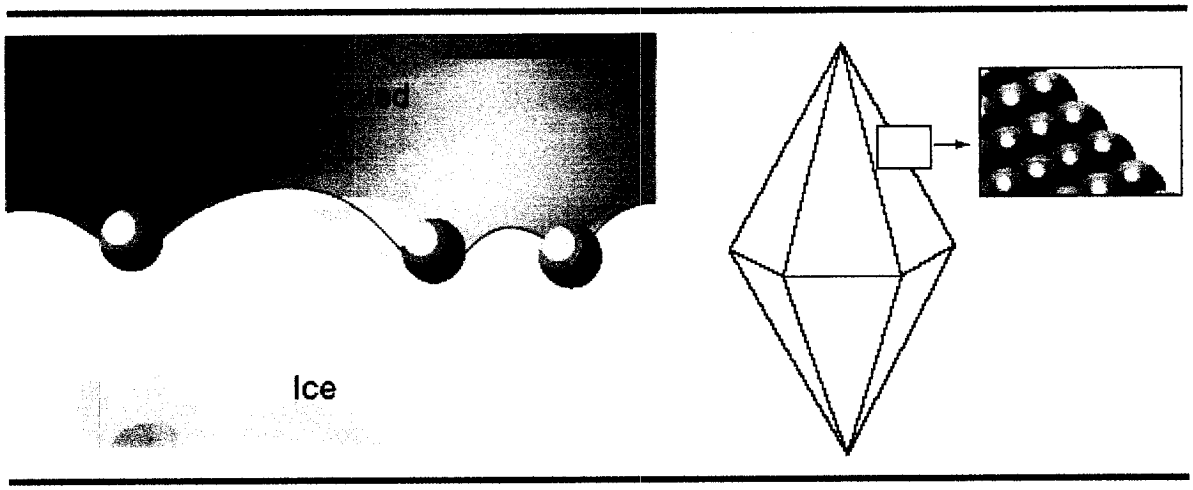


Figure 4: Diagram of the adsorption-inhibition effect[30,70]

There are two models that rationalize ice growth inhibition based on a 2D or 3D Kelvin Effect. These are the mattress model[66] and step pinning model.[69] In the mattress model (Figure 5), the adsorbed molecules prevent ice growth perpendicular to the ice surface, while in the step pinning model the molecules block the growth of a step. Both models assume irreversible adsorption of AF(G)Ps onto the ice surface.

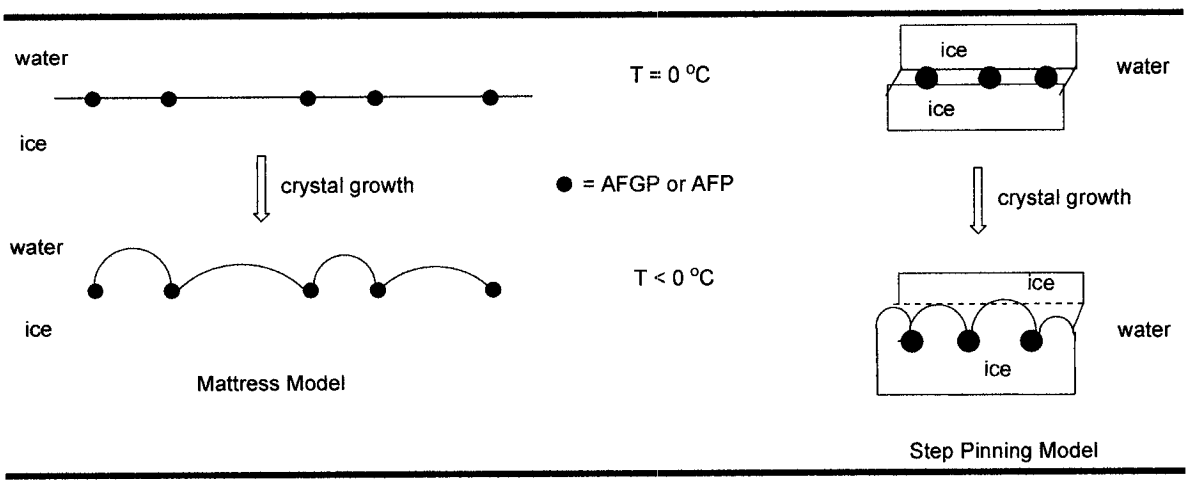


Figure 5: Adsorption inhibition mechanism[69]

It has been shown by ellipsometry[68] that the adsorption of AF(G)Ps onto ice crystals is not concentration dependant (contrary to the TH). However, neither model can explain why high levels of adsorption are not observed at low concentrations. Consequently, alternate mechanisms have been proposed. For instance, Lips[71] attributes the role of the antifreeze to lowering the step energy associated with the formation of nuclei and reducing the nucleation temperature.

Presently, the scientific community has an overall acceptable macromolecular mechanism of action. Nevertheless at the molecular level, an understanding of how these molecules inhibit ice crystal growth and how they are able to match the ice-water interface remains a source of intense debate. Interactions between biological antifreezes and the ice surface have been postulated to be hydrogen bond-based. However, modeling experiments have demonstrated that the number of hydrogen bonds formed between the antifreeze molecule and the ice surface appears to be insufficient to explain the observed tight binding of the antifreeze to ice[67] even if the hydroxyl groups of the glycoproteins are imbedded in the ice lattice (Figure 6).

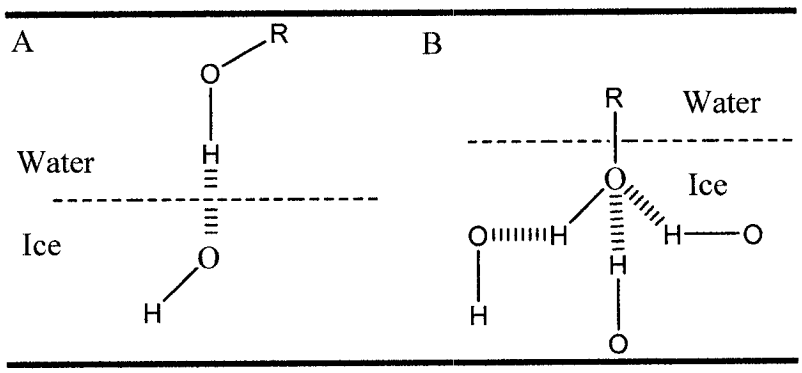


Figure 6: Theoretical hydrophilic ice binding of AF(G)Ps

A: First theory explaining the adsorption of AFGPs on the ice lattice. Two hydroxyls of the disaccharide unit can form hydrogen bonds with the ice crystal. Therefore, in this first model eight hydrogen bonds could be formed for AFGP8.

B: After different AFM experiments as well as modeling calculations, it has been shown that eight hydrogen bonds were highly insufficient to explain how tight the ice adsorption was. Consequently Devries[67] hypothesized that the hydroxyls of the disaccharide could be imbedded in the ice lattice affording three hydrogen bonds by hydroxyl. Even though 24 hydrogen bonds could be formed using this model for AFGP8, the 192 Kcal/mol involved in this adsorption process seems to be insufficient.

More recently, researchers have been divided over the importance of hydrogen bonding and its role in the mechanism of action of biological antifreezes. While it has been

proposed that the hydrophilic interactions between polar residues and the water molecules on the ice surface are extremely important, other researchers have invoked the idea that entropic and enthalpic contributions from hydrophobic residues[72-76] are essential for adsorption onto the ice surface. In order to bind to ice, biological antifreezes have to match the topology of the ice-water interface. Several experiments performed on type I AFPs[66,77,78] and extrapolated to AFGPs suggest that an α-helix structure where the hydrophilic group would be on the side exposed to the ice lattice and the hydrophobic would be exposed to the to water front, would be significant. The frequency of the α-helix turn would match different planes of the ice (Figure 7).[63] Despite the fact that significant entropic contributions are likely to be gained upon exclusion of water from the protein and ice surfaces, a detailed molecular mechanism invoking hydrophobic and/or hydrophilic interactions with emphasis on the role they play in adsorption of the antifreeze to the ice surface has failed to emerge using a static ice water interface and an irreversible adsorption process. Figure 7 shows the profile of the (2021) surface viewed from the a_2-axis. The turn of the α-helix allows an optimized position for hydrogen bonding between AFPI and the ice lattice. Indeed the Thr-Asn intervals (which afford the hydrogen bondings) are 16.7Å and the oxygen atoms at the ice facet are also separated by 16.7Å.

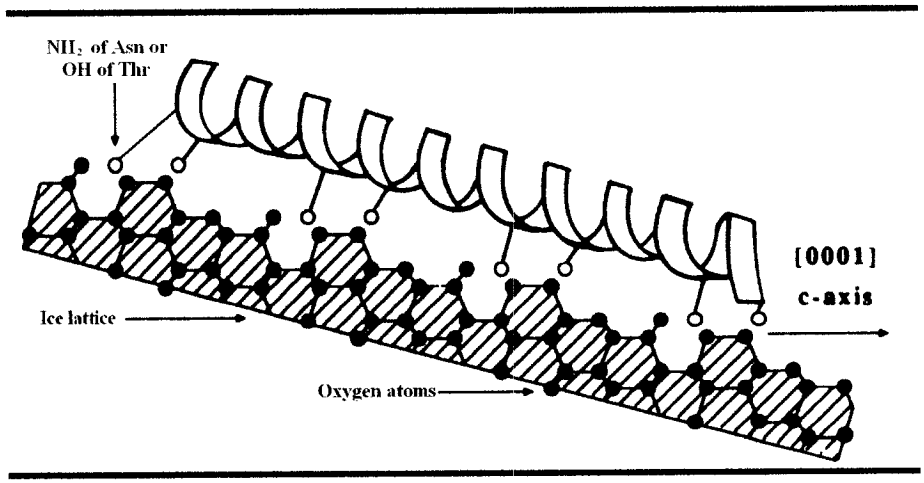


Figure 7: Helix conformation favouring hydrophilic interactions between AF(G)Ps and the ice lattice[63]

Most of the theories described previously make two assumptions. The first is that the adsorption of the AF(G)Ps molecules on ice lattice is irreversible and the second is that the

ice-water interface is an abrupt transition and this interface is static. In order to solve the puzzling problem of the molecular mechanism of action other concepts have been explored.

Indeed, to elucidate the molecular mechanism of action for AF(G)Ps it is essential for the ice-water interface to be well characterized. This is unfortunately not the case yet but several improvements have been made. Between the highly ordered ice lattice and the liquid water is a semi ordered dynamic ice layer spanning 10Å.[79,80] Most of the calculation studies made to investigate the adsorption inhibition process used a static model and therefore are highly inaccurate. Even though static models and docking techniques are useful to investigate specific binding interactions as well as the enzyme binding site, this model is oversimplified for these interactions. Despite the limitations due to the amount of calculations required by the iterative relaxation methods in order to calculate the quasi plastic water ice dynamic, recently 10 ps of this interface have been modeled.[81] These results are promising and further progress have been achieved to improve the ice-water model,[82] however, structure activity relationship studies seem to be the more adequate technique to obtain a better understanding of the ice-water-antifreeze interactions.

1.3.2 Recrystallization Inhibition (RI) Activities of AFGP's

When a frozen sample of water is left standing at sub-melting temperature or is warmed up above the melting point, the grains of ice will change in shape and size. Smaller grains with negative curvature are swallowed by larger, positive curvatured grains, to give a sample with larger average grain size (Figure 8).[83,84] This migration of the grain boundary occurs to minimize the overall grain boundary energy by increasing the surface area. There are two ways to inhibit the recrystallization;[85] altering the interface energy or decreasing the diffusion of the water. When a solution is flash frozen at -80°C, a sample of relatively small crystals is created. This process of crystallization extrudes all of the solute out of the ice crystal, creating highly concentrated boundary solutions. It has been shown that low concentrations of AFGP molecules have the ability to drastically inhibit the boundary migration of ice crystals (Figure 8). It has been argued that due to the small amount of water present at the interface, AFGPs can migrate and adsorb to the required binding site on the ice surface and therefore alter the interface energy. Knight *et al.*[86] demonstrated that some peptides that do not have antifreeze activity can also inhibit ice recrystallization by

decreasing the diffusion of water. When the solute concentration between the ice boundaries increases, these peptides will decrease water mobility; therefore affecting the reorganisation of the ice crystals and inhibiting recrystallization. Unlike these peptides, the RI activity of AFPGs is significant even at low concentrations.

Until recently, it was thought that both thermal hysteresis and recrystallization inhibition were due to the same process of adsorption on the ice crystal because they are both thought to be a direct consequence of immobilization of solid-liquid interfaces in partially frozen samples.[70] However, recent research published by Sidebottom,[87] as well as results from some of our AFGP analogues, suggests that these two properties are uncoupled.

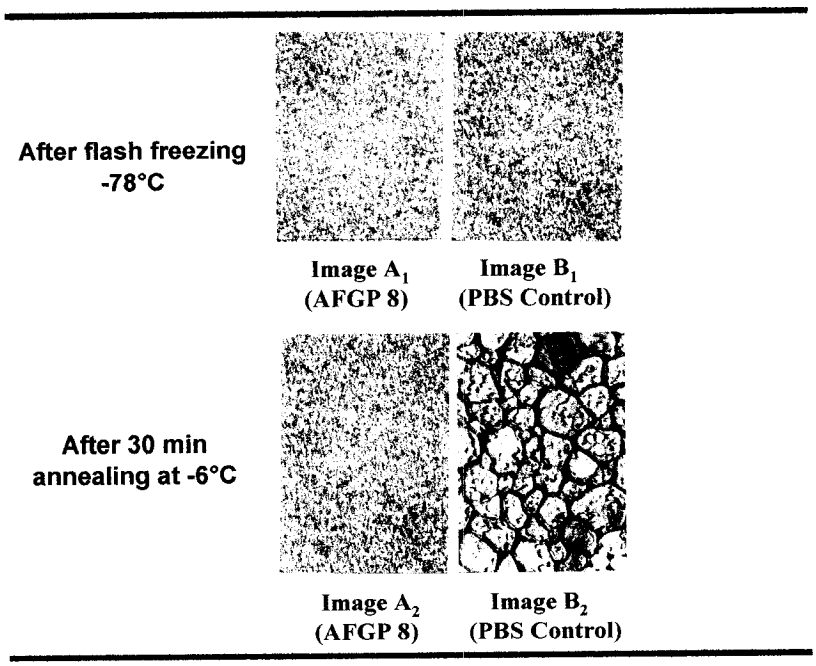


Figure 8: Example of Recrystallization-Inhibition

1.3.3 Membrane stabilization

Not only do AFGPs possess RI and TH activities, they also exhibit interesting biological properties which enhance cell survival at subzero temperatures. Rubinsky and coworkers[88,89] observed that AFGPs and AFPs possess the ability to enhance external cell membrane integrity during cooling, and protect it against rupture.[90] When bovine and porcine oocytes were cooled and re-warmed in the presence of AF(G)Ps, the cell membrane

retained its membrane potential. Schmitt *et al.* suggested that membrane integrity was enhanced because AFGP blocked the potassium and calcium ion fluxes across the membrane. Thus, it was inferred that biological antifreezes might be directly interacting with membrane-bound ion channels. This suggestion has met with some controversy.[91,92] This controversy has been amplified by the fact that AF(G)Ps have failed to enhance the preservation of rat hearts as well as to stabilize ram spermatozoa when chilled and re-warmed.

Several years later, in an effort to understand the apparent improvement in cell survival induced by the presence of AF(G)Ps, a study by Hayes *et al.*[93] confirmed that AF(G)Ps enhance membrane integrity during cooling. This work was performed using liposomes prepared from four different types of phospholipids (all phosphatidylcholine derived) that contained carboxyfluorescein (CF) probes. In these experiments it was observed that liposomes leaked up to 50% of the trapped marker as they were cooled through the thermotropic phase transition in the absence of AF(G)Ps. During this transition state (typically between 12°C and 41°C), both crystalline and liquid states of the lipid membrane coexisted.[94,95] This phenomenon provoked a mismatch of the hydrocarbon chains and ultimately facilitated leakage of the lipid membrane at the connection between the crystalline and the liquid state (Figure 9).

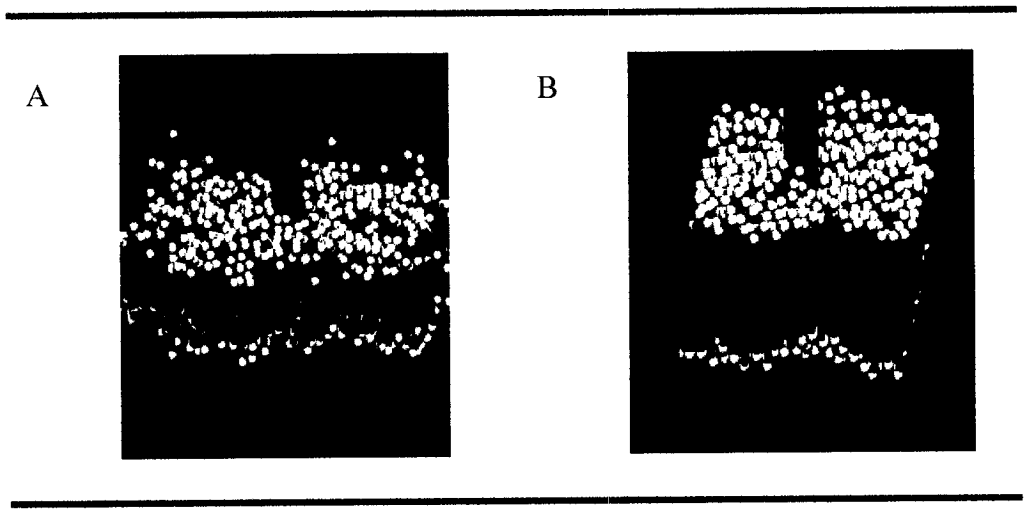


Figure 9: Modeling of a lipid bilayer membrane at different temperatures[96]
A: Lipid bilayer membrane in a liquid state at 30°C, B: Same lipid bilayer membrane in a crystalline state at 8°C. At optimum physiological conditions the membrane is in the liquid state.

In the presence of less than 1 mg/mL of AFGP, no leakage was observed during cooling or warming through the thermotropic phase transition. This suggested an alternate mechanism consistent with a nonspecific interaction with the liposome membrane.

In 2000, Wu and Fletcher[97] expanded the study by Hayes *et al.* and examined interactions of AFP type 1–3, AFGP8, and albumin interactions with liposomes as model cell membranes. Because most cell membranes are more complex than a phosphatidylcholine (PC) liposome, they prepared liposomes derived from dielaidoylphosphatidylethanolamine and dielaidoylphosphatidylglycerol. AFGPs were found to be extremely effective at preventing leakage from the liposomes as each was cooled through its respective thermotropic phase transition. It was concluded that AFGPs may interact with the lipid bilayer in one of two ways. NMR studies have suggested that AFGP adopts a threefold left handed helical conformation, with the carbohydrate moieties lining up on one side of the helix and the hydrophobic alanine residues lining up on the other.[98,99] Given this arrangement, the hydrophilic carbohydrate moieties might interact with the polar head groups of the lipid bilayer (Figure 10). Alternatively, they hypothesized that the hydrophobic backbone of AFGP may partially immerse in the lipid bilayer and that this may be sufficient to prevent disruption of the bilayer. These studies were instrumental in demonstrating that most proteins interact with DEPC membranes and actually prevent leakage upon binding. However, the ability of a protein to bind to the lipid bilayer and prevent leakage is a direct function of the nature of the polar groups of the lipids. Indeed the composition of the lipid bilayer seems to control whether or not the AF(G)Ps will insert and stabilize the membrane.

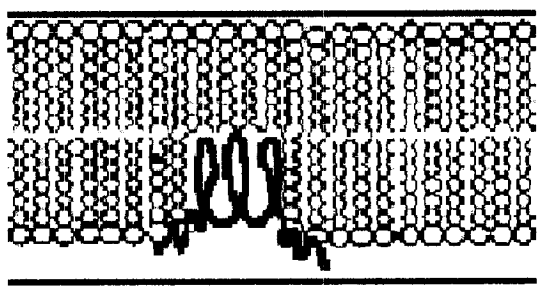


Figure 10: Membrane stabilization by antifreeze glycoproteins[99]

1.4 Applications

Any organic compound with the ability to inhibit the growth of ice has many potential medical, industrial and commercial applications. AFGPs have potential uses in all of these areas, and can also be used for cryopreservation using their membrane stabilization properties and their propensity to facilitate the vitrification process.

1.4.1 Frozen food preservation.

One promising field is certainly the different applications for food preservation. [100,101] Due to their propensity to preserve cell integrity, AF(G)Ps can diminish the manufacturing costs, preserve more homogeneous products, lower extensive heat processing, and improve the quality of food preservation. However, the availability of AF(G)Ps by extraction or synthesis is low. Therefore different patents have been deposited for DNA recombination technology by direct engineering of micro-organisms which do not need purification prior to their addition to a food product. Consequently, many foodstuffs which are sensitive to low temperature like fruits or vegetables, can be frozen without any depreciation of quality. This may make them ideal for addition to food either by direct addition or by gene transfer, however, the type of AF(G)P will have to be carefully chosen. First, the AF(G)Ps would have to be sufficiently stable to survive the processing temperatures. Secondly depending on the product to be preserved, the choice of AF(G)P as well as its concentration would have to be optimized.

1.4.2 Cryopreservation (slow freezing rate)

Cryopreservation applications include improved protection of blood platelets and human organs at low temperatures.[102] However, there are many complex processes that occur when the ambient temperature of a cell is lowered close to or below that of the freezing point. Presently, three distinct modes of cell death are known to occur upon freezing. These are physical cell rupture, necrosis, and cold induced apoptosis.[103] Although all three processes are significant, the most common form of cell death associated with

cryopreservation[104-108] is cell rupture owing to fluctuating cell volumes and intracellular ice formation.

1.4.2.1 In vivo physiological mechanisms during chilling

The first phenomenon which occurs upon chilling is the transition between a relatively dynamic lipid bilayer membrane and the crystal state of this same membrane (Figure 9). During this thermotropic transition phase, the membrane is composed of both liquid state and crystal states. At the interconnection of these two states, some mismatch between the phospholipids in the membrane can occur.[109] These mismatches result in leakage of the membrane and exclusion of the intracellular content, which ultimately leads to the cellular death. This transition state usually occurs between 40°C and 12°C depending on the composition of the membrane.

All cells are regarded as compartmentalized systems and the probability of ice nucleation is directly proportional to the degree of supercooling and volume. Consequently, as the temperature is lowered, ice nucleation is more likely to occur outside of the cell since the volume is greater and the concentration of colligatively acting substances (salts, proteins, and so forth) is lower than inside the cell. After nucleation occurs, extracellular ice growth results in an increase in solute concentration in the diminishing extracellular volume. As the concentration of these solutes increases, extracellular osmotic pressure increases.

The rate at which the extracellular osmotic pressure increases is directly proportional to the rate of supercooling, and the cell compensates for this osmotic flux by regulating the flow of water through the semi-permeable cell membrane. When the rate of extracellular ice growth is fast, the cell membrane is likely to be fractured due to water depletion or disruption of the cell membrane. This is caused either by the osmotic flux or by the large mechanical stress applied by the ice crystals on the cell membrane. Once this fracturing occurs, intracellular ice formation occurs and ultimately the cell is destroyed. Researchers have debated whether the ice or elevated salt concentration is responsible for damage during the progressive freezing of cells. Work by Mazur *et al.*[109] demonstrated the effect of the cooling rate on water transport during progressive cooling, and they correlated this with cell survival. Ultimately, this leads to the "two factor hypothesis," which states that solute

damage occurs at low-cooling rates where extracellular ice formation is innocuous to the cells. Conversely, intracellular ice formation at high cooling rates is generally lethal because the membrane permeability of each cell type can vary dramatically; this means that each cryoprotectant is associated with an optimum cooling rate and a type of cell.

If the temperature is decreased further, ice crystals will nucleate in the intracellular media, disrupt all physiological processes, and trigger cellular death. Finally, once the sample is entirely frozen, upon storage or thawing, recrystallisation can occur, increasing the size of the ice crystals and applying tremendous mechanical stress on the cell membrane. This ultimately leads to cellular death (Figure 11).

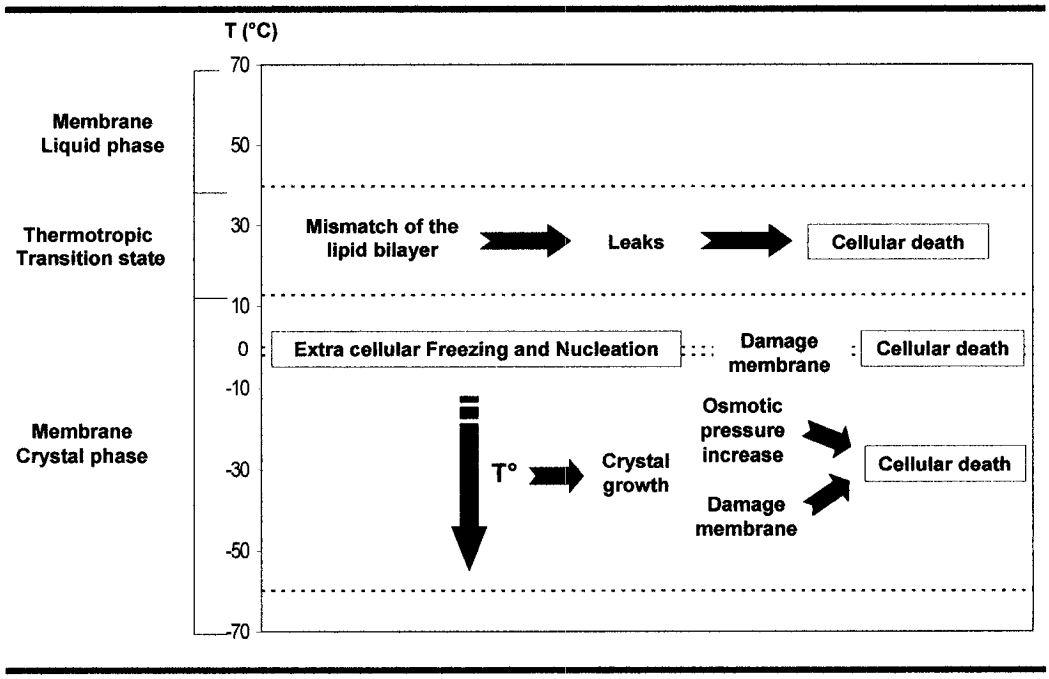


Figure 11 : In vivo temperature transition without cryoprotectants

The issues surrounding the cryopreservation of multicellular systems, such as tissues and organs, are infinitely more complex given that many cell types exist and each will differ in the requirements for optimal preservation. Because extracellular ice formation is likely to be the most significant obstacle in the cryopreservation of multicellular systems, researchers have focused on the use of various cryoprotectants that possess the ability to regulate the formation of extracellular ice.

1.4.3 Vitrification (fast freezing rate)

Vitrification is a common technique employed for the freezing of cells and tissues. In this process a sample is frozen under "ice-free" conditions that are achieved as the solution transforms into a "glass-like" solid. This is an attractive technique in that it can be used to avoid cell damage from ice formation. Indeed, by using a very fast freezing process the transition between the thermotropic and the nucleation steps is very fast (between 13 to 18 μs for the nucleation process compare to several minutes for a slow freezing process), limiting the amount of damage dealt to the cell (Figure 12).

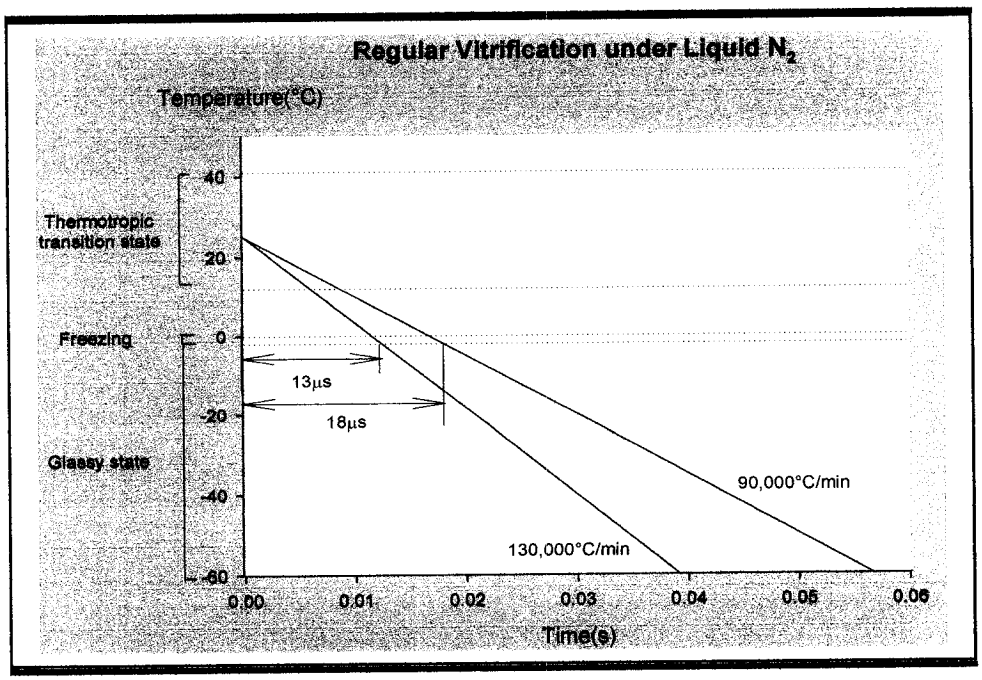


Figure 12: Typical fast freezing and vitrification profile

Despite the attractive features of vitrification, it does possess some limitations. For instance, when dealing with larger volumes, the heat transfer in cells, tissues and organs does not permit vitrification without the risk of crystallization. Consequently, a slow freezing process (0.5–100°C/min cooling rate) is often applied for preservation of large volumes.

In the vitrification process, a very rapid cooling rate (24–130,000°C/min)[110] is applied, resulting in a glassy or vitreous state which is dependant on the concentration, viscosity, volume and cooling rate of the process. Although this high-cooling rate minimizes

cellular damage (Figure 12), recrystallization can still occur during warming. To avoid recrystallization, rapid and uniform warming can be achieved using microwaves.[111,112,113]

Two problems unique to vitrification are encountered when using the methodologies outlined above. First, the necessary cryoprotectant concentration is very high and is sometimes too toxic for the cell. Although the process is feasible with lower concentrations of cryoprotectants, a higher cooling rate must be used to achieve a vitrified state. Second, some fractionation of the glassy state still occurs and the degree to which this occurs is dependant on the volume and the cooling rate. Some of the most promising cryoprotectant solutions incorporate antifreeze glycoproteins.[114] Not only do they inhibit ice formation at any subzero temperature, they also interact with the cell membrane during the thermotropic phase transition state to avoid leakage, thereby addressing two issues at once. The successful post-thaw revival is influenced by many factors that include the nature and the concentration of the cryoprotectant, temperature, the speed at which the cryoprotectant is added and removed, the rate of cooling and warming, and the storage temperature.[114-118]

1.4.4 Cryosurgery[119-122]

One of the lastest discoveries on AF(G)Ps is their potential application in cryosurgery. This surgical technique for the treatment of solid tumours has the advantage of being minimally invasive. Moreover since the arrival of ultrasound and magnetic resonance imaging this technique has been enhanced. Nevertheless before the discovery of AFP a major problem was still present. Even at high subzero temperatures some cells were able to survive. The presence of AFPs during the freezing process significantly increases cellular destruction by enhancing the formation of needle shaped ice crystals which perturb the cellular and nuclear membranes. In order to confirm these results *in vivo*, professor Rubinsky's team used a diluted injection of AFP I in the tumor and observed, that contrary to the untreated tumours, the latter have a mortality close to 100%. These encouraging results demonstrate that antifreeze proteins as well as antifreeze glycoproteins have a promising potential as effective chemical adjuvants for cryosurgery.

1.4.5 Advantages and disadvantages of using AFGPs as a cryoprotectant

Due to the accidental discovery, that glycerol enabled fowl spermatozoa to survive freezing at –70°C, much effort has been devoted in developing improved cryoprotectants and preservation techniques. In the presence of AFGPs, the membrane is stabilized during the thermotropic transition steps, which increases the survival of the cells (Figure 10). Moreover, if the required temperature of preservation is within the TH gap, the size of the crystal is kept under control (due to the TH activity), thus protecting the cell within this temperature gap.

If the temperature is decreased below the TH gap, several phenomena occur. First, antifreeze molecules will disrupt the growth and the shape of ice crystals. This dynamic ice shaping generates needle shaped ice crystals, which burst at subfreezing temperature and create genuine cryotoxicity. This is one of the main disadvantages of using AF(G)Ps as a cryoprotectant. Second, as the temperature further decreases, AF(G)Ps stabilize the membrane against osmotic shock caused by extracellular ice growth. Finally, due to their recrystallization inhibition properties, they also protect the sample during long-term storage by inhibiting the growth of ice crystals (Figure 13).

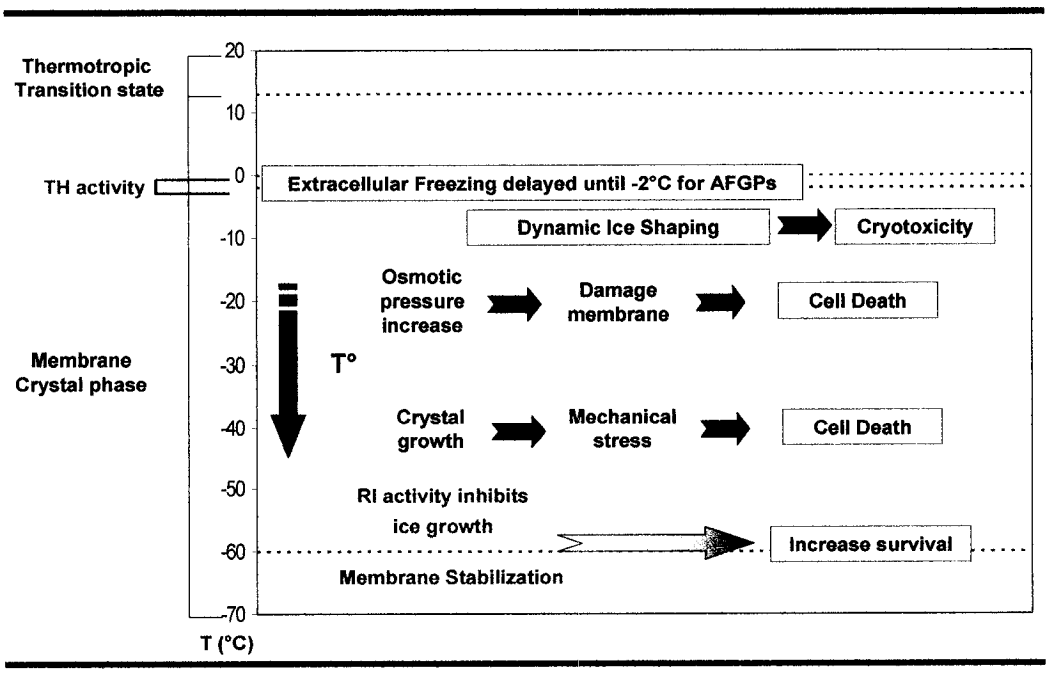
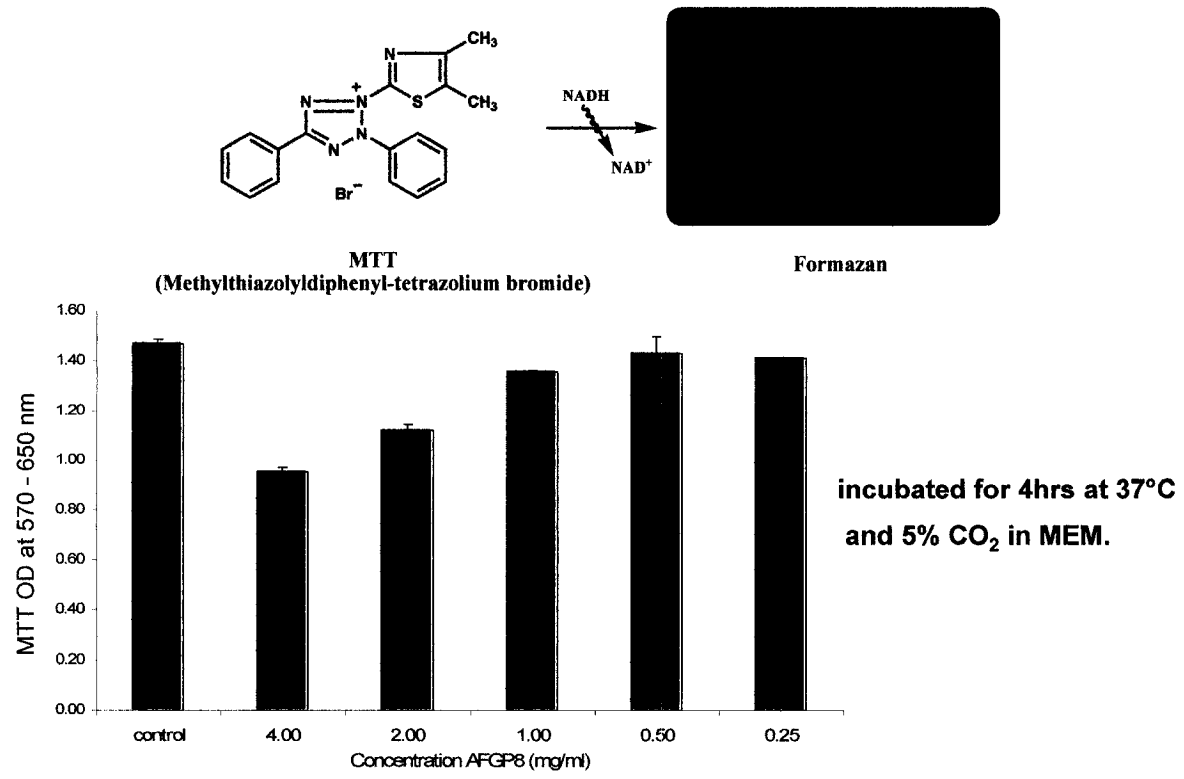


Figure 13: In vivo temperature transition with cryoprotectants

Collectively these properties make AF(G)Ps useful in order to improve the cryopreservation of cells; they also exhibit two major disadvantages. First, at temperatures below the TH gap dynamic ice shaping by AFPGs, induces cryotoxicity. Secondly, our lab has shown that AFGP8 is cytotoxic towards mammalian cells (Graph 4).[123]



Graph 4: MTT Assay of AFGP8 on human fetal liver cells[123]

Graph 4 represents the cytotoxicity of AFGP8 toward human embryonic liver cells. The MTT is reduced by mitochondrial metabolism to formazan an insoluble purple salt by healthy cells. After filtration and solubilization, the optical density of the solution is directly proportional to the number of healthy cells. The 4 mg/mL solution clearly exhibits cellular cytotoxicity compared to the control.

Consequently, the ultimate cryoprotectant, which would contain all the attractive properties without possessing any cytotoxicity and/or cryotoxicity seems to be only accessible by the synthesis of analogues of AFGPs.

1.5 Important features for antifreeze activity.

1.5.1 Influence of the length of the glycopolymer.

As shown in Table 1 the molecular weight of the AFGPs can vary from 2.6 kDa for AFGP8 to 33 kDa for AFGP 1. The smaller molecular weight AFGP6-8 form most of the circulating AFGPs in the body fluids, however, they are only responsible for two thirds of the activity. In a thorough study, using highly purified AFGPs, it has been demonstrated that AFGPs with molecular weights higher than 13 kDa had approximately three times more thermal hysteresis activity than the smaller species.

1.5.2 Structural modifications of the carbohydrate moiety and its influence.

Due to the difficulty in obtaining such compounds by purification and/or by synthetic routes,[124-129] the number of structure-activity relationship studies is limited. Figure 14 summarizes the work completed on these analogues.[15] All the following work was based on TH activity and not RI activity.

In the first experiment a β-elimination was performed to remove the disaccharide moieties and the hydroxyl group of the threonines, resulting in the total loss of thermal hysteresis activity. The peracetylation of all the hydroxyl groups to give **1** or the periodate oxidation of the carbohydrate moiety to give **2** resulted in a total loss of TH activity. This is consistent with the idea that some, if not all, hydroxyl groups are necessary for the TH activity. The oxidation of the C-6 hydroxyl groups to give the bisaldehyde **3** has no influence on the TH activity; this was interpreted as the hydroxyl groups on the C-6 position are not essential for the activity. However, the substitution by a carboxylates or sulfites results in the total loss of activity, showing that the presence of a negative charge at this position is not tolerated. Addition of sodium borate formed a mixture of **6a** and **6b** and eliminates any activity. Further experiments have been realized in order to determine if the addition of bulky groups on the C-6 position had any influence. The results indicated that bulky groups usually decrease the activity but do not inhibit it.

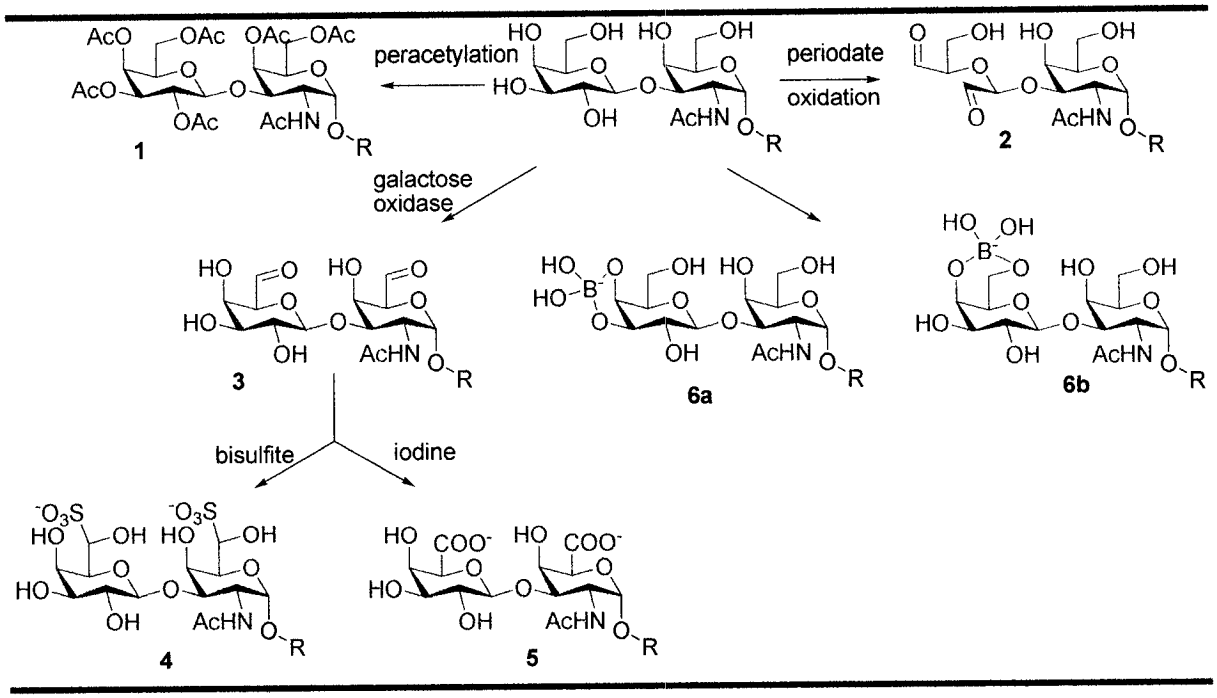
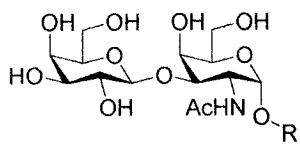
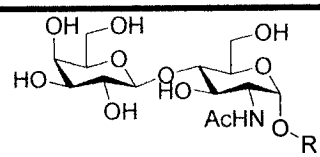
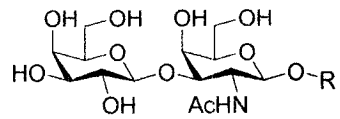
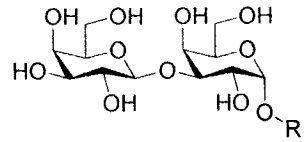
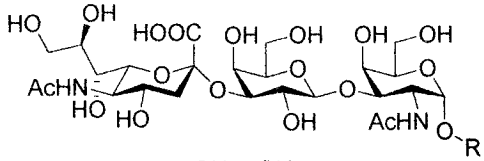
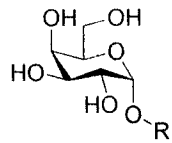
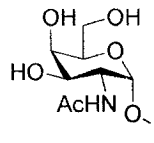
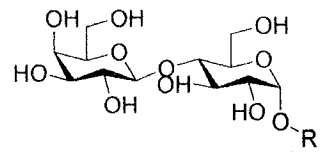


Figure 14: Summary of the different structural modifications completed on AFGPs[15]

In 2004, Nishimura published a paper where he synthesized a series of AFGP analogues,[128] and answered several important questions about the relative importance of different structural features. The different analogues that he synthesized are summarized in Table 3.

In this study Nishimura demonstrates that even though the monomer (one disaccharide unit linked to a tripeptide core Ala-Ala-Thr) does not show any specific antifreeze activity, it has the ability to exhibit some dynamic ice shaping. The average molecular weights of the compounds tested were approximately 6.6 kDa. In order to investigate the relative importance of the glycosidic linkage between the disaccharide moiety and the peptide backbone, compound **8** was tested and showed very weak dynamic ice shaping and no TH activity. This result suggests that the α-glycosidic linkage is an essential feature for the activity. Amongst the other analogues, only GalNAc **9** and LacNAc **10** exhibited a weak TH activity. To summarize, this study demonstrated that the α-glycosidic linkage, the *N*-acetyl function on the C-2 position, the γ-methyl on the threonine, the relative orientation of the different hydroxyls as well as the disaccharide unit all have a significant impact on the TH activity.

Table 3: Nishimura's AFGPs analogues[128]

	7		11
	8		12
	9		13
	10		14

1.6 Previous syntheses of AFGP analogues

Due to the fact that isolation of AFGPs is a long and costly process, that it would be difficult to isolate in industrial quantities, and that the biosynthesis has not yet been achieved, recently researchers have focused on the development of efficient synthetic routes to AFGPs. Moreover, due to the need for different analogues, these syntheses might prove to be extremely useful.

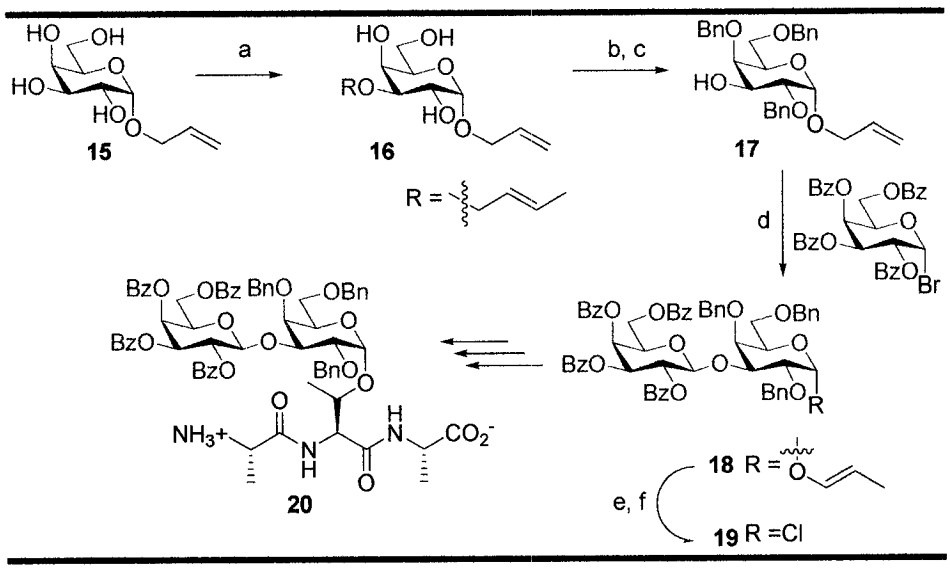
1.6.1 First synthesis of a core unit analogue

One of the first syntheses of building blocks to obtain AFGPs analogues was realized by Anderson[124] and consists on the synthesis of a disaccharide moiety connected to a tripeptide building block. In this synthesis, Anderson incorporates a β-D-galactosyl-1-3-α-D-galactose disaccharide to mimic the β-D-*N*-acetylgalactosamine disaccharide of the native system (Scheme 1).

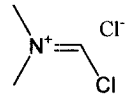
To synthesize the disaccharide moiety, the allylated monosaccharide **15** was selectively crotylated on the C-3 position in the presence of tributyl tin oxide and crotyl

bromide. The remaining hydroxyl groups were benzylated and the crotyl was selectively removed to obtain compound **17**.

Scheme 1: Anderson's synthesis[124]



Reagents : a) Bu_3SnO, crotyl bromide ; b) NaH, BnBr ; c) *Tert*-BuOK ; d) AgOTf, tetramethyl urea, 2,6-dimethylpyridine ; e) $HgCl_2$, HgO ; f) CH_2Cl_2,



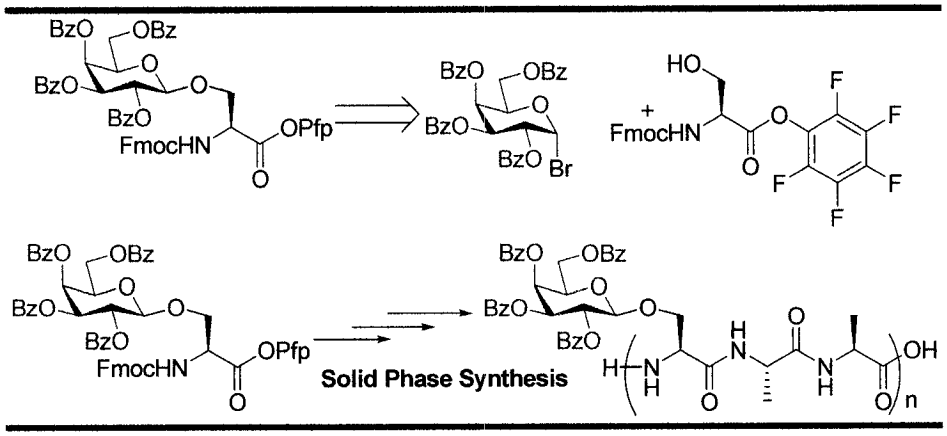
The glycosylation was performed using silver triflate and halide galactose derivative to afford the disaccharide **18**. The tripeptide unit (Ala-Thr-Ala) was synthesized separately and then coupled to the disaccharide unit. In order to obtain the desired α-glycosidic linkage between the peptide and the disaccharide, Anderson protected the *C*-2 position with a non participating group; benzyl ether which should minimize the amount of β diastereoisomer that is formed. The glycosylated tripeptide building block prepared in a convergent fashion can then be employed in conventional solid phase synthesis.

1.6.2 Pentafluorophenyl esters for temporary protection of acid in SPS of an AFGP analogue

Meldal's group synthesized a monosaccharide β-linked AFGP analogue using solid phase synthesis.[125] This synthesis provided an interesting advantage by the use of a glycosylated Fmoc-Serine which already possesses an activated ester on the C-terminus

(Scheme 2). This strategy reduced the number of steps during both the liquid phase and the solid phase sequence.

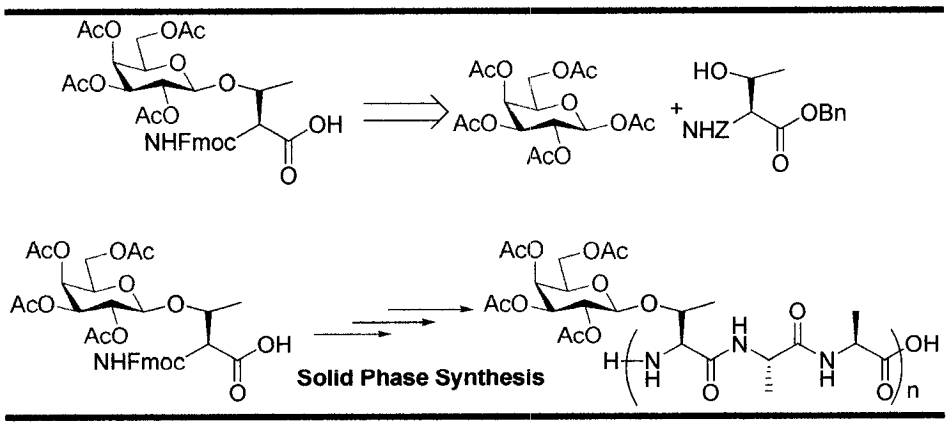
Scheme 2: Meldal's SPS retrosynthesis[125]



1.6.3 Linear solid phase synthesis of monosaccharide analogues

A similar synthesis was reported by Filira and co-workers in 1990. They reported the synthesis of AFGPs analogues containing two to seven repeating tripeptide units.[126] They used continuous flow solid phase synthesis in order to obtain different AFGPs analogues for their conformational study (Scheme 3).

Scheme 3: Filira's SPS retrosynthesis[126]



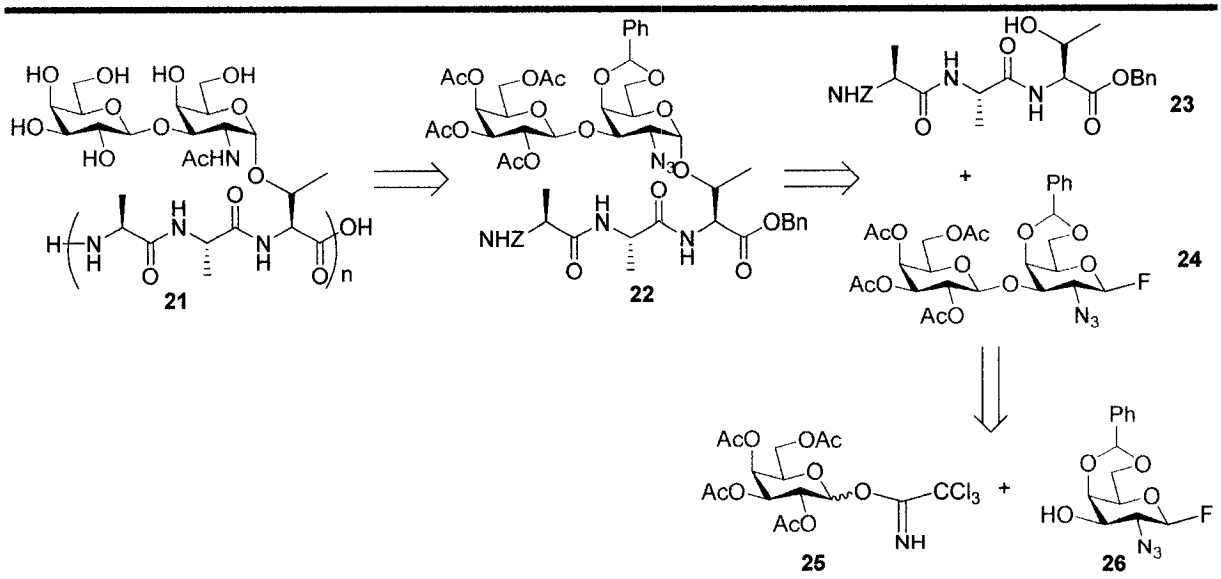
In this synthesis Filira used a monosaccharide, β-D-galactose instead of the native 3-*O*-(*β*-D-galactosyl)-D-*N*-acetyl-galactosamine to mimic the native β-(1-3) linkage between

the galactose and the galactosamine. β-D-Galactose pentaacetate and Z-Thr-OBn were coupled in the presence of boron trifluoride ethyl etherate to produce the glycosylated building block. The building block is then debenzylated and the *N*-terminus is selectively protected with Fmoc succinimide. This yields the desired building block ready for solid phase synthesis. To incorporate alanine residues, Fmoc-Ala-OPfp (Pfp = pentafluorophenol) was used for the couplings. The pentafluorophenyl ester is a preactivated ester, which allows a direct coupling on the solid phase. The polymers were removed from the Wang Resin using 95% TFA and then deacetylated to give the desired AFGPs analogues.

1.6.4 First synthesis of AFGPs

The first synthesis of a native AFGPs was described by Nishimura and co-workers in 1996, using diphenyl phosphoride azide (DPPA) mediated solution polymerization.[127] In this approach they polymerized a fully deprotected glycosylated tripeptide unit to obtain analogues containing eight to twelve repeating units (Scheme 4).

Scheme 4: Nishimura's solution phase synthesis[127]



To selectively obtain the disaccharide β anomer **24**, they reacted the galactosyl trichloroacetimidate **25** with the acceptor **26** in presence of trimethylsilyl trifluoromethylsulfonate (TMSOTf) at low temperature. In order to obtain the α linkage

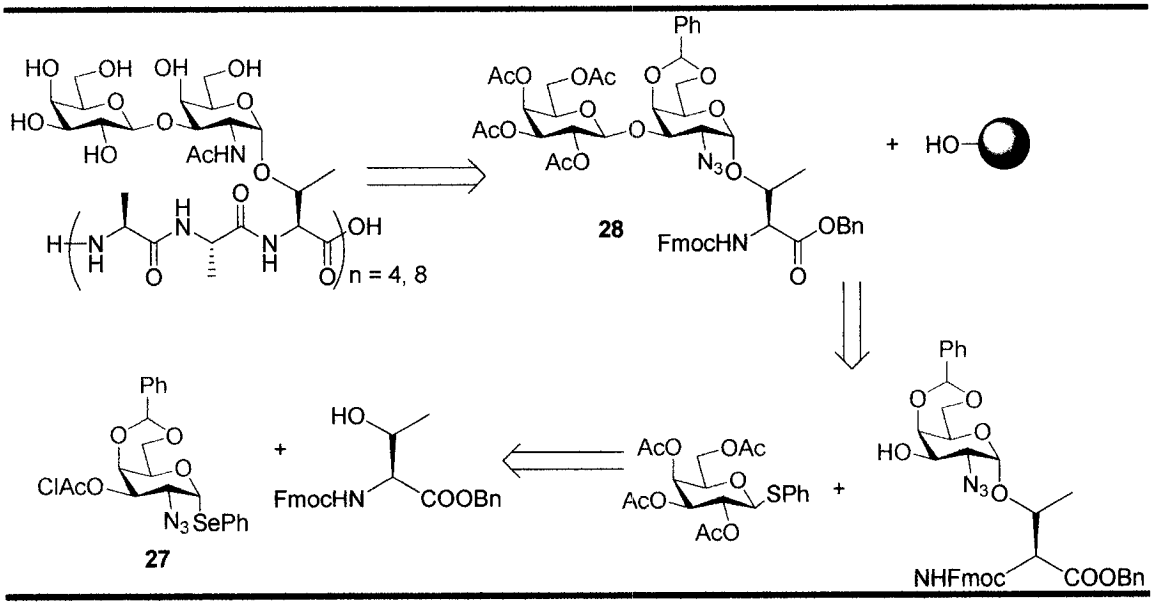
between the tripeptide unit and the carbohydrate moiety, they used a non participating azide group in the *C*-2 position. The azide group possesses two advantages, the first of which is that, it does not form an oxonium bridge during the coupling, thus minimizing the formation of the β anomer. Second, it is also a good precursor that can subsequently allow the installation of an amine group. To achieve this coupling, they used $(C_5H_5)_2ZrCl_2$ and silver perchlorate. The azide group was then reduced and acetylated to form the acetamide. Prior to the polymerization, the tripeptide building block was then fully deprotected for solubility reasons. The polymerization is mediated with 1.3 equivalents of diphenyl phosphoryl azide (DPPA) and triethylamine in dimethyl sulfoxide, resulting in polymers of molecular weight around 6 kDa. This methodology allowed the researchers to obtain different complex glycopolymers, however due to the complex mixture of various molecular weight products, it is difficult to use these solutions as a probe for SAR studies.

In a later publication, Nishimura demonstrated that this polymerization strategy is also compatible with base labile groups such as *O*-acetate, *O*-sulfate and/or *O*-phosphate giving access to more complex glycopolymers.[128] Moreover other polycondensation agents, such as 1-isobuthoxycarbonyl-2-isobuthoxy-1,2-dihydroquinoline (IIDQ) or 4-(4,6-dimethoxy-1,3,5-triazin-2-yl)-4-methylmorpholinium chloride (DMT-MM) have been tested and were proved to be equally efficient.

1.6.5 First synthesis of AFGP8

The second total synthesis of native AFGPs has been reported by Chen *et al.* Using solid phase synthesis, they were able to accurately control the length of the biopolymer synthesized (Scheme 5).[129] This approach features the coupling of the seleno derivative **27** with the protected threonine residue in the presence of silver triflate and tetramethyl urea. The resulting monosaccharide component is then deprotected and coupled the galactose donor to give disaccharide **28**. After the hydrogenation and acetylation, the desired building block is ready for solid phase synthesis to afford the desired AFGP molecules.

Scheme 5: Chen's synthesis[129]



While these approaches provide several strategies to afford AFGP analogues, few of these compounds have been used for structure-activity relationship (SAR) studies and none have been used in order to investigate RI activity. Some of the mentioned syntheses, especially the approaches that incorporate solid phase chemistry, are also adaptable to the preparation of a combinatorial library of AFGP analogues. During the past decade, antifreeze glycoproteins have demonstrated tremendous potential for many industrial, commercial and medical applications. However, realization of these applications is still a long way off, as the commercialization of such compounds is limited by their availability.

1.7 Previous work on the synthesis of *C*-linked AFGPs by our laboratory

Structurally diverse AFGP analogues could be used as probes to further elucidate the molecular mechanisms of action. Alternatively, rationally-designed chemically and biologically stable AFGP analogues[130,131,132] would be ideal molecules to facilitate membrane stabilization because they could be custom-tailored for each type of cell membrane and produced on a large scale. Moreover, such analogues may find applications as cryosurgical adjuvants where recent advances in ultrasound and magnetic resonance imaging allow for the complete destruction of tumours. During the past 10 years, the role of

AFP in such applications has been extensively investigated, but similar studies with AFGP or AFGP analogues have not been performed. Toward these goals several *C*-linked AFGP analogues have been previously prepared in our laboratory. *C*-linked carbohydrates are known for their enhanced stability toward acid and basic conditions and are also known for their ability to mimic their *O*-linked counterpart.

1.7.1 Mimicking O-linked native AFGPs with C-linked glycoconjugates

1.7.1.1 Exo-anomeric and exo-desoxyanomeric effect

Since the early 1970's many syntheses of *C*-linked carbohydrate derivatives have been reported.[133-138]

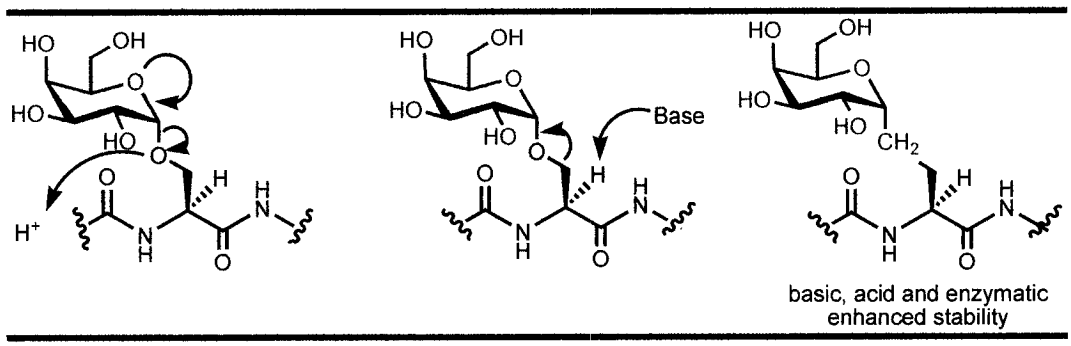


Figure 15: Enhanced stability of C-linked glycopeptides

Due to the instability of the native *O*-linked glycoproteins (Figure 15), countless studies have been performed in order to obtain stable glycoprotein mimics. Stable glycoprotein analogues with comparable biological activity are essential tools to investigate the different biological functions described previously. Their enhanced stability towards basic and acidic media as well as resistance to enzymatic degradation makes them ideally suited for this purpose.[139,140,141]

Several families of glycoprotein analogues with enhanced stability have been synthesized and tested, among them *S*-linked,[142] *N*-linked,[143] and *C*-linked analogues (Figure 16). While *C*-linked analogues represent a dramatic change in the polarity and the electronic effect of the glycoconjugates, they are by far the more stable analogues.

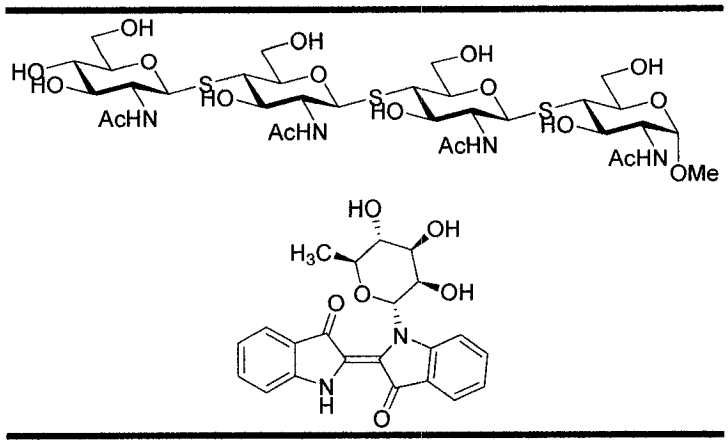


Figure 16: S-linked[142] and N-linked[143] analogues of the native molecule

Probably one of the greatest contributions by Lemieux[144] to the field of carbohydrate chemistry has been his delineation of the importance of the *exo*-anomeric effect.[145,146,147] In a simple acetal derived from a carbohydrate, the normal anomeric effect is observed, which stabilizes the axial anomer over the equatorial anomer. However, when the exocyclic alkoxy group is in the appropriate configuration, there is an antiperiplanar arrangement of a lone pair on oxygen (of OR) and the C_1-O_5 bond. This is described as the *exo*-anomeric effect (Figure 17).

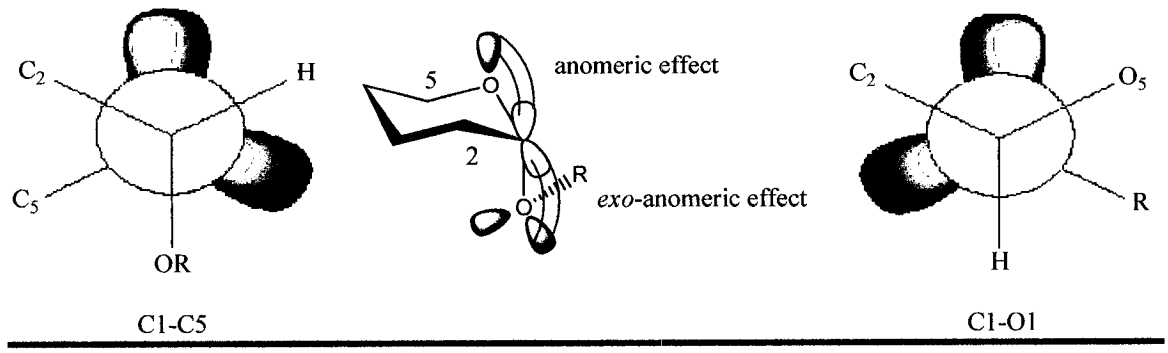


Figure 17: Anomeric and exo-anomeric effect

Because these two anomeric effects operate in opposite directions, the *exo*-anomeric effect is not considered important in such axial acetals. Where there is no contribution from a normal anomeric effect, it is the *exo*-anomeric effect that is dominant and dictates the preferred conformation of the alkoxy group at the anomeric carbon atom. The logical conclusion is that the *exo*-anomeric effect appears to be responsible for the helical shape of

many polysaccharide chains. It is certainly important in determining the shape of many biologically-important oligosaccharides.

Methylene is of similar size to an oxygen atom with its two lone pairs, and it is therefore an isostere. Kishi demonstrated that *C*-glycosides have the same conformation as *O*-glycosides.[148] The aglycon have a conformation where the interaction n-σ*, between the lone pair of the oxygen 1 and the antibonding orbital of the bond C_1-O_5 is favoured. It is the *exo*-anomeric effect. However, within the *C*-glycoside there is no *exo*-anomeric effect due to the absence of the *exo* cyclic oxygen. Using ab initio calculations on models such as 2-ethyltetrahydropyrane and 2-ethyl-3-hydroxytetrahydropyrane (model for glucose and galactose), Houk *et al.*[149] defined the *exo*-desoxyanomeric effect. The conformation of a *C*-glycoside is influenced by the minimization of the steric strain. For instance, in the case of 2-ethyltetrahydropyrane, the conformational energies of the *exo-syn* and *non-exo* forms are comparable, whereas in the case of 2-ethyl-3-hydroxytetrahydropyrane, the *non-exo* conformation is disfavored because of the 1,3-diaxial (syn-pentane) interaction between the hydroxyl group on the C-2 position and the R group. The *exo*-anomeric effect as well as the *exo*-desoxyanomeric effect therefore favour the same conformations.

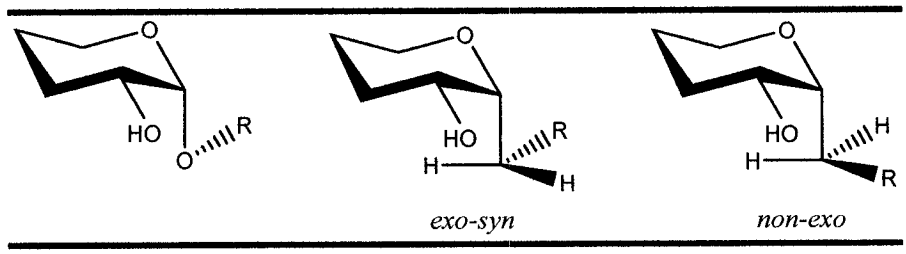


Figure 18: The exo-desoxyanomeric effect of 2-ethyl-3-hydroxytetrahydropyrane[149]

This tendency was subsequently confirmed by another study from Kishi.[150] Using NMR techniques, they demonstrated that *C*-linked pyranoglycosides tend to adopt preferentially a similar conformation to their *O*-linked analogues. These results were confirmed for both α and β anomers as well as for different oligosaccharides.

Several other studies verified this trend, for instance *C*-glycoside and *O*-glycoside solution conformations were demonstrated to be similar using receptor-ligand recognition techniques. Using X-ray crystal structure, Kishi *et al.* reported that the *C*-lactose bond to

peanut lectin is in a near perfect *exo*-anomeric conformation which also corresponds to the binding conformation of the *O*-linked analogue.[151]

1.7.1.2 Conformational and biological studies of C-glycoside analogues

Many different syntheses have been published describing methods to obtain *C*-glycosides and *C*-oligopeptides. As the conformation of analogues plays a major role in enzymatic or biological recognition process, several studies were also performed to compare *C*-linked analogues to their *O*-linked counterparts.

For instance, J. Jimenez-Barbero *et al.* studied α-*C*-Man-(1→1)-β-Gal (Figure 19) using *ab-initio* calculations,[152] and they compared the different conformations of *O* and *C*-disaccharides.

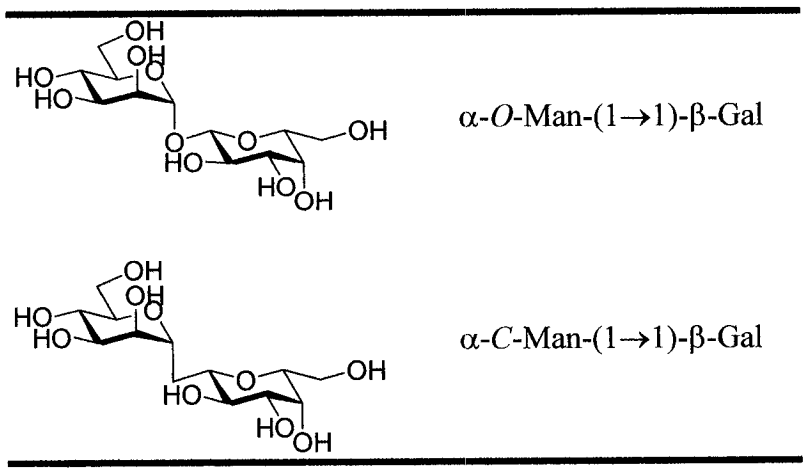


Figure 19: Carbohydrates used in Jimenez-Barbero's conformational studies[152]

Table 4 and Table 5 shows the predicted conformation for both *O*-linked and *C*-linked analogues. The different conformations were accurately predicted by calculations and were supported by NOE NMR experiments. This method of combining molecular modeling as well as NMR studies was applied to other systems.[153,154,155] These studies conclude that *C*-glycosides are usually more flexible than their *O*-glycoside analogues.

Table 4: Different conformations of the disaccharide α-O-Man-(1→1)-β-Gal[153]

		Mannose		
		Exo-syn	*Non-exo*	*Exo-anti*
Galactose	*Exo-syn*	>93%	4%	
	Non-exo			
	Exo-anti	2%	<1%	

Table 5: Different conformations of the disaccharide α-C-Man-(1→1)-β-Gal[153]

		Mannose		
		Exo-syn	*Non-exo*	*Exo-anti*
Galactose	*Exo-syn*	30%	42%	
	Non-exo	12%		
	Exo-anti	6%	10%	

The interactions of *C*-glycosides were intensively studied with glycosidases. For instance, the activity of a β-glycosidase from *E.coli* has been studied using *C*-lactose (Figure 20).[156] First, theoretical calculations tend to show that the natural substrate and the carbonated analogues have a similar conformation. Moreover in this case the *C*-lactose is well recognized by the enzyme, even though the gauche-gauche conformation adopted by this molecule in the active site has never been observed in the native substrate (β-*O*-gal(1→3)-β-glc). This important information shows that the enzyme recognizes a high energy conformation of the *C*-glycoside. It is theorized that the lactose might adopt a similar conformation when it interacts with the active site.[156]

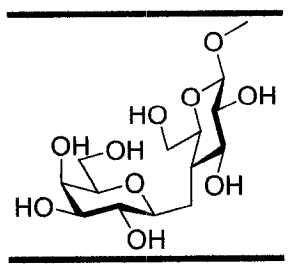


Figure 20: Methyl-C-lactose in gauche-gauche conformation[156]

Sinay *et al.* prepared several analogues of the heparin pentasaccharide (Figure 21).[157] Amongst these analogues, one with a methylene group (X= CH_2) between the D and E units was prepared.

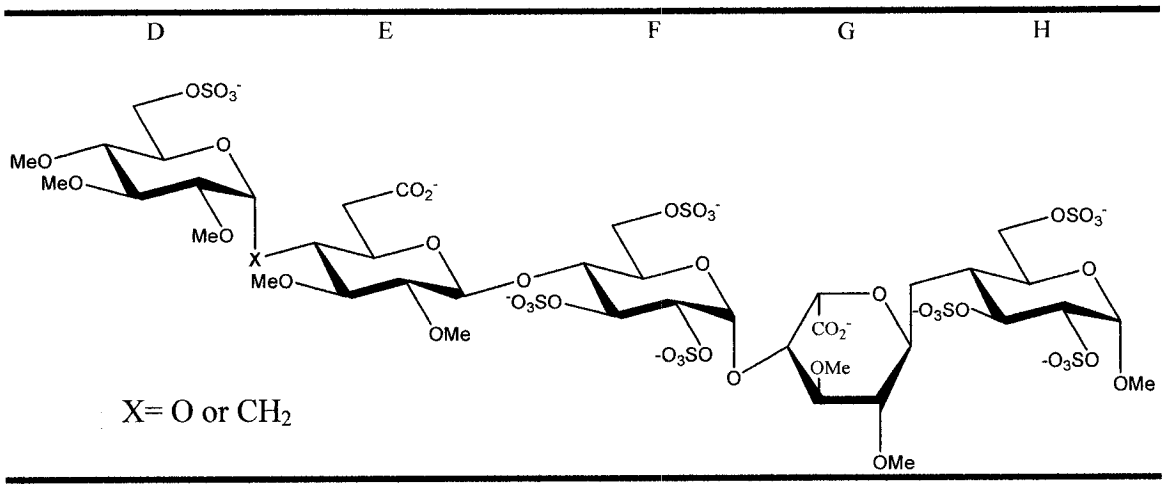


Figure 21: Heparin and its C-linked analogue synthesized by Sinay's group[157]

Heparin is a sulfated polysaccharide that activates a glycoprotein (antithrombine III) that inhibits the formation of fibrin.[157] Fibrin can block blood circulation in vital organs such as the heart and the lungs in some patients. The affinity and the activity of antithrombin III were measured and compared to the *O*-linked analogues (Table 6).

Table 6: Comparative activity of O-linked and C-linked analogues of Heparine[157]

Compounds	Affinity for AT III (*Kd* in mM)	Activity (anti-factor Xa) (IU/mg)
O-Glycoside	1.9 ± 0.1	1180 ± 30
C-Glycoside	2.8 ± 0.2	880 ± 40

This data shows that in this case the replacement of the interglycosidic oxygen with CH_2 has very little influence on the activity of the molecule. The same group also prepared different analogues of the β(1→6)-D-galactane oligomers,[158] which exhibit an interesting affinity for various immunoglobulins (Figure 22).[159]

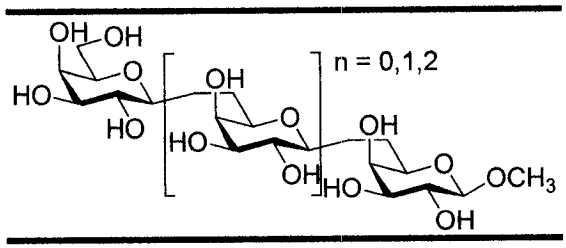


Figure 22: Example of the C-linked analogues of the β(1→6)-D-galactane oligomers[159]

Although in some cases the exchange between the intersaccharidic oxygen and carbon does not seem to affect the biological activity, several other cases have been shown to either increase or decrease this activity.

The agelsphins are glycolipids isolated from various marine sponges. These compounds were tested on mice and demonstrated anti-tumor activity. A researcher from Kirin Pharmaceuticals modified these molecules and synthesized the compound KRN 7000, which shows improved activity (Figure 23).

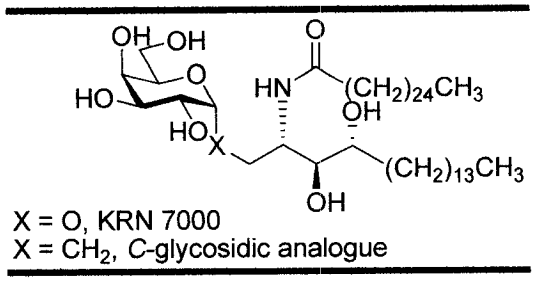


Figure 23: Structure of the compound KRN 7000 and of its C-linked analogue[160]

Franck *et al.* later synthesized the *C*-linked analogue of KRN 7000[160] and reported that their analogue was approximately 1000 fold more active than the *O*-linked analogues and remained active for longer time periods. This demonstrates the improved stability of the *C*-linked systems toward biological systems. In this example the improved activity was attributed to some new hydrophobic interactions due to the new methylene group.

L.L. Kiessling and E.J. Toone also used *C*-glycosides in order to understand biological systems.[161] Using by titration microcalorimetry and fluorescence anisotropy titration, they studied the binding of the mannose/glucose specific lectins from *Canavalia ensiformis* and *Dioclea grandiflora* to a series of *C*-glucosides and mannosides. These closely related lectins share a specificity for the trimannoside methyl 3,6-di-*O*-(α-D-

mannopyranosyl)-α-D-mannopyranoside, and are a useful model system for addressing the feasibility of differentiating between lectins with overlapping carbohydrate specificities. The ligands (Figure 24) were designed to address two issues: (1) how the recognition properties of non-hydrolyzable *C*-glycoside analogues compare with those of the corresponding *O*-glycosides and (2) the effect of having more than one saccharide recognition epitope on both affinity and specificity. Both lectins bind the *C*-glycosides with affinities comparable to those of the *O*-glycoside analogues; however, the ability of both lectins to differentiate between the *gluco* and *manno* diastereomers was diminished in the *C*-glycoside series.

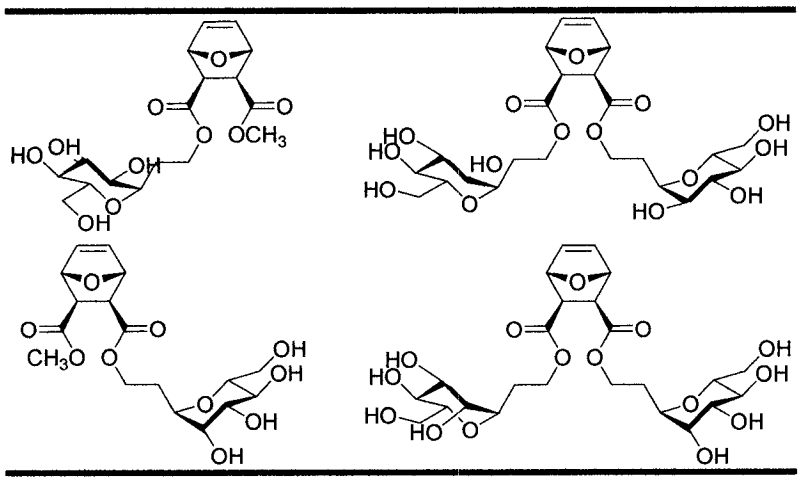


Figure 24: Different lectin ligands synthesized by the Kiessling and Toone groups[161]

These different examples show that the conversion of the interglycosidic oxygen to a methylene group can have multiple consequences. Consequently *C*-glycosides can be used:

- As molecular probes, to obtain a better understanding of the receptor
- As a ligand or inhibitor of different enzymes
- As therapeutic agents displaying better activity than the *O*-linked analogues
- As stable biomolecular tools (i.e. to improve survival of cells such as heat shock molecules, antifreezes)

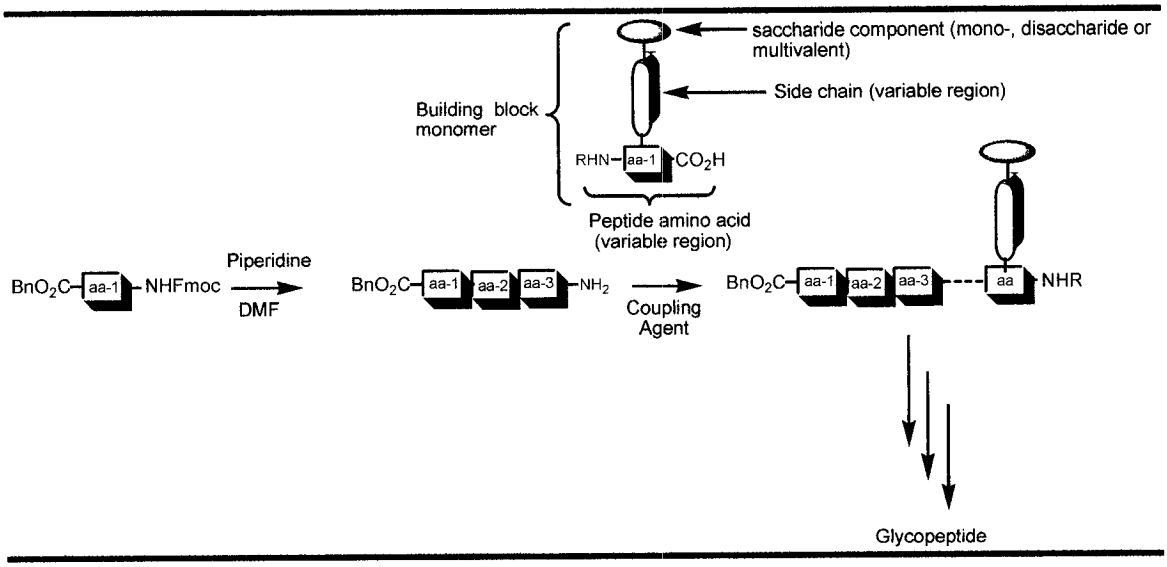
Therefore *C*-linked analogues possessed all the required properties for our purpose. Despite the many advances in the field of oligosaccharide and glycoconjugate chemistry, complex glycans require lengthy and costly syntheses.[162] Two reasons for this are the high instability of the anomeric carbon-oxygen and the need to employ orthogonal protecting group strategies.

1.7.2 Standard glycopeptides synthesis strategies

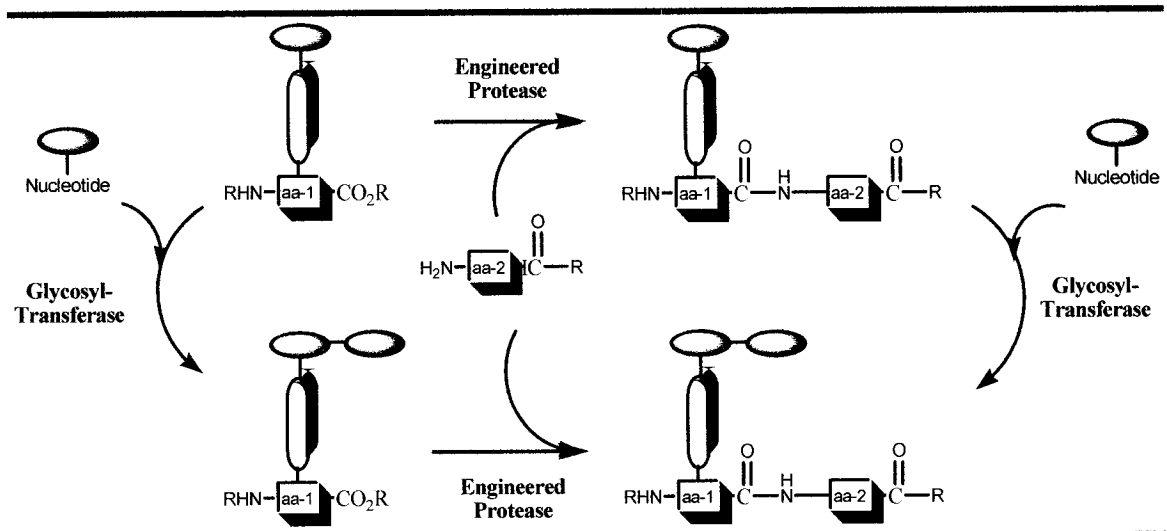
Although the study of glycopeptides has received a large amount of attention, access to these molecules remains challenging. While automation of the synthesis of sequenced nucleic acids and peptides is now a simple routine, this is not the case for the synthesis of glycoconjugates. This is mainly due to the nature of the carbohydrate moiety. Indeed, in order to carry out syntheses involving carbohydrates, careful procedures must be followed. First, carbohydrates possess several hydroxyl groups with similar reactivity; therefore differentiation of these hydroxyl groups using selective protective groups must be completed in order to realize regioselective reactions. Secondly, carbohydrate moieties are sensitive to basic and acidic media, which also decrease the range of reactions possible on these molecules. Finally, during the glycosylation different diastereoisomers (α or β linkages) often result, requiring the utilization of stereoselective glycosylation procedures that are specific to carbohydrates.

To meet the demand for homogeneous glycoproteins or glycopeptides, many synthetic methods were developed in the past decades. These methods include solution (Scheme 6) and solid phase syntheses using properly protected glycosyl amino acids as key building blocks.[163-167] Other methods use other elegant synthetic designs such as chemoenzymatic synthesis utilizing enzymatic transglycosylations or enzymatic elongation (Scheme 7) of the peptide and carbohydrate chains[168-171] and chemoselective ligations (Scheme 8) between carbohydrates and peptides.[172,173,174] Another very interesting method is the solid phase synthesis using free glycosyl amino acids as building blocks. An advantage of this last technique is that it does not need final stage carbohydrate deprotection, however it is limited to short glycopeptide structures. Finally, a new strategy using unprotected glycosyl amino acids as building blocks[175,176] and phase tags was developed.[177,178] This technique avoids the use of the traditional 95% of TFA to release the glycopeptide from the polymer support.

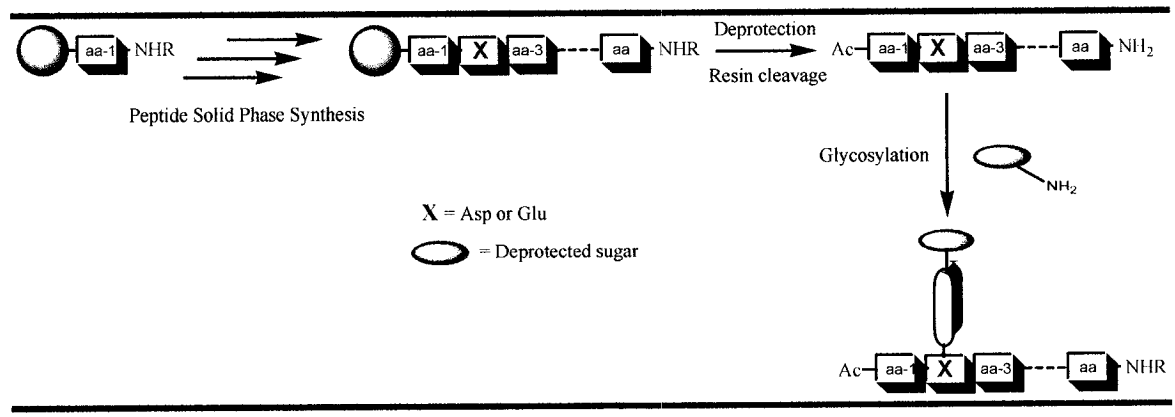
Scheme 6: Solution synthesis with protected glycosyl amino acids



Scheme 7: Enzymatic elongation of the peptide and carbohydrate chains

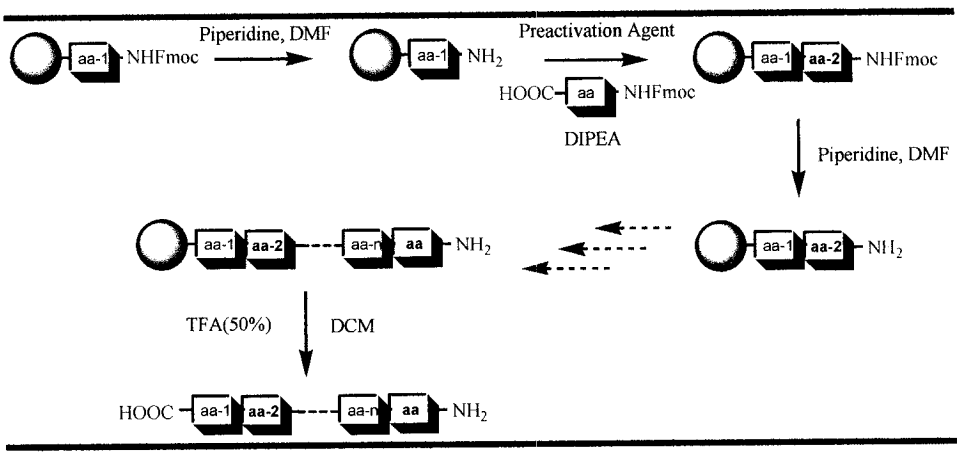


Scheme 8: Chemoselective ligation in glycopeptides synthesis



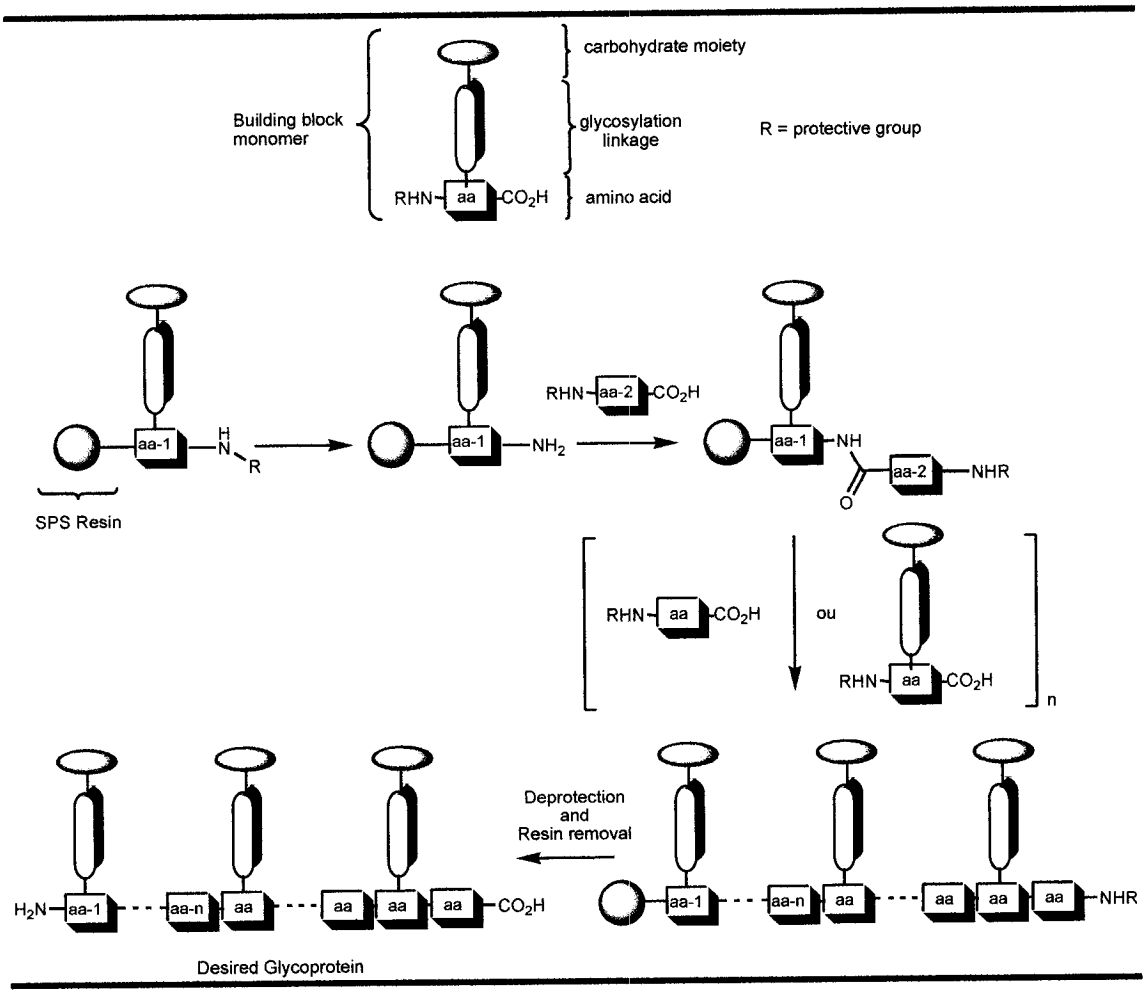
Despite all these techniques the stepwise solid phase synthesis approach that employs fully protected glycosylated amino acid building blocks is nowadays the most general and efficient method for preparing structurally diverse glycopeptides. The most compatible strategy to obtain glycopeptides is the regular peptide Fmoc strategy (Scheme 9). By protecting the α-amino group of building block monomers with fluoren-9-ylmethoxycarbonyl (Fmoc) and utilizing base-labile protecting groups (i.e. acetate) for the hydroxyl groups of the carbohydrate component, it is possible to obtain the desired glycopeptides.

Scheme 9: Standard solid phase Fmoc strategy



The β-elimination as well as epimerization side products resulting from basic conditions (i.e. sodium methoxide solution for the deprotection of acetate) were demonstrated to be minimal.[179] Some other protecting groups were found to be compatible with such syntheses. For instance, the fluorobenzoyl group and some acid-labile groups such as tert-butyldiphenylsilyl, tert-butyldimethylsilyl or 4-methoxybenzyl ether can be utilized (Scheme 10).

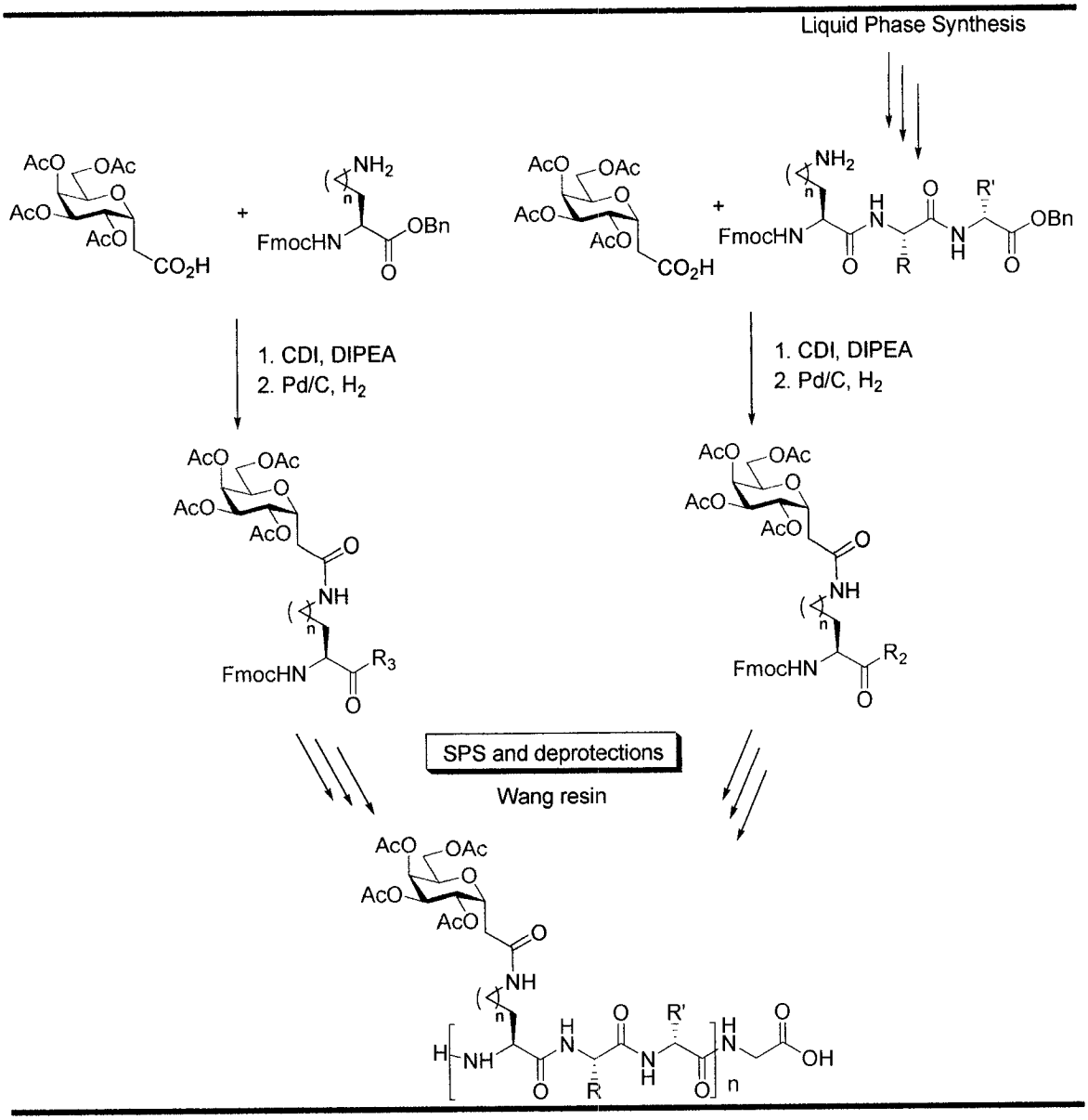
Scheme 10: General procedure for solid phase synthesis of structurally diverse glycoproteins



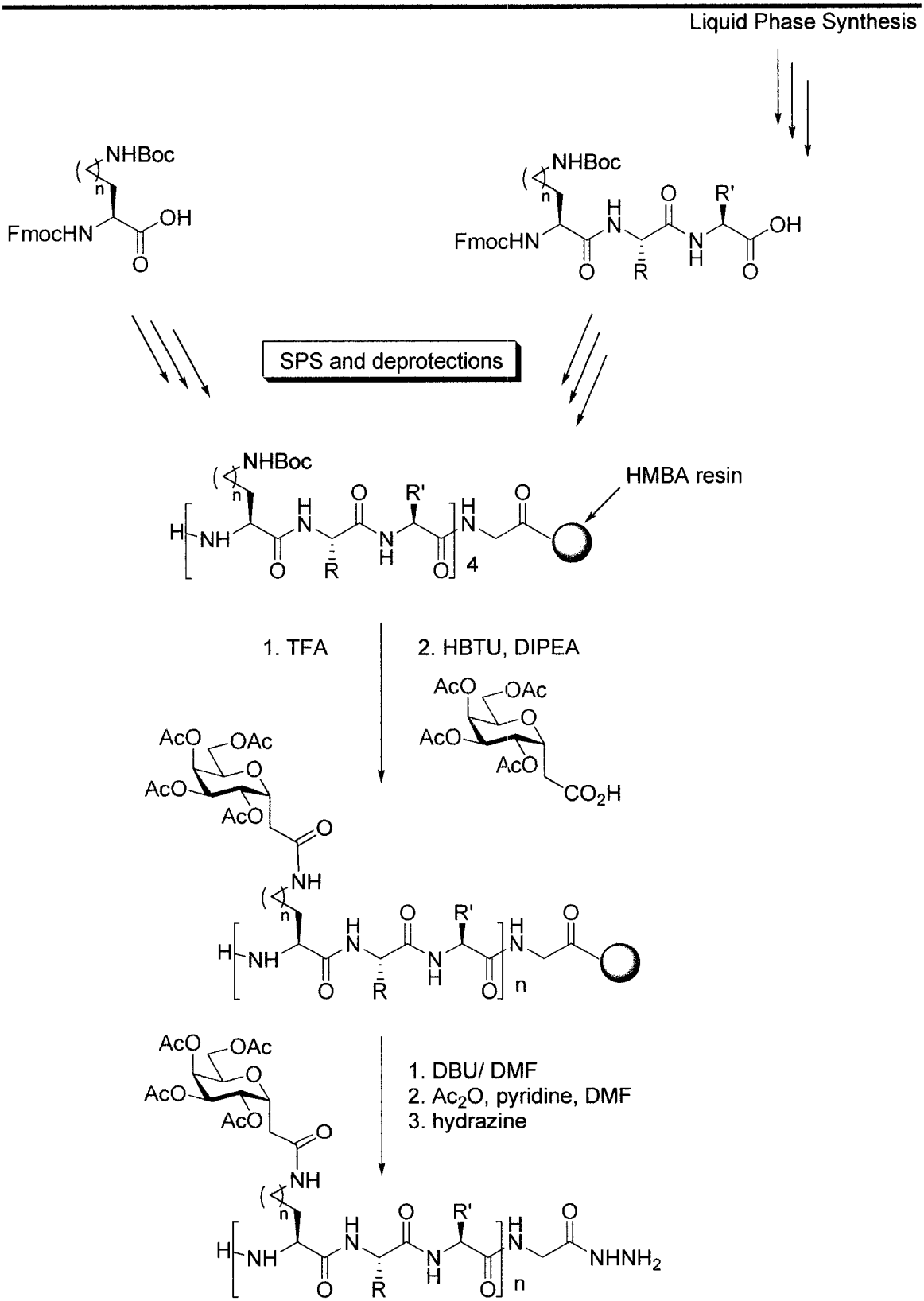
1.7.3 Our synthetic strategy for the synthesis of C-linked AFGPs

Despite many advances in *C*-linked glycoconjugate synthesis, there are still many naturally occurring systems for which *C*-linked analogues are not easily constructed. As part of our continuing efforts towards the rational design of functional AFGP analogues, our group has previously investigated and designed different general synthetic strategies[132,180] to afford structural *C*-linked analogues of AFGP. These strategies consist of a regular amide ligation between an amino acid and a carbohydrate unit followed by either a linear solid phase strategy (Scheme 11) or a convergent solid strategy (Scheme 12).

Scheme 11: Linear solid phase strategy[180]



Scheme 12: Convergent solid phase strategy[180]



The linear strategy is based on the synthesis of an amino acid or oligopeptide unit, which is then coupled to the desired carbohydrate moiety. Once the desired building block is obtained, the peptide is assembled in a linear fashion using solid phase synthesis. After cleavage from the beads, full deprotection, and purification of the glycopolymer, the AFGP analogue can be tested for recrystallization-inhibition and thermal hysteresis activity.

The convergent strategy consists of the synthesis of the full peptide backbone, cleavage from the beads, and deprotection of the lateral functionality where the desired carbohydrate moiety will be subsequently coupled. Once the peptide backbone is purified, the carbohydrate moieties are coupled to the peptide backbone using regular solution phase techniques.

The linear strategy proved to be more efficient than the convergent method. Indeed, the overall yields of the linear strategy are higher and performing the glycosylation in liquid phase is more reliable.

1.7.4 Previous C-linked AFGP8 analogues synthesized in our laboratory

1.7.4.1 Influence of the molecular weight and the carbohydrate concentration on RI[181]

a) Molecular weight:

Previous work on antifreeze glycoproteins demonstrated that the weight of the glycoproteins is an important factor for thermal hysteresis activity. AFGP1 (MW = 22 kDa) exhibits a greater thermal hysteresis gap than AFGP8 (MW = 2.6 kDa).[40] Furthermore, previous work performed in our laboratory suggests that the influence of molecular weight might be different for thermal hysteresis and recrystallisation inhibition (Figure 25). As described earlier, thermal hysteresis is associated with dynamic ice shaping, which is detrimental to cell survival at temperatures lower than the thermal hysteresis gap. Consequently, variation of the length of the peptide backbone might be one key to tailor thermal hysteresis and recrystallisation inhibition to the desired levels.

Previous Ph.D. students from the Ben laboratory have prepared analogues with different modifications toward the native system.[181] These analogues possess a

monosaccharide moiety instead of the native β-D-galactosyl-(1,3)-D-*N*-acetylgalactosamine. This modification has been implemented because of potential future industrial applications as well as to facilitate the synthesis.

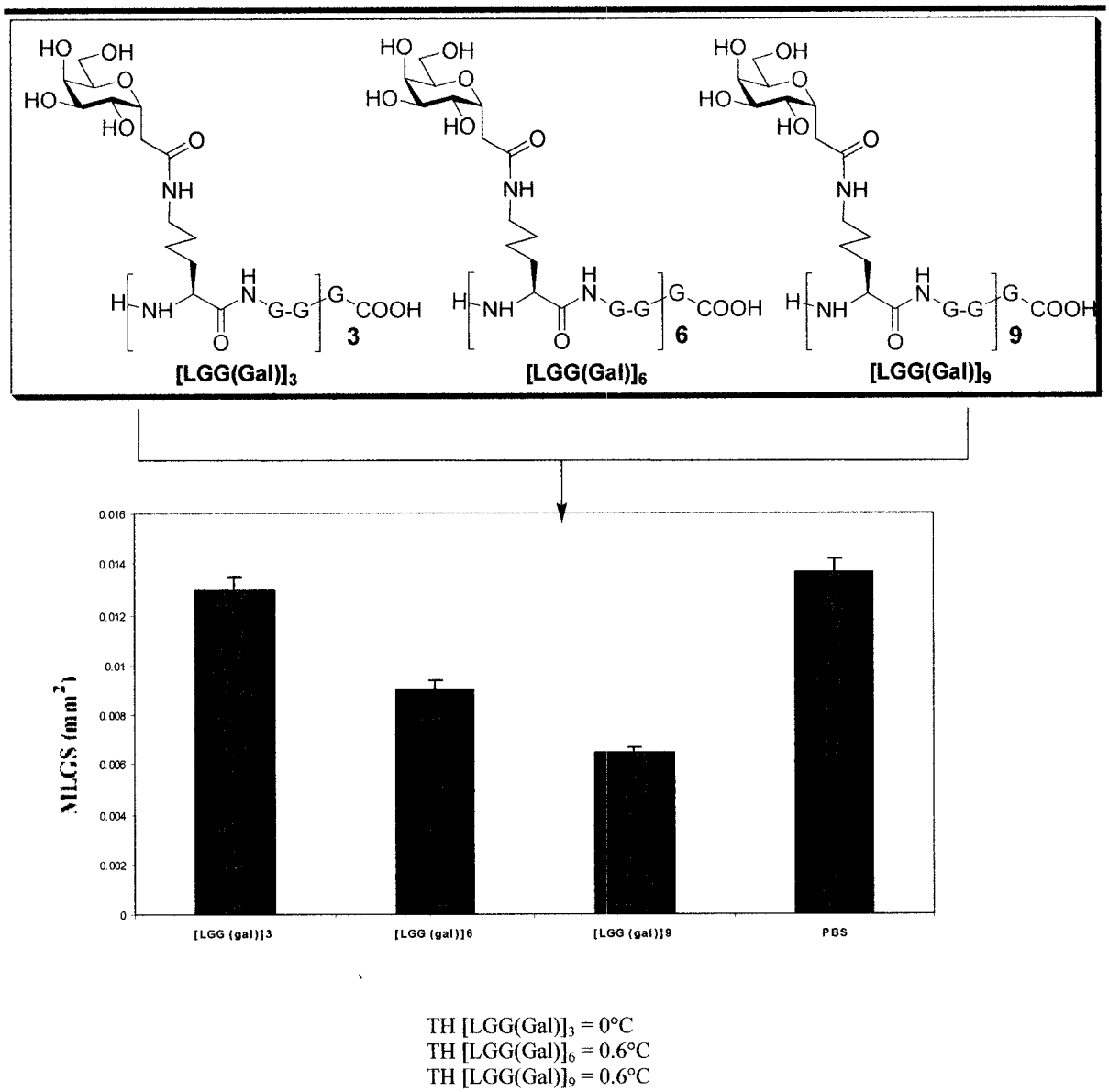


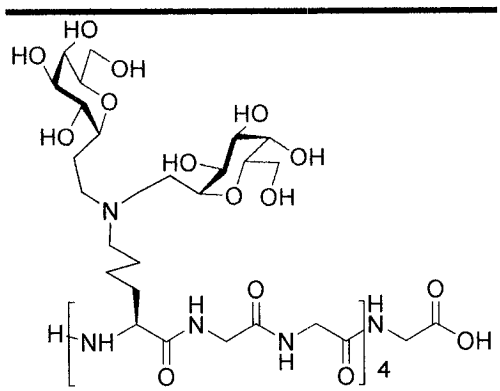
Figure 25: Analogues to study the influence of the molecular weight of the polymer[181] *(each analogue was tested with constant carbohydrate concentration).*

Figure 25 represents the recrystallisation inhibition activity where the X-axis shows the different analogues tested and the Y-axis represents the mean largest grain size (mm^2) of the ice crystals after 30 minutes annealing time. The activity of the analogue increases as the size of the crystals decreases. The results shown in Figure 25 demonstrate that glycopeptides

with higher molecular weights exhibit higher RI activity. The 9-mer is more active than the 6-mer, which is more active than the 3-mer. This correlates well with the previous TH studies. As the different analogues were tested with constant carbohydrate concentration any non specific effects on activity are ruled out. Indeed, even though $[LGG(Gal)]_9$ was tested with a peptide concentration three times lower than $[LGG(Gal)]_3$; it is still exhibited higher activity.

b) Carbohydrate concentration with respect to the individual backbone:

Nishimura studies indicate that there is a loss of thermal hysteresis activity with the use of a monosaccharide moiety compared to a disaccharide moiety. The analogue proposed in Figure 26 shows the importance of the carbohydrate concentration with respect to the individual backbone on TH and RI activity.

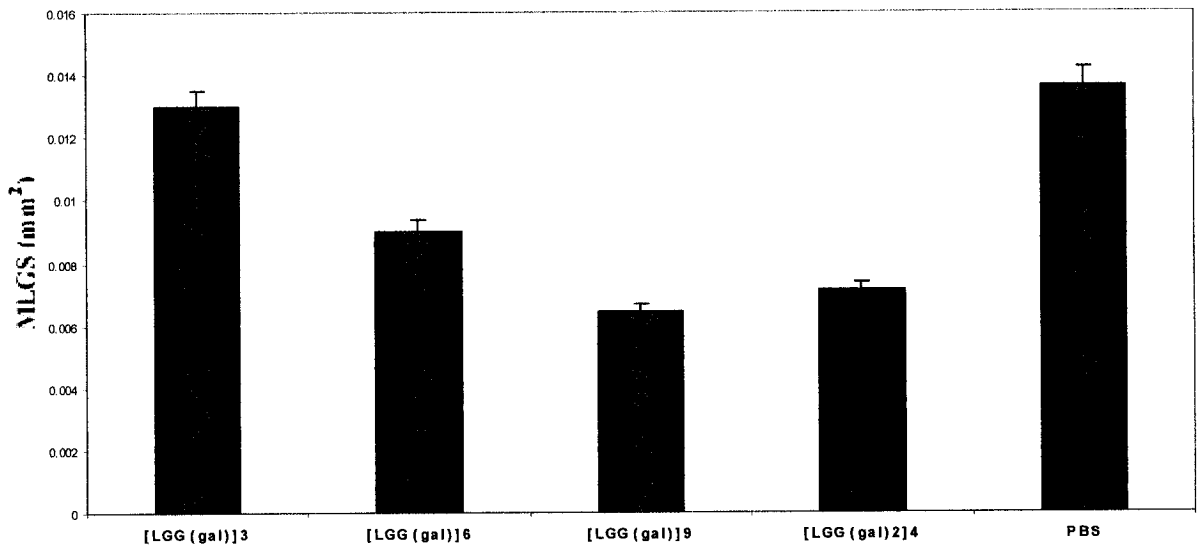


$[LGG(Gal)_2]_4$

Figure 26: Divalent analogue for the study of the influence of carbohydrate concentration[181]

In Graph 5, we also observed that the 4-mer divalent $[LGG(Gal)_2]$ is more active than the 6-mer of the monovalent $[LGG(Gal)]_6$, even though they were tested at equal carbohydrate concentrations. This could suggest that the carbohydrate concentration with respect to the individual backbone is another important factor for RI activity and that doubling the carbohydrate concentration on the peptide backbone could more than double the activity (with constant carbohydrate concentration). This might also suggest that there is some type of cumulative effect due to the carbohydrate moiety. However, to confirm this last suggestion several other analogues would need to be synthesized in order to fully compare $[LGG(Gal)]_6$ and $[LGG(Gal)_2]_4$. First, a 4-mer of the LGG(Gal) analogue would be

necessary. Second, we would also need to investigate the functional influence (i.e. hydrogen bonding) and structural influence (i.e. conformation) of the amide bond. Toward this goal the following AFGP analogues would be necessary (Figure 27).



	$[LGG(Gal)]_3$	$[LGG(Gal)]_6$	$[LGG(Gal)_2]_4$	$[OGG(Gal)]_4$
Peptide Concentrations (mmol/L)	0.007	0.0033	0.00105	0.0055
Carbohydrate Concentrations (mmol/L)	0.021	0.021	0.021	0.022

Graph 5: Recrystallisation inhibition of different molecular weight analogues[181]

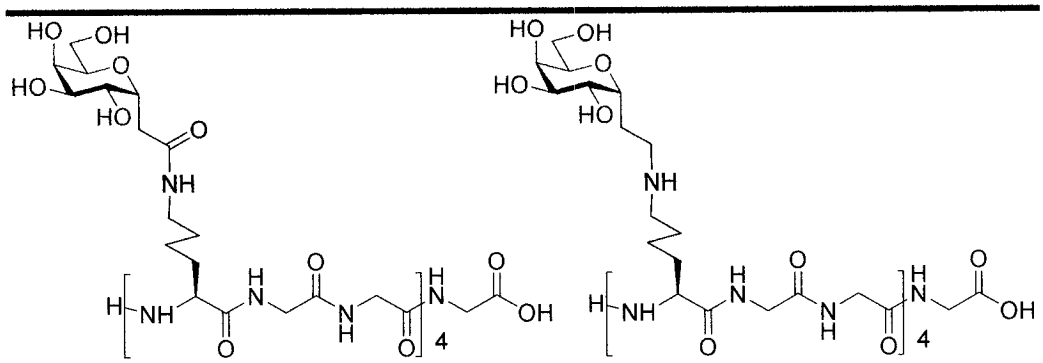


Figure 27: Analogues for the study of the influence of the carbohydrate concentration

1.7.4.2 *Influence of the stereochemistry of the carbohydrate moiety and the length of the peptide side chain on RI activity*[181]

a) *Influence of the stereochemistry of the carbohydrate moiety*

Different carbohydrates were tested in order to elucidate the importance of the conformation of the carbohydrate moiety as well as to observe the influence of the relative position of hydroxyl groups (hydrophile) (Figure 28) on recrystallization inhibition and thermal hysteresis activity.

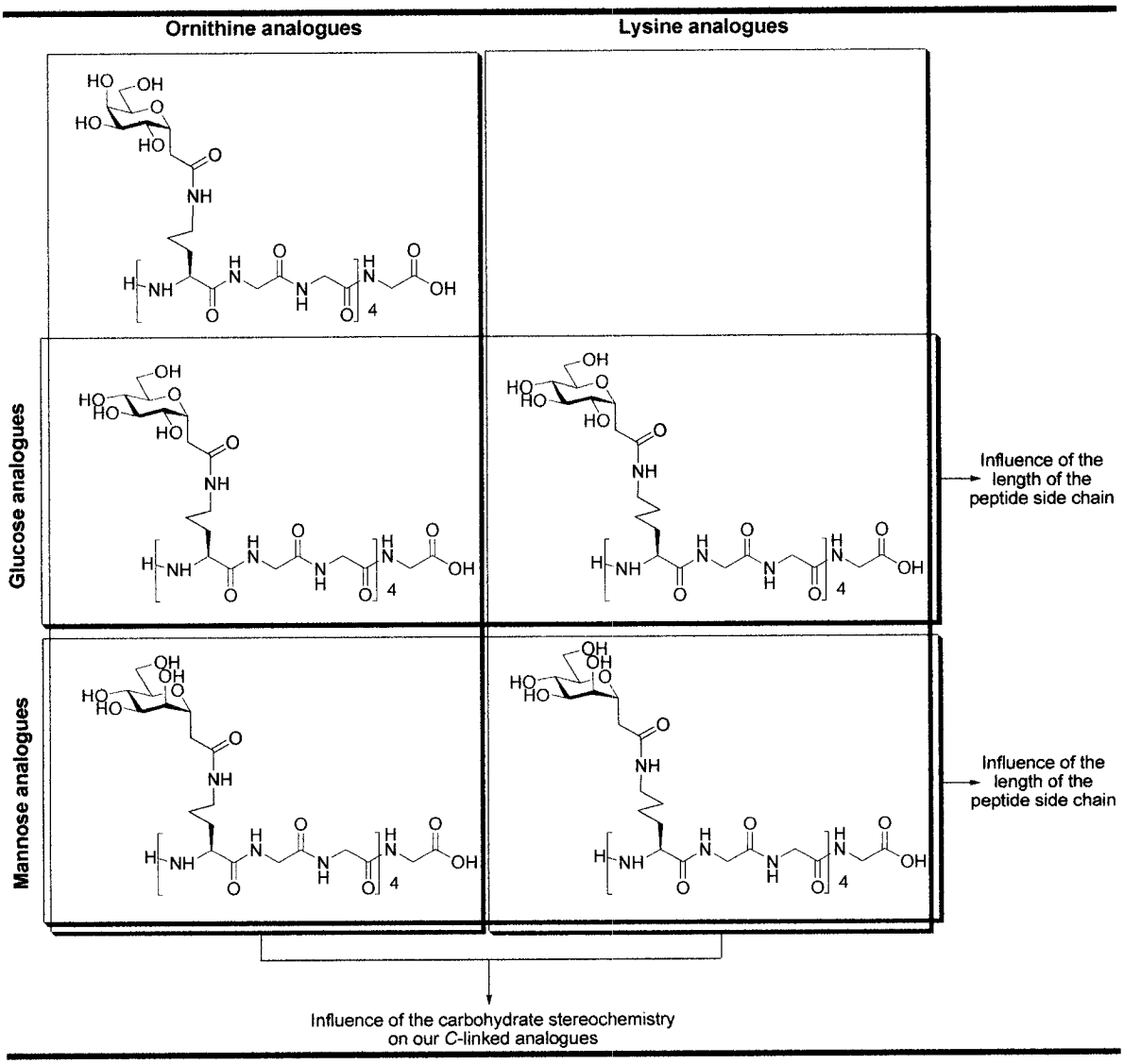
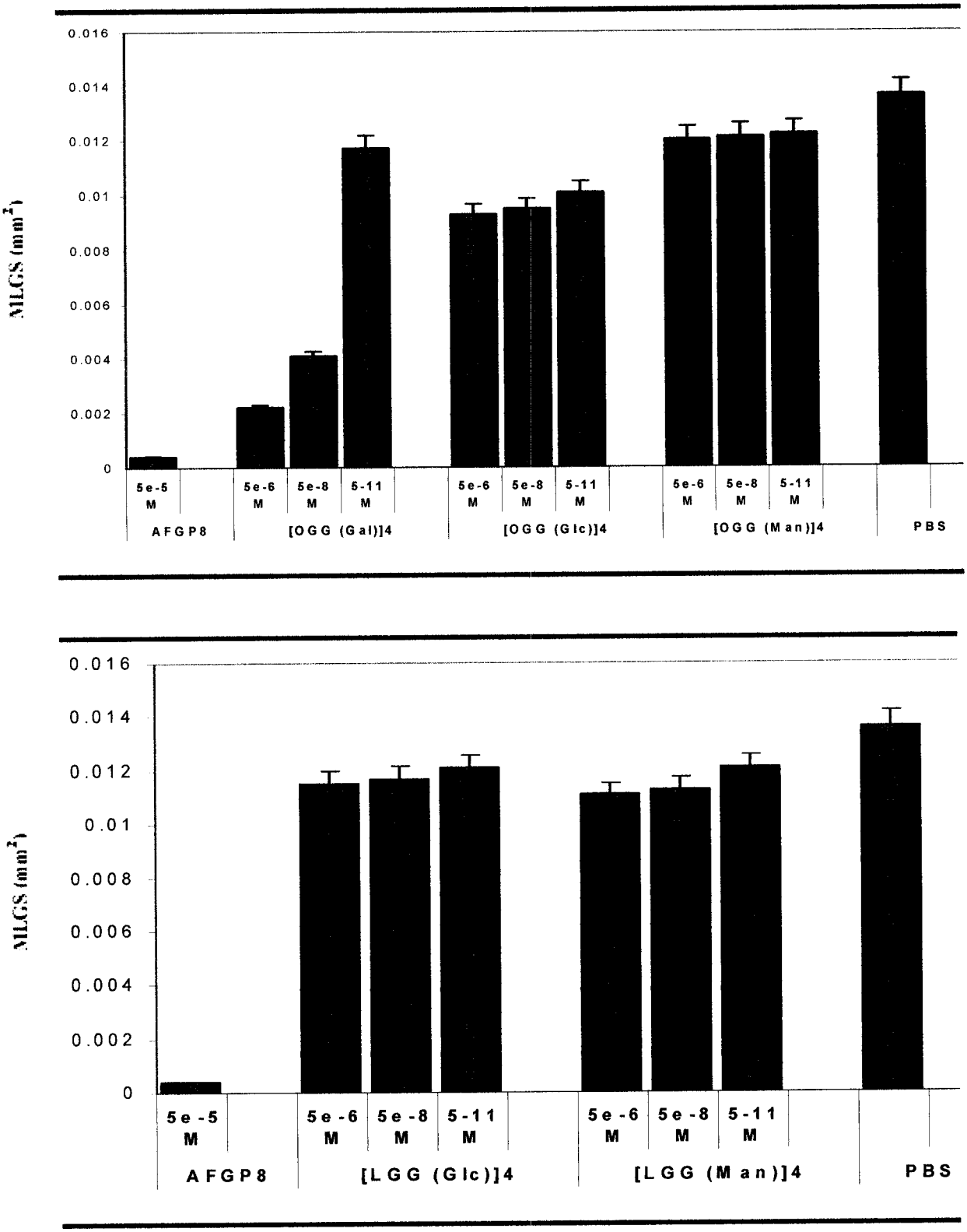


Figure 28: Analogues designed to test the influence of the stereochemistry of the carbohydrate moiety[181]



Graph 6: Recrystallisation inhibition of different carbohydrate analogues[181]

The results of Graph 6 demonstrate that the ornithine based glycopeptides containing galactose tend to be more active than those with glucose or mannose. The synthesis of $[LGG(Gal)]_4$ would provide evidence for or refute this hypothesis by adding another set of AFGP8 analogues. It also seems that using mannose as a carbohydrate moiety results in complete loss of activity. An interesting fact is that none of these analogues exhibit TH or dynamic ice shaping (DIS).

b) Influence of the length of the peptide side chain:

After realizing that these analogues possessed antifreeze activities, the first modification was to investigate the influence of the length of the linkage between the carbohydrate moiety and the peptide backbone (using ornithine or lysine). To accomplish this, three sets of analogues were prepared and compared (Figure 28).

In this study the $[OGG(man)]_4$ and $[LGG(man)]_4$ are both within each other SEM (standard error of the mean) error bars, therefore they exhibit the same activity. $[OGG(glc)]_4$ is more active that its lysine counterpart and the $[OGG(Gal)]_4$ is more active than $[LGG(Gal)]_6$. This suggests that ornithine analogues which are closer in length to arginine analogues (Figure 1) are more active than their lysine counterparts. To confirm this $[LGG(Gal)]_4$ should be synthesized and studied. None of these analogues exhibit any TH activity; therefore the length of the peptide side chain might also be one of the keys in order to tailor TH and RI activities.

1.7.4.3 Influence of the structural features of the side chain of the on RI

a) New side chain linkage (serine and serine analogues)[123]

The amide linkage using ornithine was a good mimic of the guanidinium functionality found in the native system (glycosylated arginine analogue (Figure 1)). A second generation of *C*-linked analogues was then designed in order to mimic the threonine *O*-linked analogues. For this study it was decided to prepare an analogue containing serine instead of threonine (Figure 29). The reason for this modification was the synthetic difficulties relating to the additional stereocenter in the threonine structure. Up to now, only

one synthesis of a glycoside *C*-linked to threonine has been reported and it involved more than 18 steps and generated only the β anomer. [182]

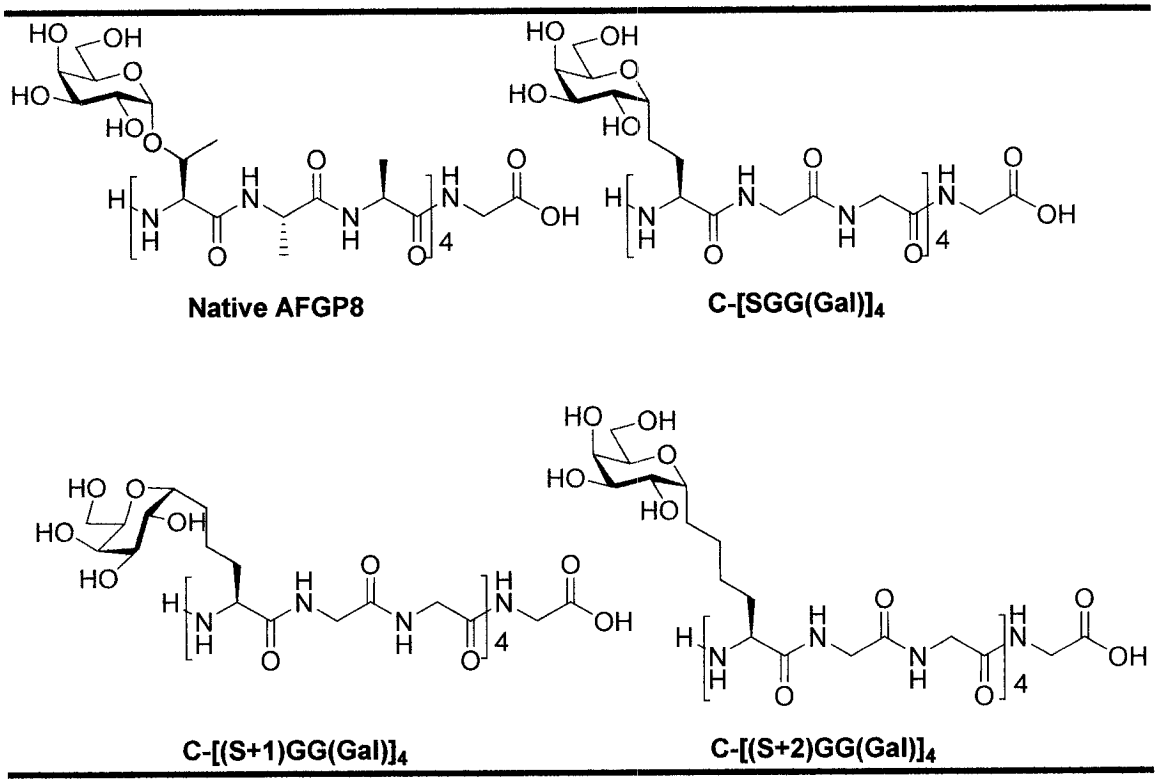
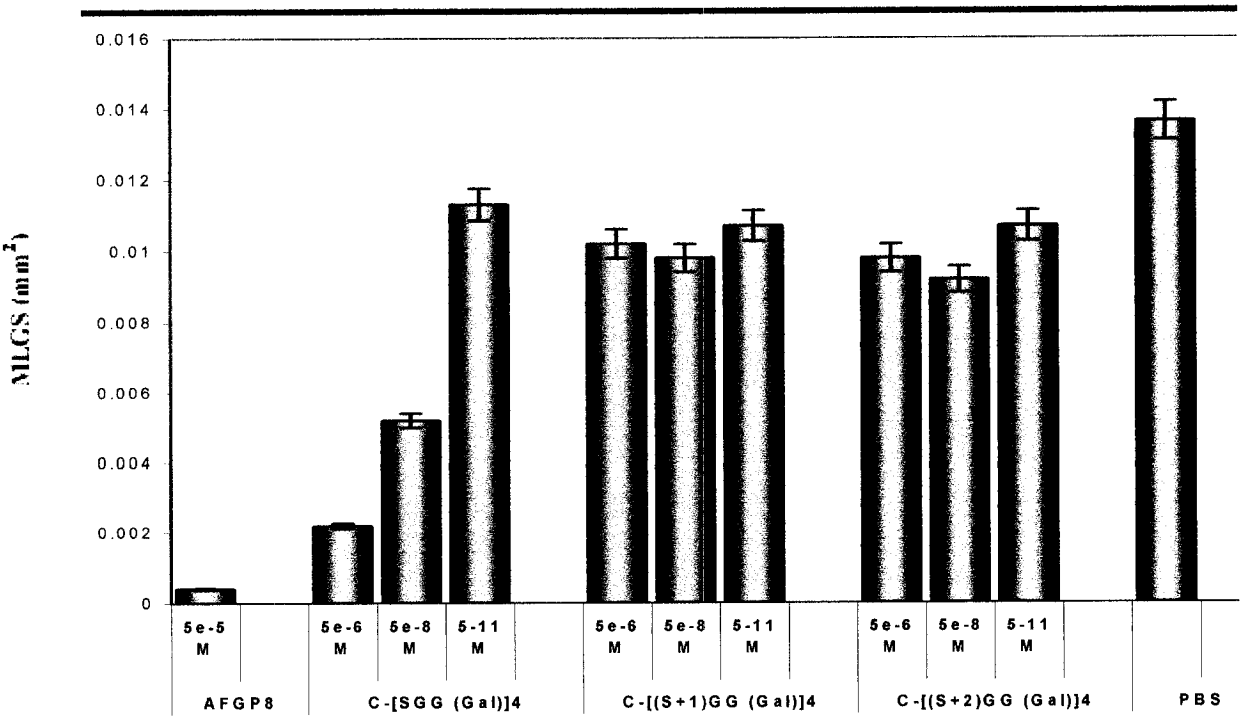


Figure 29: Second generation of AFGP8 analogues: C-linked serine and serine analogues [123]

Although all three *C*-linked serine analogues demonstrated recrystallization inhibition activity relative to the PBS negative control, the serine *C*-linked analogue was the most potent with an activity close to the native AFGP8 (Graph 7). This increased RI activity suggests that an optimal distance between the carbohydrate and the peptide moieties exists and plays a key role. The serine *C*-linked analogue (*C*-[SGG(Gal)]$_4$) possesses the same number of atoms between the carbohydrate and peptide moieties as native AFGP.



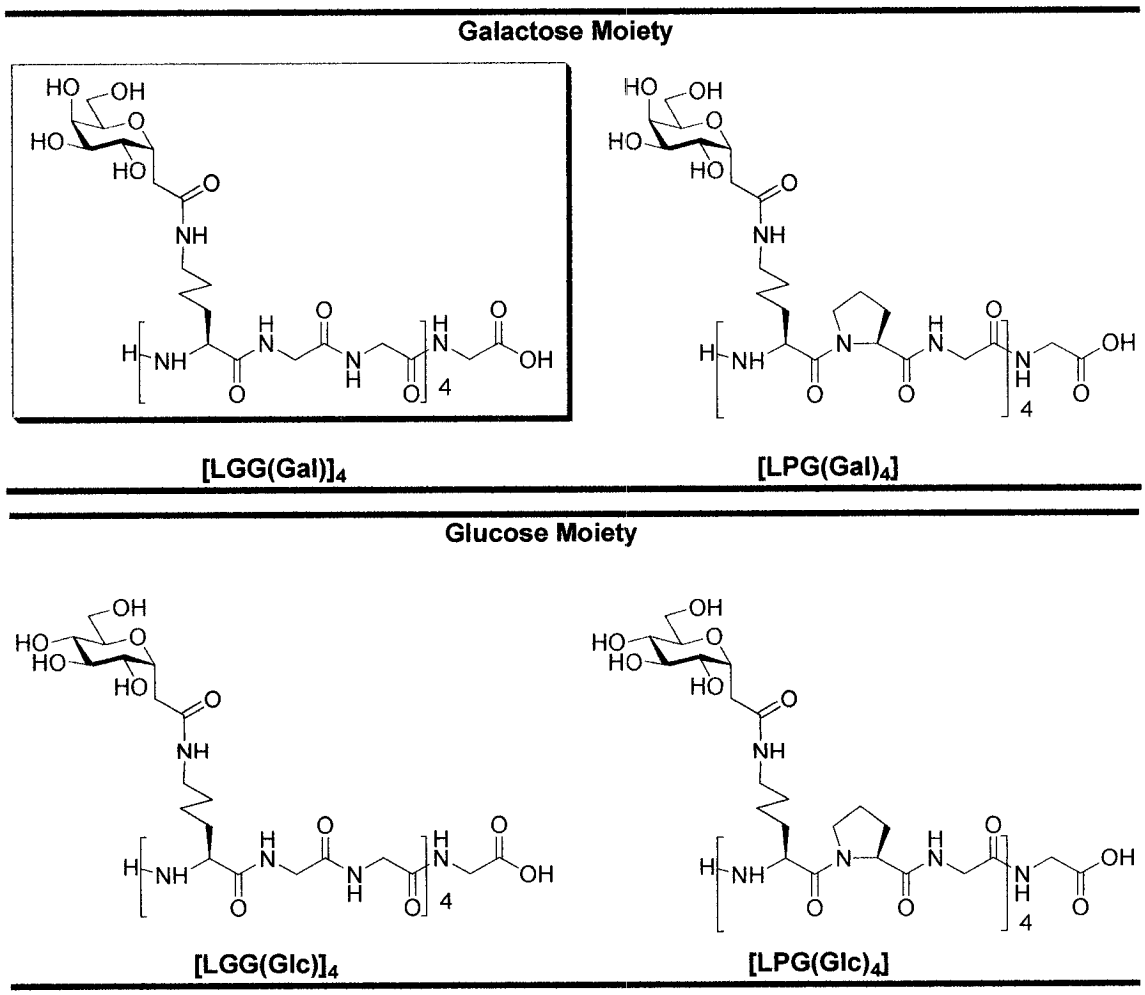
Graph 7: Recrystallisation inhibition of different side chain length on C-linked serine and serine-type analogues[123]

In these previously prepared analogues, it seems that a shorter side chain length, which is similar to that in native AFGP8, is important for *C*-linked AFGP analogues to show potent RI activity. However, the side chain length is not the only factor that can affect RI activity. Previously prepared analogue $[OGG(Gal)]_4$ contains a side chain that is much longer even than that in analogue *C*-$[(S+2)GG(Gal)]_4$, but still demonstrated similar activity to analogue *C*-$[SGG(Gal)]_4$. Compared to the flexible sidechains in serine analogues, a relatively rigid amide bond with partial double bond character exists in the side chain of analogue $[OGG(Gal)]_4$. The rigidity of the side chain in AFGP analogues can directly affect the orientation of the carbohydrate moieties, and previous studies have indicated that the RI activity of *C*-linked AFGP analogues is proportional to the carbohydrate concentration. Therefore, the potent RI activity of $[OGG(Gal)]_4$ analogue is presumably due to the fact that the rigid side chain can limit the extent of free rotation of the carbohydrate moiety and stabilize it in a certain orientation, so as to increase vicinal carbohydrate concentrations. On the other hand, in analogues *C*-$[(S+2)GG(Gal)]_4$ and *C*-$[(S+2)GG(Gal)]_4$, the side chains are

too flexible to keep the carbohydrate moiety in a certain orientation, thus causing these two analogues to show very limited RI activities.

1.7.4.4 Influence of the peptide sequence on RI

In AFGP7 and AFGP8 alanine can be substituted by proline. The activity of the resulting compounds is quite similar to the one possessing alanine. In order to examine this modification, previous Ph.D. students prepared different analogues containing the amino acid proline (Figure 30 and Graph 8). This is the first step in a larger combinatorial investigation of the influence of the peptide sequence as well as the influence of this sequence on the conformation of our analogues. These observations should determine which conformations are favourable for TH and RI activities.



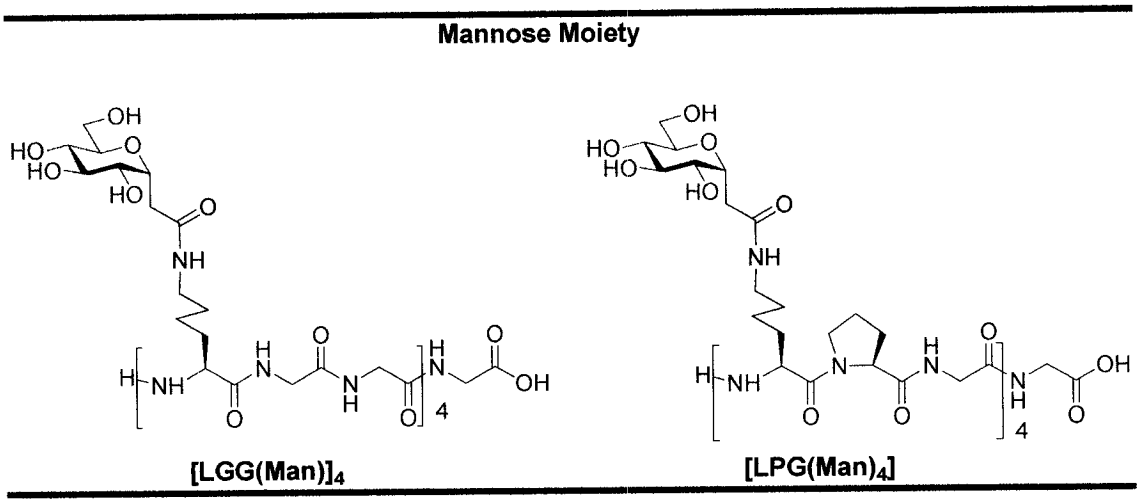
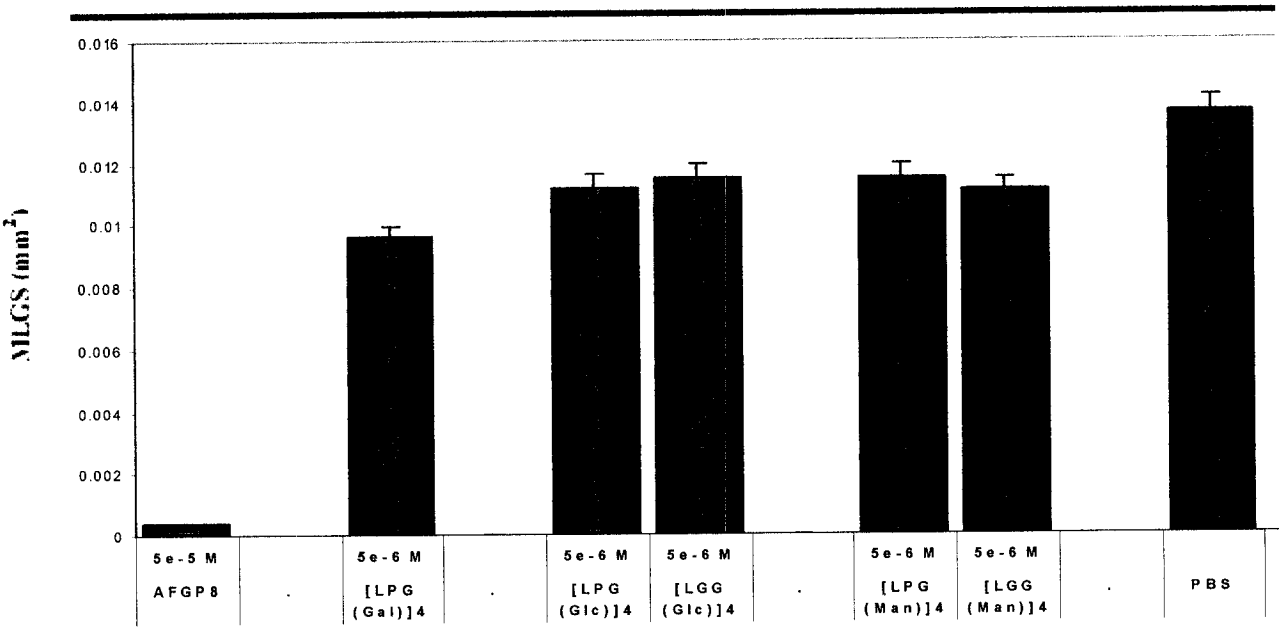


Figure 30: Compounds designed to study of the influence of the peptide sequence[181]



Graph 8: Recrystallization inhibition of different peptide sequence analogues[181]

The small peptide modification to replace glycine with proline (present in the native system) does not seem to have an effect in the glucose or the mannose analogues. However, it will be interesting to synthesize $[LGG(Gal)]_4$ to see its impact on the galactose analogues to confirm this trend.

1.7.4.5 Influence of different structural modifications on TH and dynamic ice shaping.

Interestingly only the [LGG(Gal)] 9-mer and 6-mer exhibit TH, suggesting that the length of the glycopeptide has a strong influence on TH activity (Graph 9). This suggests that the 4-mer would be a good compromise between low TH and high RI activity. Moreover the disaccharide unit seems to be essential for the TH activity but not necessary for the RI activity. Although $[OGG(Gal)]_4$ exhibits dynamic ice shaping within the TH gap, this compound does not spiculate at lower temperatures and should not damage the cells during freezing. So far $[OGG(Gal)]_4$ and *C*-$[SGG(Gal)]_4$ seem to be the more promising compounds in order to tailor thermal hysteresis and recrystallisation inhibition activity.

Analogues	n	Ice Shaping	TH	RI
Lys-Gly-Gly (Gal)n	3	None	×	✓
Lys-Gly-Gly (Gal)n	6	Hexagon	✓(0.06°C)	✓
Lys-Gly-Gly (Gal)n	9	Hexagon	✓ (0.06°C)	✓
Lys-Gly-Gly (Glc)n	4	None	×	×
Lys-Gly-Gly (Man)n	4	None	×	×
Orn-Gly-Gly (Gal)n	4	Hexagon	×	✓
Orn-Gly-Gly (Glc)n	4	None	×	×
Orn-Gly-Gly (Man)n	4	None	×	×
Lys-Pro-Gly (Gal)n	4	None	×	✓
Lys-Pro-Gly (Glc)n	4	None	×	×
Lys-Pro-Gly (Man)n	4	None	×	×
Lys-Gly-Gly bis(Gal)n	4	None	×	✓
C-[SGG(Gal)]n	4	None	×	✓
C-[(S+1)GG(Gal)]n	4	None	×	✓
C-[(S+2)GG(Gal)]n	4	None	×	✓

Graph 9: Thermal hysteresis, dynamic ice shaping and recrystallisation inhibition of our different analogues[123,181]

[1] Devries, A.L. *Biochemical and Biophysical Perspectives in Marine Biology* London; New-York : Academic Press **1974**, 289.

[2] Duman, J.G.; Olsen, T.M. *Cryobiology* **1993**, 30, 322.

[3] Yeh, Y.; Feeney, R.E. *Chem. Rev.* **1996**, 96, 601.

[4] Graham, L.A.; Liou, Y.-C.; Walker, V.K.; Davies, P.L. *Nature* **1997**, 388, 727.

[5] Cheng, C.C.; DeVries, A.L. *In Life Under Extreme Conditions* Springer **1991**, 1.

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

[7] DeVries, A.L.; Komatsu, S.K.; Feeney, R.E. *J. Biol. Chem.* **1970**, 245, 2901.

[8] Fletcher, G.L.; Hew, C.L.; and Davies, P.L. *Annu. Rev. Physiol.* **2001**, 63, 359.

[9] Graham, L. A.; Davies P. L. *Science* **2005**, 310, 5747.

[10] Scholander, P. F.; Flagg, W.; Walters, V.; Irving, L. *Physiol. Zool.* **1953**, 26, 67.

[11] Scholander, P.F.; Dam, L.V.; Kanwisher, J.; Hammel, T.; Gordon, M.S. *J. Cell. Compar. Physiol.* **1957**, 49, 5.

[12] Gordon, M.S.; Amdur, B.H.; Scholander, P.F. *Biol. Bull.* **1962**, 22, 52.

[13] Scholander, P.F.; Maggert, J.E. *Cryobiology* **1971**, 8, 371.

[14] DeVries, A. L.; Wohlshlag, D. E. *Science* **1969**, 163, 1073.

[15] Komatsu, S. K.; DeVries, A. L.; Feeney, R. E. *J. Biol. Chem.* **1970**, 245, 2909.

[16] DeVries, A. L.; Vandenheede, J.; Feeney, R. E. *J. Biol. Chem.* **1971**, 246, 305.

[17] Duman, J.G.; DeVries, A.L. *Cryobiology* **1972**, 9, 469.

[18] Lin, Y.; Duman, J.G.; DeVries, A.L. *Biochem. Biophys. Res. Commun.* **1972**, 46, 87.

[19] Raymond, J.A.; DeVries, A.L. *Cryobiology* **1972**, 9, 541.

[20] DeVries, A.L. *Ann. Rev. Physiol.* **1983**, 45, 245.

[21] Feeney, R.E.; Burcham, T.S.; Yeh, Y. *Ann. Rev. Biophysics. Biophys. Chem.* **1986**, 15, 59.

[22] DeVries, A.L. *Methods Enzymol.* **1986**, 127, 293.

[23] Knight, C.A.; DeVries, A.L. *Atmospheric Aerosols and Nucleation (*Vali, G. &Wagner, P.E., eds), **1988,** 717.

[24] Hardling, M.M.; Anderberg, P.I.; Haymet, A.D.J. *Eur. J. Biochem.* **2003**, 270, 1381.

[25] Davies, P.L.; Sykes, B.D. *Curr. Opin. Struct. Biol.* **1997**, 7, 828.

[26] Hew, C.L.; Yang, D.S.C. *Eur. J. Biochem.* **1992**, 203, 33.

[27] Harding, M.M.; Ward, L.G.; Haymet, A.D.J. *Eur. J. Biochem.* **1999**, 264, 653.

[28] Feeney, R.E.; Yeh,Y. *Trends Food Sci. Technol.* **1998**, 9, 102.

[29] Ewart, K.V.; Lin, Q.; Hew, C.L. *Cell. Mol. Life Sci.* **1999**, 55, 271.

[30] Jia, Z.; Davies, P.L. *Trends Biol. Sci.* **2002**, 27, 101.

[31] Fletcher, G.L.; Hew, C.L.; Davies, P.L. Annu. *Rev. Physiol.* **2001**, 63, 359.

[32] DeVries, A.L. *Science* **1971**, 172, 1152.

[33] Feeney, R.E. *Am. Sci.* **1974**, 62, 712.

[34] O'Grady, S.M.; Schrag, J.D.; Raymond, J.A.; DeVries, A.L. *J. Exp. Zool.* **1982**, 224, 177.

[35] Fletcher, G.L.; Hew, C.L.; Joshi, S.B. *Can. J. Zool.* **1982**, 60, 348.

[36] Burcham, T.S.; Osuga, D.T.; Rao, B.N.N.; Bush, C.A.; Feeney, R.E. *J. Biol. Chem.* **1986**, 261, 6384.

[37] Hew, C.L.; Slaughter, D.; Fletcher, G.L..; Joshi, S.B. *Can. J. Zool.* **1981**, 59, 2186.

[38] Geoghegan, K.F.; Osuga, D.T.; Ahmed, A.I.; Yeh ; Y. Feeney, R.E. *J. Biol. Chem.* **1980**, 255, 663.

[39] Wohrmann, A.P.A. *Mar. Ecol. Prog. Series* **1996**, 130, 47.

[40] Wu,Y.; Banoub, J.; Goddard, S.V.; Kao, M.H.; Fletcher, G.L. *Comp. Biochem. Physiol.* **2001**, B 128, 265.

[41] Hsiao, K.C.; Cheng, C.H.C.; Fernandes, I.E.; Detrich, H.W.; DeVries, A.L. *Proc. Natl. Acad. Sci. U.S.A.* **1990**, 87, 9265.

[42] Ewart, K.V.; Lin, Q.; Hew, C.L. *Cell. Mol. Life. Sci.* **1999**, 55, 271.

[43] Duman, J.G.; DeVries, A.L. *Nature* **1974**, 247, 237.

[44] Duman, J.G. ; DeVries, A.L. Comp. *Biochem. Physiol.* **1976**, B 54, 375.

[45] Ng, N.F.L.; Trinh, K.-Y.; Hew, C.L. *J. Biol. Chem.* **1986**, 261, 15690.

[46] Ng, N.F.L.; Hew, C.L. *J. Biol. Chem.* **1992**, 267, 16069.

[47] Jia, Z.C.; Deluca, C.I.; Davies, P.L. *Protein Sci.* **1995**, 4, 1236.

[48] DeLuca, C.I.; Davies, P.L.; Ye, Q.; Jia, Z. *J. Mol. Biol.* **1998**, 275, 515.

[49] Deng, G.J.; Andrews, D.W.; Laursen, R.A. *FEBS Lett.* **1997**, 402, 17.

[50] Deng, G.J.; Laursen, R.A. *Biochim. Biophys. Acta* **1998**, 1388, 305.

[51] O'Grady, S.; Clark, A.; DeVries, A.L. *Exp. Zool.* **1982**, 220, 179.

[52] Alghren, J.A.; Cheng, C.-H.C.; Schrag, J.D.; DeVries, A.L. *J. Exp. Biol.* **1988**, 137, 549.

[53] Valerio, P.F.; Kao, M.H.; Fletcher, G.L. *Can. J. Zool.* **1990**, 68, 1065.

[54] Valerio, P.F.; Kao, M.H.; Fletcher, G.L. *J. Exp. Biol.* **1992**, 164: 135.

[55] Hobbs, P. V. *Ice Physics*; Clarendon Press: Oxford, **1974.**

[56] Franks, F.; Darlington, J.; Schenz, T.; Mathias, S. F.; Slade, L.; Levine, H. *Nature* **1987**, 325, 146.

[57] Wilson, P. W.; Leader, J. P. *Biophys. J.* **1995**, 68, 2098.

[58] Kuroda, T. In Proc. 4th Topical Conference on Crystal Growth Mechanisms; Hokkaido Press: Japan, **1991**, 157.

[59] Knight, C.A.; DeVries, A.L. *J. Cryst.Growth* **1994**, 143, 301.

[60] Brown, R.A.; Feeney, R.E. *Biopolymers* **1985**, 24, 1265.

[61] Wierzbicki A.; Taylor M.S.; Knight C.A.; Madura J.D.; Harrington J.P.; Sikes C.S. *Biophys. J.* **1996**, 71, 8.

[62] Raymond J.A.; Wilson P.; DeVries A.L. *Proc. Natl. Acad. Sci.* **1989**, 86: 881.

[63] Wen D.; Laursen R.A. *Biophys. J.* **1992**, 63: 1659.

[64] Sonnichsen F.D.; DeLuca C.I.; Davies P.L.; Sykes B.D. *Structure* **1996**, 4: 1325.

[65] Raymond, J.A.; DeVries, A.L. *Proc. Natl. Acad. Sci. USA* **1977**, 74: 2589.

[66] Knight, C.A.; Cheng, C.-H.C.; DeVries, A.L. *Biophys. J.* **1991**, 59: 409.

[67] Knight, C.A.; Driggers, E.; DeVries, A.L. *Biophys. J.* **1993**, 64: 252.

[68] Wilson, P.W.; Beaglehole, D.; Devries, A.L. *Biophys. J.* **1993**, 64, 1878.

[69] Wilson, P. *Cryo-letters*, **1993**, 14, 31.

[70] Knight, C.A.; *Nature*, **2000**, 406, 250.

[71] Hall, D.G.; Lips, A. *Langmuir* **1999**, 15, 1905.

[72] Fairley, K.; Westman, B.J.; Pham, L.H.; Haymet, A.D.J.; Harding, M.M.; Mackay, J.P. *J. Biol. Chem.* **2002**, 277, 24073.

[73] Haymet, A.D.J.; Ward, L.G.; Harding, M.M.; Knight, C.A. *FEBS Lett.* **1998**, 430, 301.

[74] Chao, H.; Houston, M.E.; Hodges, R.S.; Kay, C.M.; Sykes, B.D.; Loewen,M.C.; Davies, P.L.; Sönnichsen, F.D. *Biochemistry* **1997**, 36, 14652.

[75] Haymet, A.D.J.; Ward, L.G.; Harding, M.M. *J. Am.Chem. Soc.* **1999**, 121, 941.

[76] Baardsnes, J.; Kondejewski, L.H.; Hodges, R.S.; Chao, H.; Kay, C.; Davies, P.L. *FEBS Lett.* **1999**, 463, 87.

[77] Chou, K.C. *J. Mol. Biol.* **1992**, 223, 509.

[78] Jorgensen,H.; Mori, M.; Matsui, H.; Kanaoka, M.; Yanagi, H.; Yabusaki, Y.; Kikuzono, Y. *Protein Eng.* **1993**, 6, 19.

[79] Karim, O.A.; Haymet, A.D.J. *J. Chem. Phys.* **1988**, 89, 6889.

[80] Bryk T.; Haymet A.D.J. *J. Mol. Liq.* **2004**, 112 (1-2), 47.

[81] Haymet A.D.J. Unpublished results **2000**.

[82] Madura, J.; Baran, K.; Wierzbicki, A. *J. Mol. Recognit.* **2000**, 13, 101.

[83] Knight, C.A.; Hallett, J.; DeVries, A. L. *Cryobiology* **1988**, 25, 55.

[84] McKown, R.L.; Warren, G.J. *Cryobiology* **1991**, 28, 474.

[85] Yeh, Y.; Feeney, R. E.; McKown, R.L.; Warren, G.J. *Biopolymers* **1994**, 34, 1495.

[86] Knight, C.A.; Wen, D.-Y.; Laursen, R.A. *Cryobiology* **1995**, 32, 23.

[87] 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.

[88] Rubinsky, B.; Arav, A.; Devries, A.L. *Cryobiology* **1992**, 29, 69.

[89] Storey, K.B.; Bischof, J.; Rubinsky, B. *Am. J. Physiol.* **1992**, 263, R185.

[90] Hincha, D.K.; Devries, A.L.; Schmitt, J.M. *Biochim. Biophys. Acta* **1993**, 1146, 258.

[91] Cheng, C.; Devries, A.L. *Cryobiology* **1992**, 29, 783.

[92] Payne, S.R.; Oliver, J.E.; Upreti, G.C. *Cryobiology* **1994**, 31, 180.

[93] Hays, L.; Feeney, R.E.; Crowe, L.M.; Crowe, J.H.; Oliver, A.E. *Proc. Natl. Acad. Sci. USA* **1996**, 93, 6835.

[94] Quinn, P.J. *Cryobiology* **1995**, 22, 128.

[95] Clerc, S.G.; Thompson, T.G. *Biophys. J.* **1995**, 68, 2333.

[96] Excerpted from the website http://www.physik.uni-bielefeld.de/~olenz/

[97] Wu, Y.; Fletcher, G.L. *Biochim. Biophys. Acta* **2000**, 1524, 11.

[98] Knight, C. A.; Driggers, E.; Devries, A. L. *Biophys. J.* **1993**, 64, 252.

[99] Franks, F.; Morris, E.R. *Biochim. Biophys. Acta.* **1978**, 540, 346.

[100] Griffith, M.; Ewart, K.V. *Biotech. Adv.* **1995**, 13, 375.

[101] Feeney, R.E.; Yeh, Y. *Trends Food Sci. Technol.* **1998**, 9, 102.

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

[103] Baust, J.M. *Cell Preservation Technol.* **2002**, 1, 17.

[104] Glander, A.J.; Schaller J. *Mol. Hum. Reprod.* **1999**, 5, 109.

[105] Baust, J.M.; Van Buskirk, R.G.; Baust, J.G. *In Vitro Cell. Dev. Biol. Animal* **2000**, 36, 262.

[106] Fowke, K.R.; Behnke, J.; Hanson, C.; Shea, K.; Cosentino, M. *J. Immunol. Meth.* **2000**, 244, 139.

[107] Hilbert, S.L.; Luna, R.E.; Zhang, J.; Wang, Y.; Hopkins, R.A.; Yu, Z.X.; Ferran, V.T. *J. Thorac. Cardiovasc. Surg.* **1999**, 117, 454.

[108] Villalba, R.; Pena, J.; Luque, E.; Gomez-Villagran, J.L. *Cryobiology* **2001**, 43, 81.

[109] Mazur, P. *J. Gen. Physiol.* **1963**, 47, 347.

[110] Arav, A.; Yavin, S.; Zeron, Y.; Natan, D., Dekel, I.; Gacitua, H. *Mol. Cell. Endocrinol.* **2002**, 187, 77.

[111] Marsland, T. P.; Evans, S.; and Pegg, D. E. *Cryobiology* **1981**, 24, 311.

[112] Robinson, M. P.; Pegg, D. E. *IEEE Trans. Biomed. Eng.* **1999**, 46, 1413.

[113] Pegg, D. E. *Semin. Reprod. Med.* **2002**, 20, 5.

[114] Rubinsky, B.; Arav, A.; Devries, A. L. *CryoLetters* **1991**, 12, 93.

[115] Pickering, S.J.; Braude, P.R.; Johnson, M.H.; Can, A.; Currie, J. oocyte. *Fertil. Steril.* **1990**, 54, 102.

[116] Pickering, S.J.; Johnson, M.H. *Hum. Reprod.* **1987**, 2, 207.

[117] O'Neil, L.; Paynter, S.J.; Fuller, B.J.; Shaw, R.W.; DeVries, A.L. *Cryobiology* **1998**, 37, 59.

[118] Vincent, C.; Johnson, M.H. *Oxford Rev. Reprod. Biol.* **1992**, 14, 73.

[119] Arnott, J. *On the Treatment of Cancer by Regulated Application of an Anesthetic Temperature.* Churchill, London, **1851**.

[120] Koushafar, H.; Rubinsky, B. *Urology* **1997**, 49, 421.

[121] Pham, L.; Dahiya, R.; Rubinsky, B. *Cryosurgery* **1999**, **38**, 169.

[122] Wang, J.-H. *Cryobiology* **2000**, 41, 1.

[123] Liu S. Ph.D. thesis, University of Ottawa, Ottawa, **2006**.

[124] Anisuzzaman, A.K.M.; Anderson L. *Carbohydr. Res.* **1988**, 174, 265.

[125] Meldal, M.; Jensen, K. J. *J. Chem. Soc. Chem. Commun.* **1990**, 483.

[126] Filira, F.; Biondi, L.; Scolaro, B.; Foffani, M.T.; Mammi, S.; Peggion, E.; Rocchi, R. *Int. J. Biol. Macromol.* **1990**, 12, 41.

[127] Tsuda, T.; Nishimura, S.I. *Chem. Commun.* **1996**, 24, 2779–2780.

[128] Matsushita T.; Hinou H.; Kurogochi M.; Shimizu H.. Nishimura S.I. *Org. Lett.* **2005**, 7, 877.

Tachibana Y.; Monde K.; Nishimura S.I. *Macromolecules* **2004**, 37, 6771.

Tachibana Y.; Matsubara N.; Nakajima F.; Tsuda T.; Tsuda S.; Monde K.; Nishimura S.I. *Tetrahedron. Lett.* **2002**, 58, 10213.

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

[129] Tseng, P.H.; Jiiang, W.T.; Chang, M.Y.; Chen, S.T. *Chem. Eur. J.* **2001**, 7, 585.

[130] Enaide, A.; Ben, R. N. *Biomacromolecules* **2001**, 2, 557.

[131] Ben, R. N.; Enaide, A.; and Hauer, L. *Org. Lett.* **1999**, 1, 1759.

[132] Eniade, A.; Murphy, A.V.; Landreau, G.; Ben, R.N. *Bioconjugate Chem.* **2001**, 12, 817.

[133] Postema, M. H. D. *Tetrahedron* **1992**, 48, 8545.

[134] Postema, M. *C-Glycoside Synthesis*; CRC Press: Boca Raton, **1995**.

[135] Levy, D.; Tang, C. *The Chemistry of C-Glycosides*; Pergamon: Oxford, **1995**.

[136] San Martin, R.; Tavassoli, B.; Walsh, K. E.; Walter, D. S. and Gallagher, T. *Org. Lett.* **2000**, 2, 4051 and references therein.

[137] Grant, L.; Liu, Y.; Walsh, K. E.; Walter, D. S. and Gallagher T. *Org. Lett.* **2002**, 4, 4623 and references therein.

[138] For a recent review on "Synthetic methods of amino *C*-Glycosides" see Xie, J. *Recent Res. Devel. Organic Chem.* **1999**, 3, 505.

[139] Varki, A. *Glycobiology* **1993**, 3, 97.

[140] Lis, H.; Sharon, N. *Eur. J. Biochem.* **1993**, 218, 1.

[141] Mc Ever, R. *Glycoconjugate J.* **1997**, 14, 585.

[142] Wang, L.X.; Lee, Y.C. *J. Chem. Soc. Perk. Trans. 1*, **1996**,6 ,581.

[143] Hein, M.; Michalik, D.; Langer, P. *Synthesis* **2005**, 20, 3531.

[144] Lemieux, R.U. *Exploration with Sugars: How sweet it was,* American Chemical Society, Washington DC. **1990**.

[145] Haines, A.H.; *Adv. Carbohydr. Chem. Biochem.* **1976**, 33,11.

[146] Williams, J. M.; Richardson, A.C. *Tetrahedron* **1967**, 23, 1329.

[147] Carey, F.A.; Hodgson, K.O. *Carbohydr. Res.* **1970**,10, 463.

[148] Wu T.; Goekjan P.G.; Kishi Y. *J. Org. Chem.* **1987**, 52, 4819.

[149] Houk, K.N.; Eksterowicz, J.E.; Wu, Y.D.; Fuglesang, C.D.; Mitchell, D.B. *J. Am. Chem. Soc.* **1993**, 115, 4170.

[150] Wei, A.; Boy, K. M.; Kishi, Y. *J. Am. Chem. Soc.* **1995**, 117, 9432.

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

[152] Asencio J.L.; Canada F.L.; Cheng, X.; Khan, N.; Mootoo, D.R.; Jimenez-Barbero J. *J. Chem. Eur.* J. **2000**, 6, 1035.

[153] Mikkelsen, L.M.; Hernaiz M.J., Martin-Pastor M. Skrydstrup T.; Jimenez-Barbero J. *J. Am. Chem. Soc.* **2002**, 124, 14940.

[154] Jarreton, O.; Skrydstrup T.; Espinosa J.F.; Jimenez-Barbero J.; Beau J.M. *Chem. Eur. J.* **1999**, 5, 430.

[155] Espinosa, J.F.; Bruix, M.; Jarreton, O.; Skrydstrup T.; Beau, J.M.; Jimenez-Barbero J. *Chem. Eur. J.* **1999**, 5, 442.

[156] Espinosa, J.F., Canada F.J., Asencio, J.L.; Dietrich, H.; Martin-Lomas M., Schmidt R.R.; Jimenez-Barbero J. *Angew. Chem., Int. Ed. Engl.* **1996**, 35, 303.

[157] Petitou, M.; Herault, J.P.; Lormeau, J.C.; Helmboldt, A.; Mallet, J.M.; Sinay, P.; Herbert J.M. *Bioorg. Med. Chem.*, **1998**, 6, 1509.

[158] Xin, Y.C.; Zhang Y.M., Mallet, J.M.; Glaudemans, C.P.J.; Sinay, P. *Eur. J. Org. Chem.*, **1999**, 471.

[159] Glaudemans, C.P.J. *Chem. Rev.* **1991**, 91, 25.

[160] Yang, G.L.; Schmieg, J.; Tsuji, M.; Franck, R.W. *Angew. Chem. Int. Ed.* **2004**, 43, 3818.

[161] Weatherman, R.V.; Mortell, K.I.; Chervenak, M.; Kiessling, L.L.; Toone, E.J. *Biochemistry*, 1996, 35, 3619.

[162] Lowary, T.; Meldal, M.; Helmboldt, A.; Vasella, A.; Bock, K. *J. Org. Chem.* **1998**, 63, 9668.

[163] Herzner, H.; Reipen, T.; Schultz, M.; Kunz, H.; *Chem Rev.* **2000**, 100, 4495.

[164] Seitz, O. *ChemBioChem.* **2000**, 1, 214.

[165] Kunz, H. *Pure Appl. Chem.* **1993**, 65, 1223.

[166] Meldal, M.; Bock, K. *Glycoconjugate J.* **1994**, 11, 59.

[167] Paulsen H. *Angew. Chem. Int. Eg. Engl.* **1990**, 29, 823.

[168] Witte, K.; Seitz, O.; Wong, C.-H. *J. Am. Chem. Soc.* **1998**, 120, 1979.

[169] Wong, C. H.; Schuster, M.; Wang, P.; Sears, P. *J. Am. Chem. Soc.* **1993**, 115, 5893.

[170] Mizuno, M.; Haneda, K.; Iguchi, R.; Muramoto, M. T. K.; Aimoto, S.; Yamamoto, K.; Inazu, T. *J. Am. Chem. Soc.* **1999**, 121, 284.

[171] Deras, I. L.; Takegawa, K.; Kondo, A.; Kato, I.; Lee, Y. C. *Bioorg. Med. Chem. Lett.* **1998**, 8, 1763.

[172] Roberge, J. Y.; Beebe, X.; Danishefsky, S. J. *J. Am. Chem. Soc.* **1998**, 120, 3915.

[173] Cohen-Anifeld, S. T.; Lansbury, P. T. J. *J. Am. Chem. Soc.* **1993**, 115, 10531.

[174] Marcaurelle, L. A.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2001**, 123, 1587.

[175] Reimer, K. B.; Meldal, M.; Kusumoto, S.; Fukase, K.; Bock, K. *J. Chem. Soc., Perkin Trans.* **1993**, 1, 925.

[176] Laczko, I.; Hollosi, M.; Urge, L.; Ugen, K. E.; Weiner, D. B.; Mantsch, H. H.; Thurin, J.; Otvos Jr, L. *Biochemistry* **1992**, 31, 4282.

[177] Shao, N. Xue, J. Guo, Z. *J. Org. Chem.* **2003**, 68, 9003.

[178] Xue, J.; Guo, Z. *J. Org. Chem.* **2003**, 68, 2713.

[179] Sjölin, P.; Elofsson, M.; Kihlberg, J. *J. Org. Chem.* **1996**, 61, 560.

[180] Eniade, A.; Purushotham, M.; Ben, R.N.; Wang, J.B.; Horwath K. *Cell Biochem. Biophys.* **2003**, 38, 115.

[181] Murphy A.V. Ph.D. Thesis, State University of Binghamton, Binghamton, **2004**.

[182] Gustafson, T.; Saxin, M.; Kihlberg, J. *J. Org. Chem.* **2003**, 68, 2506.

Chapter 2

GOALS AND OBJECTIVES

2 GOALS AND OBJECTIVES

Since the 1970's, AFGPs molecules have been thoroughly studied in order to understand their molecular mechanism of action. It is believed that conformation of these compounds plays an essential function during the ice adsorption process which is required for antifreeze activity. While several conformation studies of AFGP8 molecules have been published using different spectroscopic methods, an accepted conformation has failed to emerge and its role on antifreeze activity needs to be addressed.

Although structure activity relationship (SAR) studies seem to be one of the most appropriate tools in order to understand the molecular mechanism of action of AFGP molecules, there are few syntheses of AFGP analogues published in literature. Amongst these publications, even fewer have tested their analogues for thermal hysteresis activity and none have tested them for recrystallization inhibition. Recrystallization inhibition might be considered as the most relevant activity for cell storage at sub-zero temperature, while dynamic ice shaping (DIS) and thermal hysteresis (TH) potentially damage cell membranes.[1] Recently, Sidebottom[2] demonstrated that there was no correlation between these two activities; however, the necessary structural features of AFGPs for RI and TH activities still require to be elucidated.

Although AFPG8 displays attractive RI activity, it also exhibits undesired properties for cryopreservation. For instance, it favours the formation of needle shape ice crystals below the TH gap. Furthermore, AFGP molecules are not truly stable under basic, acid or enzymatic conditions due to the exocyclic oxygen. Consequently, more stable analogues with potent RI activity and no DIS activity should be ideal for cryopreservation.

2.1 AFGP8 conformational studies

2.1.1 Previous conformational analyses of AFGPs

Antifreeze glycoproteins are members of the family of glycoproteins. They possess a standard *O*-glycosidic linkage, where the carbohydrate moiety is a regular β-D-galactosyl-(1,3)-D-*N*-acetylgalactosamine linked to a threonine. These glycoproteins are found in many

deep sea teleost fish.[3,4] While the physiological concentration of these glycoproteins varies throughout the year, they efficiently protect these organisms against cryoinjury and death. The mechanism by which these compounds function is an adsorption-inhibition process where they bind in an irreversible fashion to the surface of an individual ice crystal and effect a localized freezing point depression.[5] Although the macroscopic mechanism of action is well understood, the molecular mechanism of action remains a source of intense debate amongst experts in the field. The role of specific functional groups as well as hydrophilic and/or hydrophobic interactions with ice are at the center of this issue.[6] Another issue which has also received attention is the importance of the solution conformation and its role in the mechanism of action.[4]

During the last twenty years, researchers have attempted to elucidate the secondary and higher order structures of AFGPs in solution. Several complementary techniques have been employed including nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy. Recent NMR studies suggest that random unordered conformations appear to predominate in AFGP8 while both random and 3_{10} helical structures predominate in AFGPs 2-5 at a concentration of 0.03 mM.[7,8,9,10,11,12] However, in the supercooled state, the amount of unordered structure dramatically decreases while extended structures and β-turns increase in AFGP8. In contrast, the unordered conformation dominates in the supercooled state for AFGPs 2-5. In the presence of ice, both AFGP8 and AFGPs 2-5 possess a high percentage of β-turns reflecting a high degree of molecular order. This duality between the degree of order in the liquid state suggests that AFGP requires dynamic flexibility for the purpose of exposing the maximum number of "ice-binding" groups. In contrast, recent NMR and CD studies implied that a significant portion of higher molecular weight AFGPs adopts a left handed polyproline type II (PP II) helix.[13] Although the detailed mechanism of action is still unclear, it is believed that the amphiphatic nature of AFGP in this PP-II conformation is essential for antifreeze protein-specific activity.

AFM studies performed in our laboratory suggested that unlike AFGP 1-5, AFGP8 forms aggregates in solution prior to adsorption onto both hydrophilic and hydrophobic surfaces.[14] While the issue of glycoprotein aggregation is itself very interesting, the implications of AFGP8 aggregation and its effects on solution conformation and antifreeze

activity have not been studied and this remains as a key issue in the field of biological antifreezes.

2.1.2 Goals and objectives for investigating the role of conformation in antifreeze activity

As part of our ongoing studies aimed at elucidating the molecular mechanism of action of AFGP and AFGP analogues, we will study:

- The solution aggregation of AFGP8 using dynamic light scattering (DLS), in order to obtain a better understanding of the aggregation pattern observed in our previous AFM studies.[14]
- The conformations of AFGP8 using circular dichroism (CD) for better insight of the different conformations present in solution. These studies have been performed as a function of both temperature and concentration.

2.2 Synthesis and SAR study of *O*-linked analogues of AFGP8

2.2.1 Glycoconjugates

In general, glycoconjugates are extremely important in molecular biology, biochemistry and pharmacology. They are indeed widely distributed in biological organisms and are involved in a large array of cell functions. The interest in these compounds stems from their applications as glycosidase inhibitors[15] and their attractiveness as intermediates for probing carbohydrate-peptide and/or carbohydrate-lipid interactions[16] in biological systems.[17] In addition the presence of the carbohydrate component can also protect the protein from proteolytic cleavage.[18]

Most of the native glycoconjugates contain either *O*-glycosidic or *N*-glycosidic linkages. Generally, most of the glycosylation sites involved are L-Threonine (Thr) or L-Serine (Ser) in *O*-glycosides and L-asparagine (Asn) in *N*-glycosides. The most common motifs are β-*N*-acetyl-glucosamine or α-*N*-acetyl-galactosamine linked to either serine or

threonine for *O*-linked glycoconjugates and β-*N*-acetyl-glucosamine *N*-linked to asparagines for *N*-linked glycoconjugates (Figure 1).

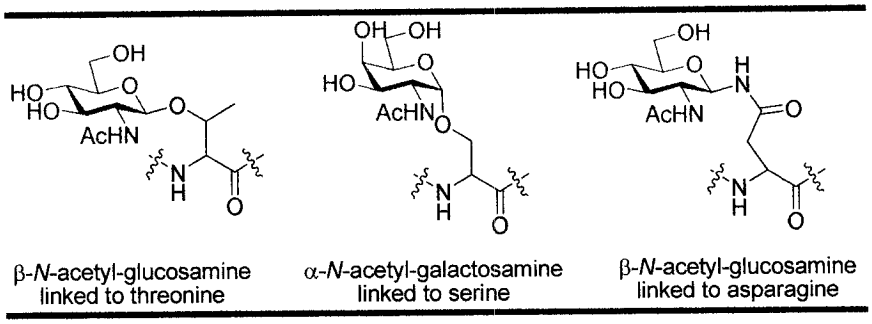


Figure 1: Most common glycosidic linkages

In addition to *O*-linked and *N*-linked glycoproteins, an unusual glycosidic linkage (*C*-linked α-mannoside of tryptophan) was observed in human RNaseU (Figure 2).[19] This unique *C*-linked glycoconjugate is involved in RNA digestion and contains an α-D-mannopyranose residue directly linked to the C-2 atom of the indole ring of the tryptophan.

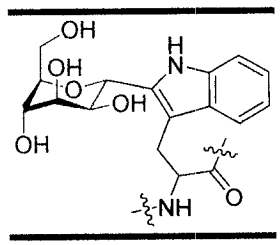


Figure 2: Native C-glycoside of tryptophan in human RNaseU[19]

2.2.2 Synthetic strategies for the preparation of O-linked glycoconjugates.

There currently exist powerful chemical and enzymatic methods for the construction of complex mucin-type glycopeptides.[20] The recent application of techniques developed by protein chemists, such as native and expressed protein ligation, has also permitted the generation of full length glycoproteins. One of the main challenges in the synthesis of *O*-glycosyl amino acids is achieving high stereoselectivity in the formation of the core α-Ser/Thr linkage. Even with simple monosaccharide donors, the outcome of the glycosylation reaction can be difficult to predict. As a result the most commonly employed method for the

synthesis of complex *O*-glycosyl amino acids involves installation of the desired α-*O*-Ser/Thr linkage prior to the elaboration of additional sugars.[21,22]

The attachment of preformed oligosaccharides to simple peptides by chemoselective ligation is an attractive method for the rapid assembly of peptides carrying complex *O*-glycans. Chemoselective ligation reactions are mild and selective, allowing for the coupling of unprotected biomolecules, such as peptides and carbohydrates, in an aqueous environment.[23]

Native chemical ligation[24] as well as expressed protein ligation[25] were also successful in order to obtain such compounds (Figure 3).

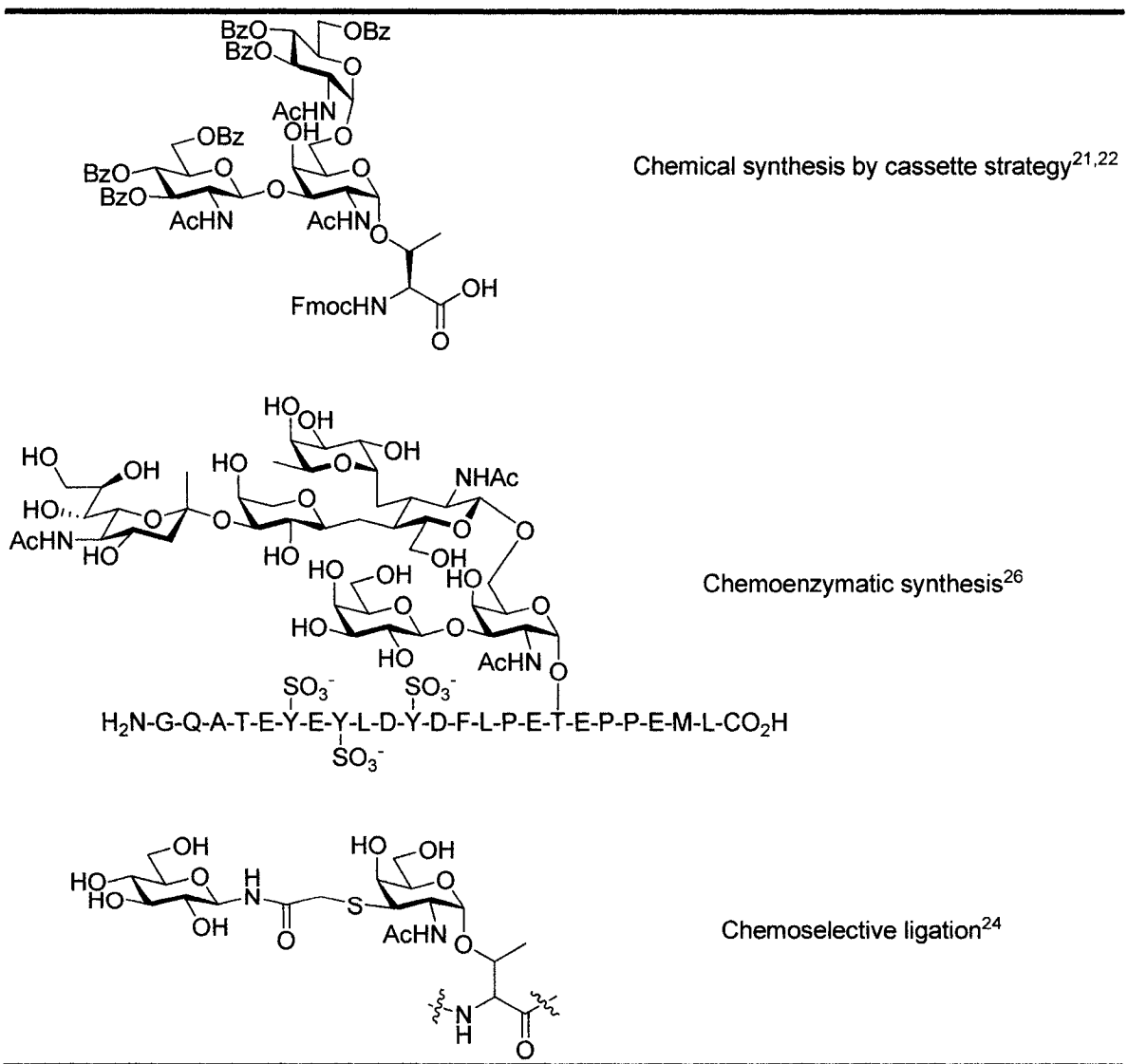


Figure 3: Examples of previously synthesized glycopeptides using different strategies[21,22,24,26]

2.2.3 Goals and objectives.

In order to have a better understanding of the necessary structural features for thermal hysteresis and recrystallization inhibition, a stepwise solid phase synthesis strategy will be used to synthesize different *O*-linked AFGP8 analogues. Our laboratory has designed a family of *O*-linked analogues very similar to the native system (Figure 4). The purpose of these analogues is to investigate the importance of different structural modifications on thermal hysteresis and recrystallization inhibition (Figure 5, Figure 6 and Figure 8).

The following issues will be addressed:

- Determination of the necessity of the disaccharide unit and investigation of the ability of a monosaccharide analogue to exhibit antifreeze activities. The synthesis of monosaccharide analogues would simplify the synthesis and reduce the cost for later industrial applications. This choice is also based upon precedent from earlier structure activity relationship observations by the Feeney group where they reported that the terminal galactose residue is crucial for specific antifreeze activity.[27] Furthermore, in complex polysaccharides it has been recently demonstrated that the terminal carbohydrate is the most important for the binding to receptors.[28]
- Determination of the role of the stereochemistry of the glycosidic linkage on the antifreeze activity (α or β).
- Investigation of the influence of the relative stereochemistry of the hydroxyl groups on TH and RI activities.
- Investigation of the influence of the length of glycopeptide analogues on TH activity. Several 11-mer *O*-linked analogues have been previously tested for TH activity by Nishimura.[29] $[TAA(GalNAc)]_{11}$ exhibits similar TH activity to native AFGP6 of similar length. Comparing the TH activity of the $[TAA(GalNAc)]_4$ to the native AFGP8 of the same length might prove to be interesting.
- Examine the role of the *N*-acetyl functional group in the C-2 glycosidic position on the antifreeze activity and its influence on the conformation.

- Elucidation of the role of the γ-methyl (on the threonine residue) on RI and TH activities. Conformation has been hypothesized to have a strong influence on antifreeze activity. The γ-methyl on the threonine is a strongly hydrophobic group that probably has significant impact on the conformation of AFGPs and their hydrophilic interaction with ice. Consequently the presence or absence of the γ-methyl should significantly impact RI and TH activities. Synthesis and study of an analogue lacking this γ-methyl should help to elucidate its role on such antifreeze activity.
- Analyzis of the role of secondary structure and its relationship to antifreeze activity.

2.2.4 O-linked analogues of AFGP8 and SAR study.

In order to realize these objectives we designed eight monosaccharidic glycopeptides to serve as AFGP analogues (Figure 4).

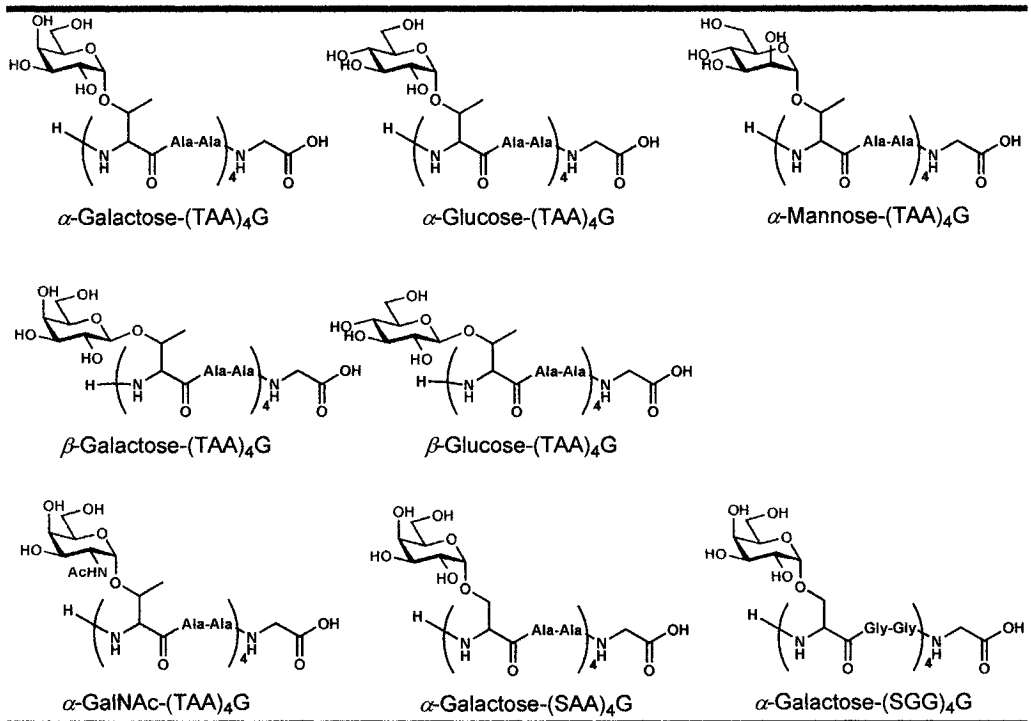


Figure 4: Description of our O-linked analogue targets

Figure 5, Figure 6 and Figure 8 represent how each analogue will be used to examine the influence of different structural modifications on RI and TH.

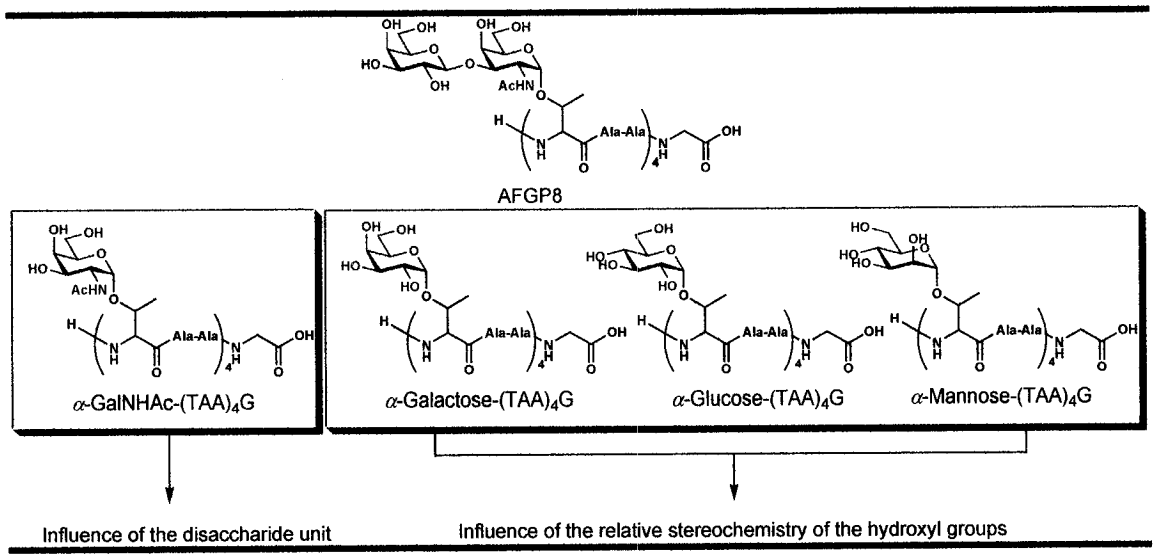


Figure 5: Influence of different structural features on RI and TH activity

Synthesizing analogues with α and β linkages should give us information about the importance of the relative stereochemistry of the carbohydrate on the specific antifreeze activities. The GalNAc analogue should inform us about the action of the *N*-acetyl function on the C-2 position (Figure 5) and finally the serine analogue $[SAA(Gal)]_4$ should give information on the influence of the hydrophobicity and steric of the γ-methyl of the threonine on conformation, RI and TH activities (Figure 5 and Figure 7).

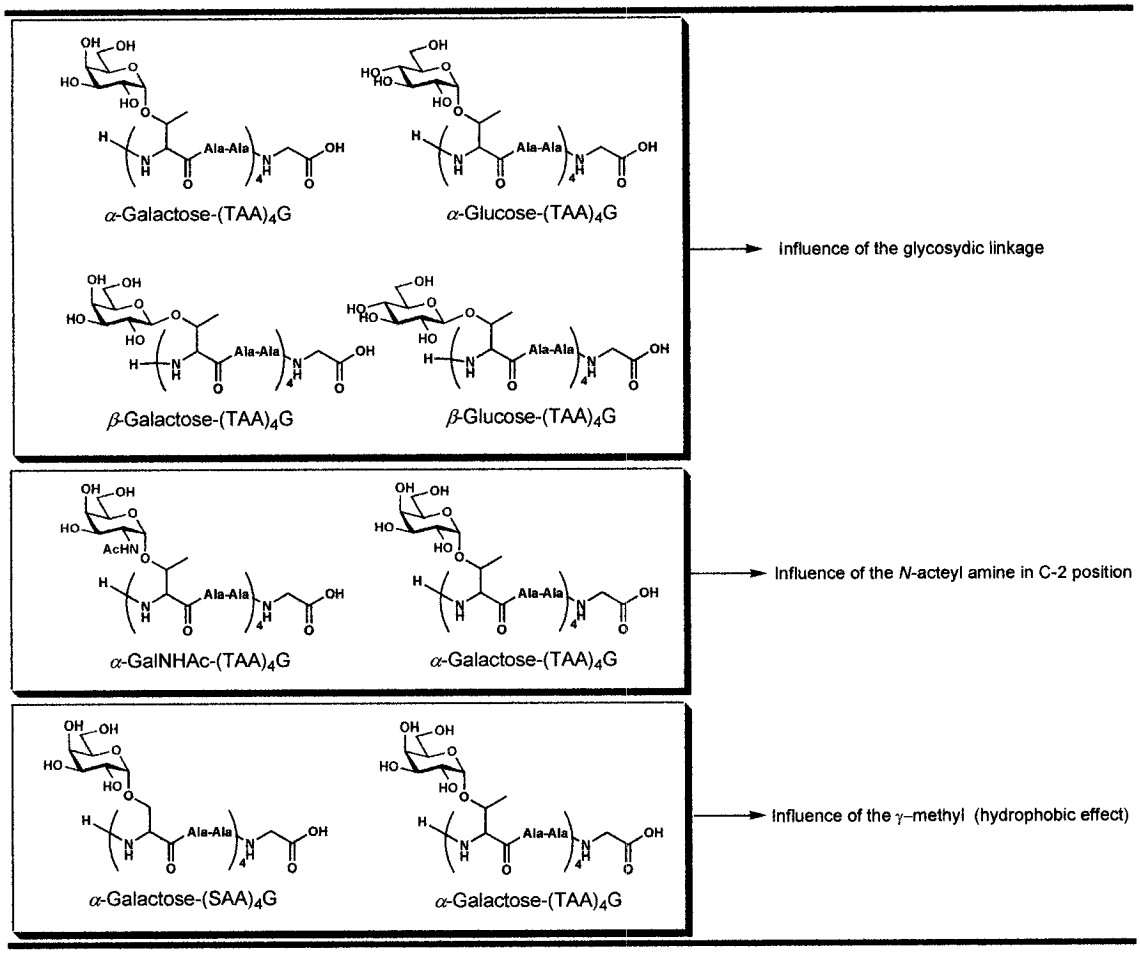


Figure 6: Influence of different structural features on RI and TH activity (continued)

The *C*-[SGG(Gal)]$_4$ analogue was previously prepared and tested in our laboratory.[30] It exhibited interesting RI activity without any TH or dynamic ice shaping. By synthesizing α-Galactose(SGG)$_4$G and compare it to its *C*-linked counterpart *C*-[SGG(Gal)]$_4$, we could obtain precious information on the influence the carbohydrate exo-cyclic oxygen (Figure 8). The same α-Galactose(SGG)$_4$G analogues compared to α-Galactose(SAA)$_4$G could also give us interesting information about the influence of the hydrophobic methyl group of the alanine on conformation and antifreeze activity (Figure 7 and Figure 8).

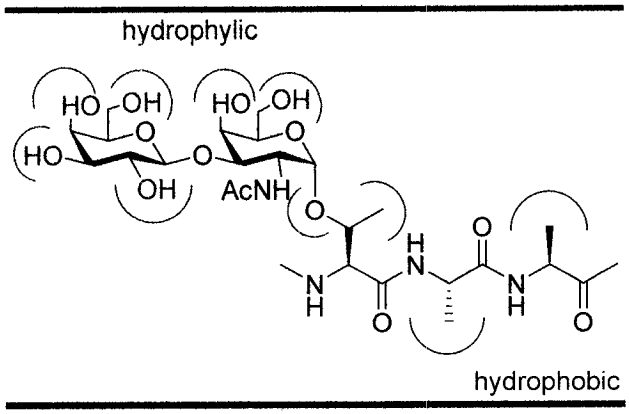


Figure 7: The amphiphilic nature of AFGPs

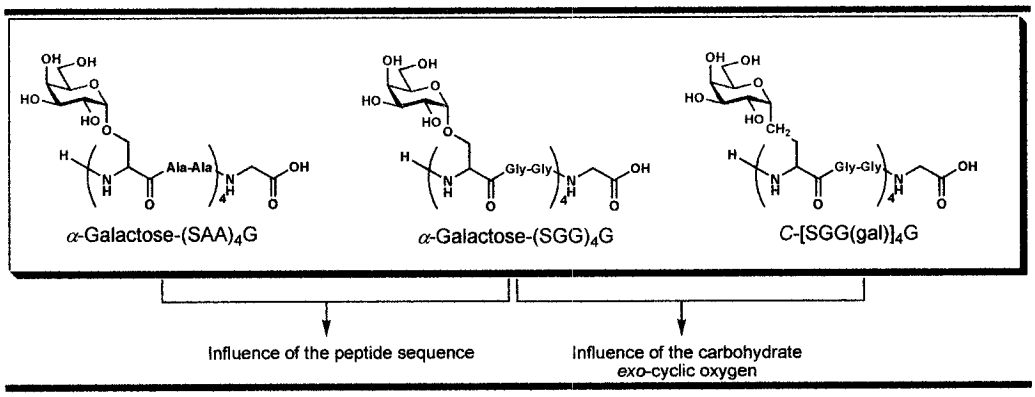


Figure 8: Influence of different structural features on RI and TH activity (continued)

2.3 Synthesis and SAR study of *C*-linked analogues of AFGP8

2.3.1 Synthetic strategies for the preparations of C-linked glycoconjugates.

Since the early 1970's many syntheses of *C*-linked carbohydrate derivatives have been reported.[17,31] Some of them provide useful synthetic handles *en route* to *C*-linked glycoamino acid. One such approach described by Bednarski *et al.*[32] incorporated a β-*C*-linked serine glycoside in the automated solid phase synthesis of a peptide. Their strategy allowed the preparation of *C*-linked amino acid derivatives as a mixture of isomers at the α-amino acid center. The key step in preparation of the *C*-linked analogues was a coupling with β-*C*-linked galactosyl aldehyde using a chiral Wittig reagent. Another approach to produce *C*-linked serine relies on catalytic asymmetric hydrogenation. Toone *et al.*[33] prepared a *C*-glycosyl enamide precursor for asymmetric hydrogenation using Horner-

Emmons olefination. The key step used a chiral rhodium catalyst and successfully produces both the D and L-amino acid diastereoisomer in high diastereomeric excess.

More recently several alternative approaches were reported, for instance Arya *et al.*[34] featured asymmetric enolate methodology to prepare α and β *C*-linked derivatives. Roy *et al.*[35] demonstrated that olefin-metathesis methodology[36,37] was compatible with the synthesis of a glycosyl amino acid,[38] and the next year Marra *et al.*[39] described the successful synthesis of different *C*-linked glycoconjugates using second generation Grubbs catalyst.

Finally in 2002, Nolen *et al.* described a straightforward stereoselective synthesis of the *C*-linked D-gluco- and D-galactopyranosyl L-serines in their β and α forms.[40] This synthesis required the conversion of the allyl *C*-glycopyranosides into their iodoethyl derivatives, which then underwent substitution with the Williams' chiral glycine enolate equivalent.

2.3.2 C-linked analogues of AFGP8 and SAR study

AFGP molecules are composed of three different units: the peptide backbone unit, the disaccharide moiety and the side chain linker, which connects the two previous parts (Figure 9).

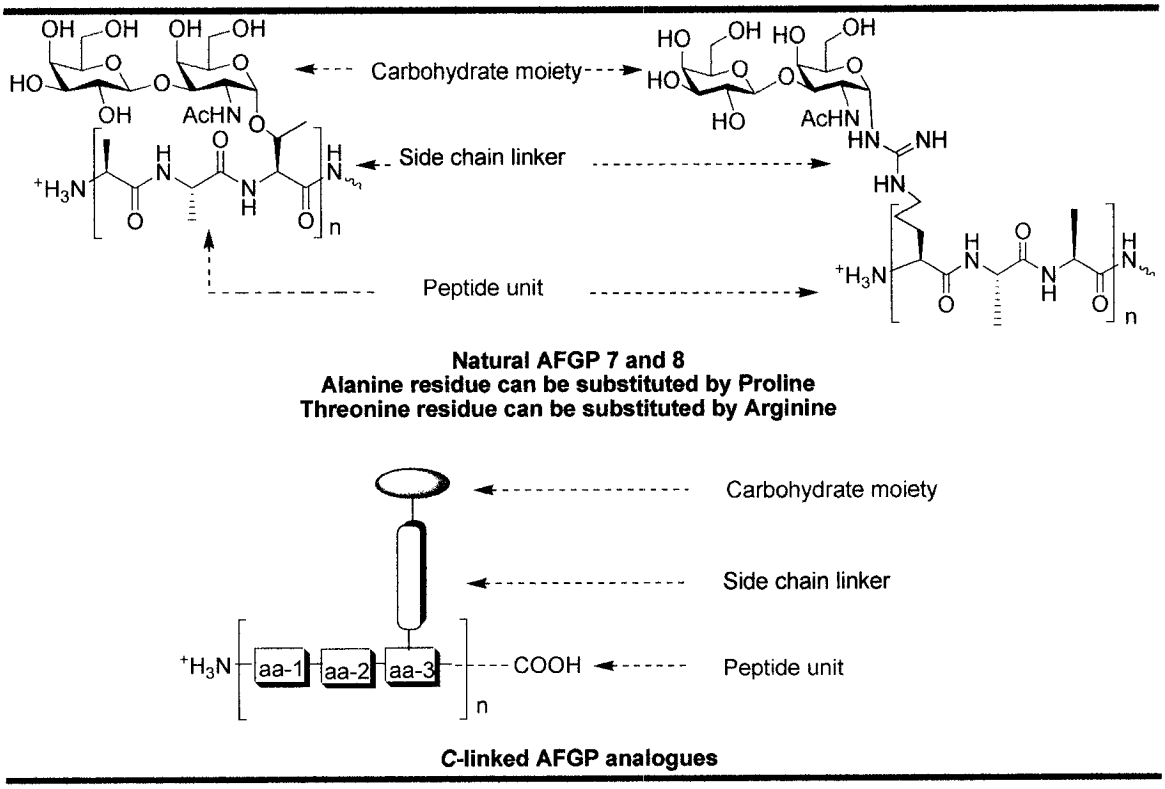


Figure 9: Structure of AFGP8 analogues.

All of these regions possess naturally occurring modifications that were described in Chapter 1. Our linear solid phase strategy allows large variations in the different units (carbohydrate moiety, side chain linker or peptide unit), and is therefore highly amenable for the synthesis of structurally diverse AFGP analogues.

A general synthesis has been developed and a serie of *C*-linked AFGP analogues has been already successfully prepared in our laboratory.[41,42] In order to obtain this first generation of glycopeptides, we decided to combine the following two assumptions. First, threonine (typical linker in the native AFGPs) is occasionally substituted by arginine in AFGP 7 and 8. Figure 10 shows what the native system would look like if this arginine were glycosylated. Second, as mentioned earlier, chemical ligation using an amide bond to form a *C*-linked glycopeptide is a well controlled and established technique. Using this information we decided to prepare our first generation of AFGP8 analogues. As shown in Figure 10, the *C*-linked amide linkage using ornithine or lysine would be a good mimic of the guanidinium moiety found in the arginine present in the native system. Moreover the use of either the convergent or the linear solid phase synthesis in order to obtain our analogues allows us to

easily modify the carbohydrate moiety (galactose, glucose and mannose) as well as to modify the polypeptide backbone.

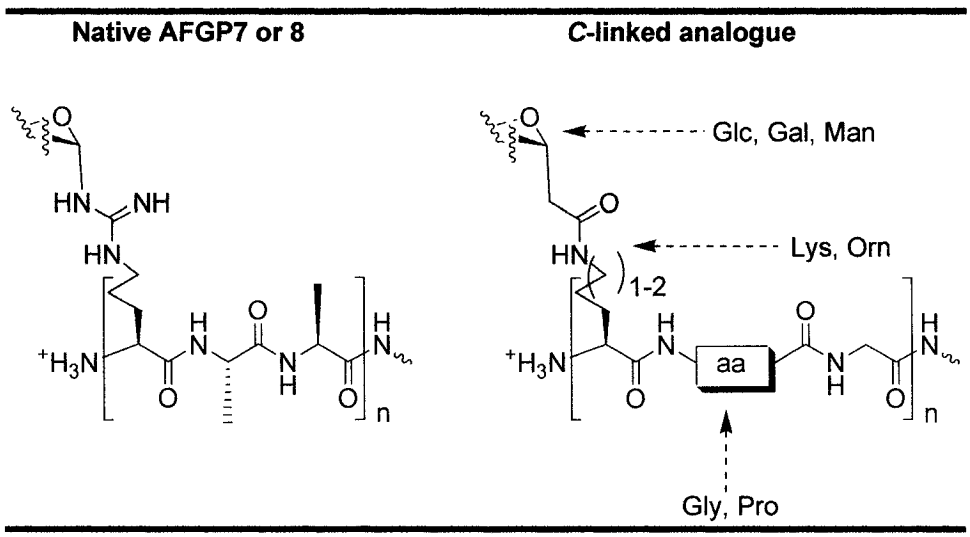


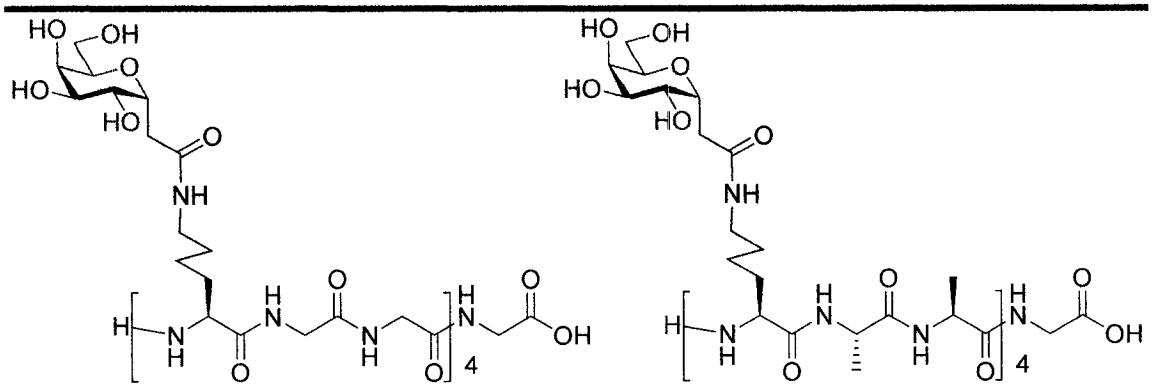
Figure 10: Concept of our C-linked analogues[41]

2.3.3 Goals and objectives

- Development of an efficient synthesis of the *C*-linked *N*-acetyl galactosamine AFGP analogues [OGG(GalNAc)]$_4$ and [OAA(GalNAc)]$_4$ (Figure 11).
- Creation of an efficient synthesis of the *C*-linked disaccharide (β-*C*-Gal(1→3)-α-GalNAc) amino acid.
- Synthesize and study [(LGG(Gal)]$_4$.
    - Determination of the length of the ideal polymer by comparing the RI and TH activity of the [(LGG(Gal)] family of analogue (3-mer, 4-mer, 6-mer and 9-mer)
    - Investigation of the favorable peptide side chain length for RI activity, through comparison of RI activity of the glycosylated ornitine [OGG(Gal)]$_4$ and lysine analogue [LGG(Gal)]$_4$.
    - Determination of whether or not the substitution of the proline in the peptide backbone [LPG(Gal)]$_4$ will influence the RI and TH activities

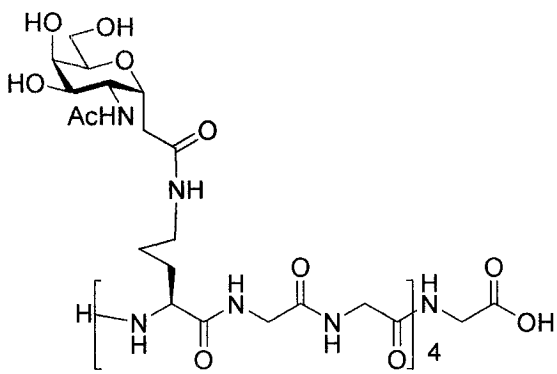
of our *C*-linked analogues $[LGG(Gal)]_4$. Proline was chosen because of its natural occurrence in the native AFGP8.

- Examination of the role of the *N*-acetyl function in the C-2 glycosidic position. Compared to $[OGG(Gal)]_4$ analogue, the $[OGG(GalNAc)]_4$ analogue should give us information about the influence on TH and RI activity of this *N*-acetyl function on our *C*-linked analogues.
- Examination of other substitutions in primary amino acid sequence. The peptide backbone of the native system is mainly composed of alanine, whereas our *C*-linked analogues contain glycine. Investigating the influence of alanine substitution on our *C*-linked analogues might explain the role of the alanine and its influence on RI and TH activities in our *C*-linked analogues. This can be obtained by synthesizing $[LAA(Gal)]_4$ and $[OAA(GalNAc)]_4$ and comparing them to their glycine counterparts $[LGG(Gal)]_4$ and $[OGG(GalNAc)]_4$.
- Analyzis of the influence of conformation of these analogues on antifreeze activity using circular dichroism

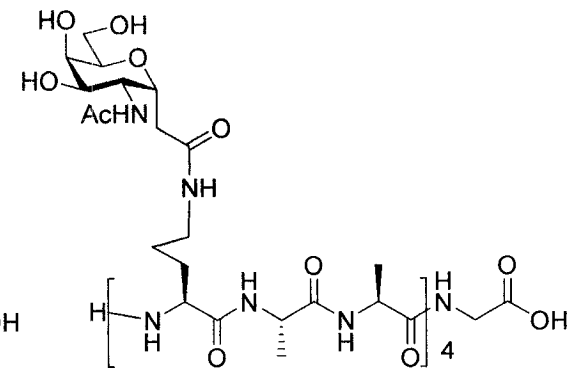


[LGG(Gal)]$_4$

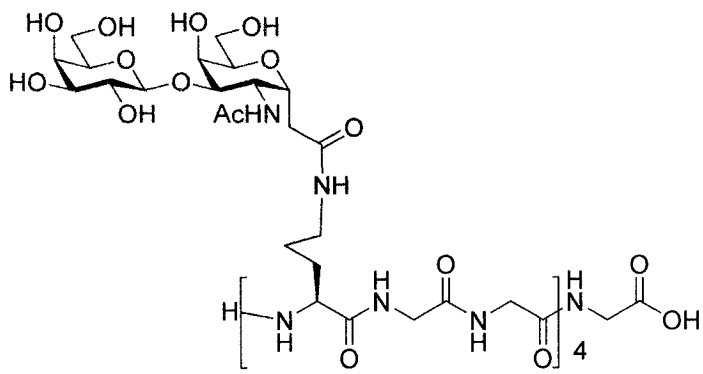
[LAA(Gal)]$_4$



[OGG(GalNAc)]$_4$



[OAA(GalNAc)]$_4$



[OGG(β-Gal(1-3)GalNAc)]$_4$

Figure 11: C-linked targets for our SAR study

[1] Rubinsky, B.; Arav, A.; Devries, A.L. *Cryobiology* **1992**, 29, 69.

[2] 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.

[3] Davies, P. L.; Sykes, B. D. *Curr. Opin. Struct. Biol.* **1997**, 7, 828.

[4] Fletcher, G. L.; Hew, C. L.; Davies, P. L. *Annu. Rev. Physiol.* **1999**, 63, 359.

[5] Wilson, P. W. *Cryo-Letters* **1993**, 14, 31.

[6] Knight, C. A. *Nature* **2000**, 406, 249.

[7] Filira, F.; Scolaro, B. B. ; Foffani, M. T.; Mammi, S. ; Peggion, R. ; Rocchi, R. *Int. J. Of Biol. Macromol.* **1990**, 12, 41.

[8] Franks F.; Morris E. R. *Biochem. Biophys. Acta.* **1978**, 540, 346.

[9] Bush C. A.; Feeney R. E.; Osuga D. S. T.; Talapati S.; Yeh Y. *J. Peptide Protein Res.* **1981** 17, 125.

[10] Bush C. A.; Feeney R. E. *Int. J. Peptide Protein Res.* **1986**, 28, 386.

[11] Rao, B. N.; Bush,C. A. *Biopolymers* **1987**, 26, 1227.

[12] Tachibana, Y.; Matsubara, N.; Nakajima, F.; Tsuda, T.; Tsude, S.; Monde, K.; Nishimura, S.-I. *Tetrahedron* **2002**, 58, 10213.

[13] Lane, AN. N. ; Hays, L. M.; Tsvetkova, N.; Feeney, R. E.; Crowe, L. M.; Crowe, J. H. *Biophys. J.* **2000**, 78, 3195.

[14] Sarno, D.; Murphy, A. V.; DiVirgilio, E. S.; Jones, W. E., Jr.; Ben, R. N. *Langmuir* **2003**, 19, 4740.

[15] (a) Shulman, M. L.; Shiyan,S. D. and Khorlin, A. Y. *Carbohydr. Res.* **1974**, 33, 229. (b) Chmielewski, M.; Bemiller, J. N. and Ceretti, D. P. *Carbohydr. Res.* **1981**, 97, C1-C4. (c) Myers, R. W. and Lee, Y. C. *Carbohydr. Res.* **1986**, 152, 143.

(d) Bemiller, J. N.; Yadav, M. P.; Kalabokis, V. N. and Myers, R. W. *Carbohydr. Res.* **1990**, 200, 111.

(e) Kuan, S. F.; Byrd, J. C.; Basbaum, C.; Kim, Y. S. *J. Biol. Chem.* **1989**, 264, 19271.

(f) Byrd, J. C.; Dahiya, R.; Huang, J.; Kim, Y. S. *Eur. J. Cancer* **1995**, 31A, 1498.

(g) Hennebicq-Reig, S.; Lesuffleur, T.; Capon, C.; De Bolos, C.; Kim, I.; Moreau, O.; Richet, C.; Hemon, B.; Recchi, M. A.; Maes, E.; Aubert, J. P.; Real, F. X.; Zweibaum, A.; Delannoy, P.; Degand, P.; Huet, G. *Biochem. J.* **1998**, 334, 283.

(h) Zanetta, J. P.; Gouyer, V.; Maes, E.; Pons, A.; Hemon, B.; Zweibaum, A.; Delannoy, P.;Huet, G. *Glycobiology* **2000**, 10, 565.

[16] (a) Lee, Y. C.; Lee, R. T. *Acc. Chem. Res.* **1995**, 28, 321.

(b) Dwek, R. A. *Chem. Rev.* **1996**, *96*, 683.

(c) Imperiali, B.; Shannon, K. L.; Rickert, K. W. *J. Am. Chem. Soc.* **1992**, 114, 7942.

(d) Imperiali, B. *Acc. Chem. Res.* **1997**, 30, 452. (e) Seitz, O. *ChemBioChem* **2000**, 1, 214.

[17] (a) Postema, M. H. D. *Tetrahedron* **1992**, 48, 8545.

(b) Postema, M. *C-Glycoside Synthesis*; CRC Press: Boca Raton, **1995**.

(c) Levy, D.; Tang, C. *The Chemistry of C-Glycosides*; Pergamon: Oxford, **1995**.

(d) Togo, H.; He, W.; Waki, Y.; Yokoyama, M. *Synlett* **1998**, 700.

(e) Skrydstrup, T.; Vauzeilles, B.; Beau, J.-M. *In Carbohydrates in Chemistry and Biology. The Chemistry of Saccharides*; Ernst, B., Hart, G. W., Sinay¨, P., Eds.; Wiley-VCH: New York, **2000**; Vol. 1, Chapter 20, pp 495-530.

[18] Seeberger, P. H.; Danishefsky, S. J. *Acc. Chem. Res.* **1998**, 31, 685.

[19] Debeer, T.; Vliegenthart, J. F. G.; Löffler, A. Hofsteenge, J. *Biochemistry* **1995**, 34, 11785.

[20] Marcaurelle L.A.; Bertozzi, C.R. *Glycobiology* **2002**, 12, 69R.

[21] Mathieux, N.; Paulsen, H.; Meldal, M.; Bock, K. *J. Chem. Soc. Perkin Trans.* **1997**, 1, 2359.

[22] Meinjohanns, E.; Meldal, M.; Schleyer, A.; Paulsen, H.; Bock, K. *J. Chem. Soc. Perkin Trans.* **1996**, 1, 985.

[23] Lemieux, G.A.; Bertozzi C.R. *Trends Biotechnol.* **1998**, 16, 506.

[24] Marcaurelle L.A.; Bertozzi, C.R. *J. Am. Chem. Soc.* **2001**, 123, 1587.

[25] Muir, T.W. *Synlett*, **2001** 6, 733.

[26] Lepannen, A.; White S.P.; Helin, J.; McEver, R. P.; Cummings, R.D. *J. Biol. Chem.* **2000**, 275, 39569.

[27] Komatsu, S. K.; DeVries, A. L.; Feeney, R. E. *J. Biol. Chem.* **1970**, 245, 2909.

[28] St. Hilaire, P.M.; Lowary, T.L.; Meldal, M.; Bock, K. *J. Am. Chem. Soc.* **1998**, 120, 13312.

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

[30] Liu S. Ph.D. thesis, University of Ottawa, Ottawa, **2006**.

[31] Du, Y.; Linhard, R.J.; Valhov, I.R. *Tetrahedron*, **1998**, 54, 9913.

San Martin, R.; Tavassoli, B.; Walsh, K. E.; Walter, D. S. and Gallagher, T. *Org. Lett.* **2000**, 2, 4051.

Grant, L.; Liu, Y.; Walsh, K. E.; Walter, D. S. and Gallagher T. *Org. Lett.* **2002**, 4, 4623. For a recent review on "Synthetic methods of amino *C*-Glycosides" see Xie, J. *Recent Res. Devel. Organic Chem.* **1999**, 3, 505.

[32] Bertozzi, R.; Hoeprich, Jr., P.D.; Bernaski, M.D. *J. Org. Chem.* **1992**, 57, 6092.

[33] Debenbaum, S.D.; Debenbaum, J.S.; Burk, M.J.; Toone, E.J. *J. Am. Chem. Soc.* **1997**, 119, 9897.

[34] Ben, R.N.; Orellana, A.; Arya, P.J. *J. Org. Chem.* **1998**, 63, 4817.

[35] Dominique, R.; Liu, B.C.; Das, S.K.; Roy, R. *Synthesis* **2000**, 862.

[36] Roy, R.; Das, S.K. *Chem. Commun.* **2000**, 519

[37] Jorgensen, M.; Hadwiger, P.; Madsen, R.; Stutz, A.E.; Wrodnigg, T.M. *Curr. Org. Chem.* **2000**, 4, 565.

[38] Mc Garvey, G.J.; Benedum, T.E.; Schmidtmann, F.W. *Org. Lett.* **2002**, 4, 3591.

[39] Dondoni, A.; Giovannini, P.P.; Marra, A. *J. Chem. Soc. Perkin Trans.*, **2001**, 1, 2380.

[40] Nolen, E.G.; Watts, M.M.; Fowler, D.J. *Org. Lett.* **2002**, 4, 3963.

[41] Ben, R. N.; Eniade, A.; Hauer, L. *Org. Lett.* **1996**, 1, 1759.

[42] Eniade, A.; Ben, R.N. *Biomacromolecules* **2001**, 2, 557.

Chapter 3

AGGREGATION OF ANTIFREEZE GLYCOPROTEIN FRACTION 8 AND ITS EFFECT ON ANTIFREEZE ACTIVITY

Bouvet, V. R.; Lorello, G.; Ben, R. N. " *Biomacromolecules* **2006,** 7, 565.

3 AGGREGATION OF ANTIFREEZE GLYCOPROTEIN FRACTION 8 AND ITS EFFECT ON ANTIFREEZE ACTIVITY

3.1 Previous conformational studies of AFGPs

While structure-function studies have implicated specific structural motifs as essential for antifreeze activity in AFGPs, the relationship between solution conformation and antifreeze activity is poorly understood. As described in a previous chapter, the mechanism of action by which these compounds function is an adsorption-inhibition process where AFGPs bind irreversibly to the surface of an ice crystal and elicit a localized freezing point depression.[1] The structural moieties crucial for the AFGP-ice interaction remain a source of debate amongst experts in the field.[2,3,4] Recent structure-function studies have implicated the disaccharide, the β-methyl group of threonine and the C-2 acetamide moieties are essential for potent antifreeze activity.[5]

The relationship between AFGP activity and solution conformation has been previously examined by a number of research groups in a concerted effort to understand how AFGPs elicit a localized freezing point depression.[6,7,8,9] Towards this end, a number of complementary techniques have been employed and have generated somewhat contradictory results. Early circular dichroism (CD) studies suggest AFGPs adopt a random coil conformation in solution,[7,10,11] while vacuum ultraviolet CD experiments imply a three-fold left-handed helix[12] (but the data is not inconsistent with a random coil structure). Quasi-elastic light scattering (QELS or DLS) studies suggest an extended coil is the predominant conformation,[13] but Raman spectroscopy implies a random coil with significant α-helix and β-conformations.[14] Analysis of the solution conformation of AFGP1-4 and AFGP8 by ^{1}H NMR supports the existence of a three-fold left-handed helix[8,15] previously observed by vacuum UV circular dichroism. Early ^{13}C NMR experiments propose a random coil structure[16] as well as a flexible coil structure.[17] In lieu of these conflicting results, a more distinct picture has recently emerged through the use of high field NMR, IR spectroscopy, and molecular dynamics simulations.[18,19] These studies suggest that AFGP1-5 are a dynamically disordered molecules showing neither long nor short range order. In contrast, AFGP8 appears to possess no long-range order but does possess local order consistent with a

polyproline II structure. While these studies suggest that the solution conformation of AFGP is very complicated, the relationship between activity and conformation is still not understood and additional studies are required. Our laboratory have previously reported differences in the adsorption of AFGP8 onto both hydrophobic (Highly Oriented Pyrolytic Graphite: HOPG) (Figure 1) and hydrophilic (mica) surfaces (Figure 2).[20] In these studies AFGP8 (2.6 kDa) that was isolated from the cold ocean Teleost species *Gadus ogac* (Rock Cod) was supplied from A/F Protein Inc. The lyophilized glycoprotein was 95% pure (HPLC) with small amounts of AFGP7 and other naturally occurring plasma proteins present.[21]

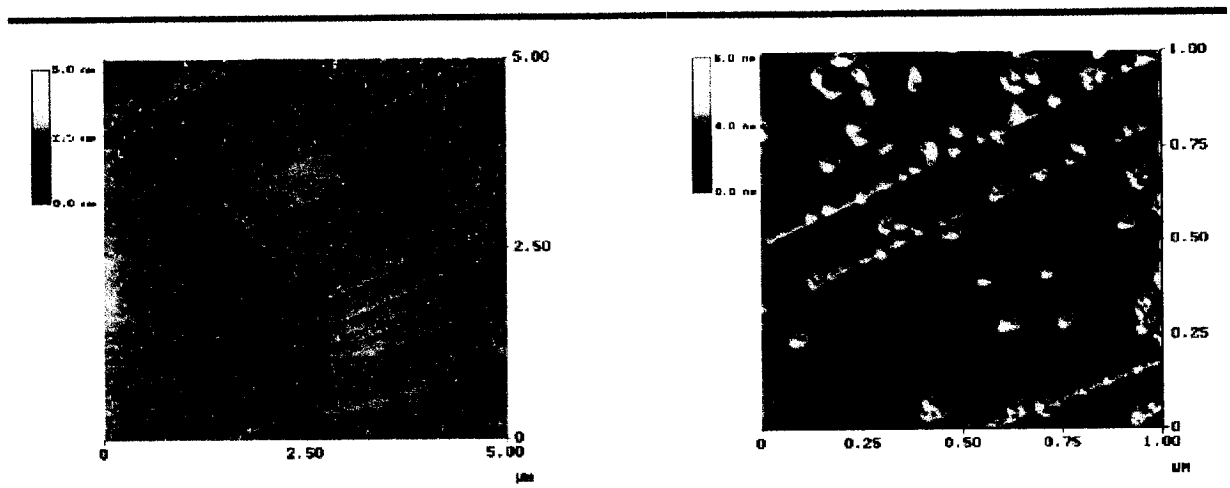


Figure 1: AFM images of AFGP8 deposited from aqueous solutions *(1.0 x 10-5g/ml) on a freshly cleaved HOPG surface: (a) scan rate, 1.20 Hz (b) scan rate 1.65 Hz*[20]

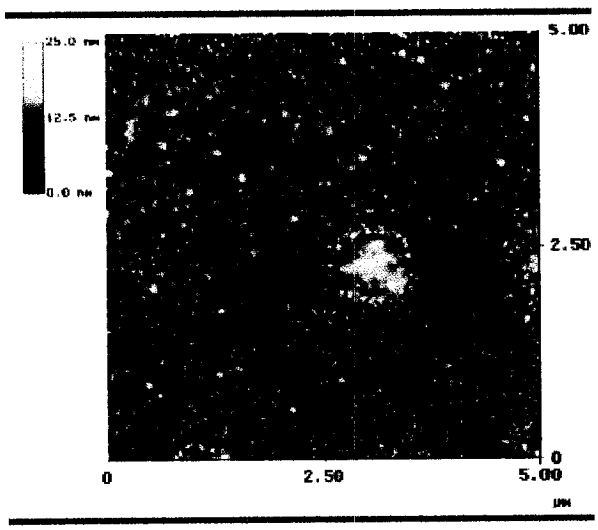


Figure 2: AFM images of AFGP8 deposited from aqueous solutions *(1.0 x 10-5g/ml) on a freshly cleaved mica surface: scan rate 1.80 Hz*[20]

Figure 1 and Figure 2, show the adsorption of AFGP8 on a hydrophobic surface (HOPG) and a hydrophilic surface (mica). It appears that while AFGP8 may be regarded as an amphipathic glycoprotein (both hydrophilic and hydrophobic), our studies indicate that when presented with both hydrophilic and hydrophobic binding sites (i.e. step edges or planes of HOPG) AFGP8 binds preferentially to the hydrophilic step edges (AFGP8 particles are oriented in straight line on HOPG and these lines correspond to hydrophilic oxidized nanometer-size step edges). It was also observed that the dispersion on both surfaces is homogeneous and that the adsorption of AFGP8 onto HOPG and mica results in homogeneous particle sizes. This suggested that unlike AFGP1-5, AFGP8 has the ability to form aggregates in solution prior to adsorption onto both hydrophilic and hydrophobic surfaces. The particles at this time were found to be approximately 40 times the size of a single AFGP8 molecule suggesting the formation of aggregates.

While the issue of glycoprotein aggregation is itself very interesting, the implications of AFGP8 aggregation and its effects on solution conformation and antifreeze activity have not been studied. Since atomic force microscopy is not a suitable technique to explore the aggregation of AFGP, we sought to examine this phenomenon using dynamic light scattering (DLS) and circular dichroism (CD), two techniques that have a time scale in between that of IR and NMR.

3.2 Dynamic light scattering experiments (DLS)

3.2.1 Introduction to Dynamic Light Scattering[22,23]

Dynamic light scattering, by recording how light scatters off of particles in suspension, has the potential to give different information about particles in solution, including their size. Brownian motion is a phenomenon that is fundamental to this experiment. It describes the way in which very small particles move in fluid suspensions, where the size of the fluid is much smaller than the suspended particles (solubilized molecules). The direction and the magnitude of the motion of these suspended molecules fluctuate with time and is directly correlated to the random collisions between the molecules. When a laser beam is shone through a solution with suspended particles, the beam scatters off of those particles in all directions, resulting in a scattering-angle-dependent intensity

pattern which is directly related to their Brownian motion. Measuring the intensity fluctuations at a given scattering angle can yield amongst other information to the "hydrodynamic radius" of the suspended particles.

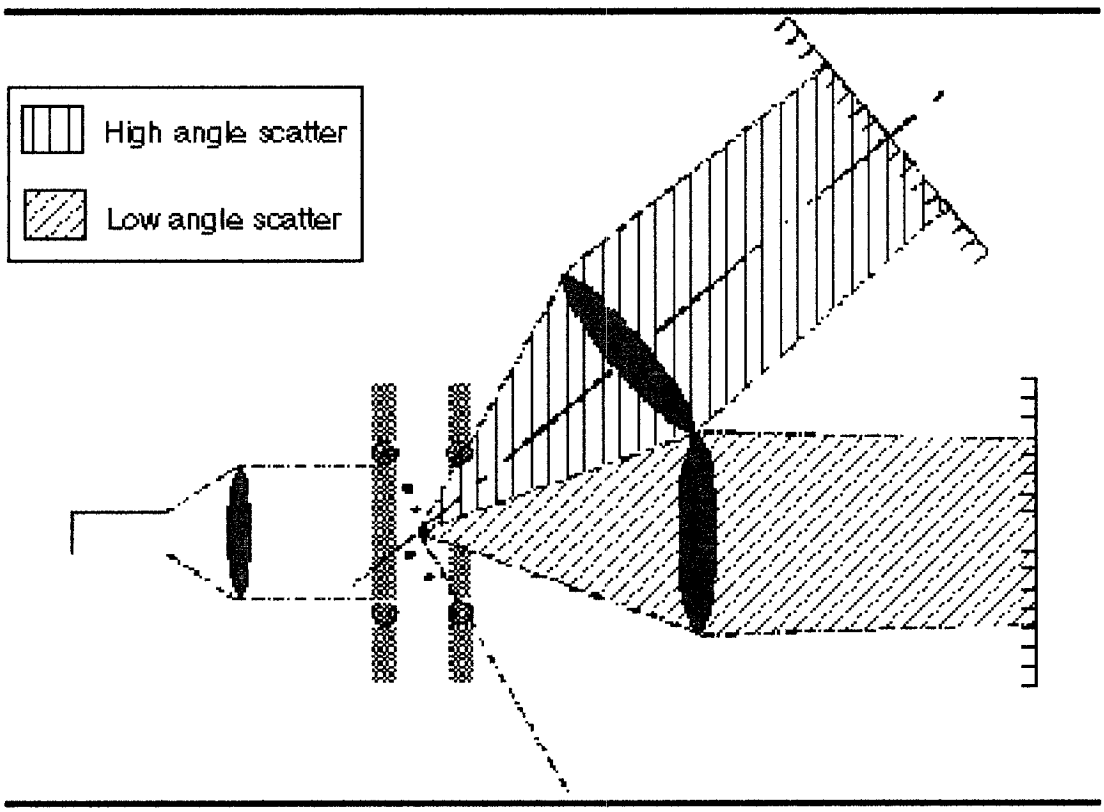


Figure 3: Principles of dynamic light scattering (DLS).[23] *The Laser Diffraction method of measuring particle size takes advantage of an optical principle, which dictates that small particles in the path of a light beam scatter the light in characteristic, symmetrical pattern. Given a certain pattern of scattered light intensity as a function of angle to the axis of the incident beam, the distribution of particle sizes can be deduced.*

The hydrodynamic radius of a particle (the effective radius of an irregularly shaped particle) is used when describing the manner in which particles in suspension diffuse through the suspending medium. The variation of the fluctuation of the intensity of scattered light allows us to use random statistical methods (correlation) to analyze this scattering pattern. If the intensity at a given scattering angle is recorded over a small sample time, the fluctuation of the intensity arising from Brownian motion can be also expected to be small. If two such recordings are made from the same scattering angle simultaneously, the two samples can be

compared with one another through cross-correlation. A measure of how quickly the scattered light intensity changes with time can be obtained.

A correlation function (R_I) can then be obtained and corresponded to a time-averaged function of products of intensities of the signals subjected to various temporal shifts (Δt). I_1 corresponds to the scattering intensity after the first laser impulsion and I_2 after the second.

$$R_I(\Delta t) = \{I_1(t).\ I_2(t+\Delta t)\}$$

Cross-correlation, when applied to a physical system, yields a time-difference dependent intensity expression of the form

$$R_I(\Delta t) = A + Be^{-\Gamma \Delta t}$$

where R_I is the square intensity function given by cross-correlation. If we consider the inverse Fourier transform of the correlation function; in this case a Lorentzian, which is known as the power spectrum:

$$\mathbf{F}^{-1}\left(R_I(\Delta t)\right) = S(\Delta\omega) = \frac{1}{\pi}\frac{\left(\Gamma/2\right)^2}{(\Delta\omega-\omega_0)^2+\left(\Gamma/2\right)^2}$$

In the case of the power spectrum, the decay constant (Γ) is described by

$$\Gamma = 2K^2D$$

where K is a constant related to the frequency of the scattering light (λ), scattering angle (θ) and index of refraction (n), and D is the diffusion constant unique to the scattering solution.

$$K = \frac{2\pi n}{\lambda_0}\sin\left(\frac{\theta}{2}\right)$$

$$D = \frac{K_B T}{3\pi\eta d}$$

The diffusion constant D is derived from the Stokes-Einstein relation, and is dependent upon temperature (T), viscosity (η), and scattering particle diameter (d). When S is plotted with the respect to frequency, the decay constant can be obtained, and from there we can simply derive the diameter of the scattering particle.

3.2.2 Dynamic light scattering experiments (DLS) with AFGP8

Dynamic light scattering experiments were performed on a DynaPro dynamic light scattering apparatus in distilled water at concentrations ranging from 0.4 to 40 mM. The correlation curves were interpreted using the Dynamics Software v.6.1.08. This software can be used to fit correlation curves exhibiting multiple exponential decays (i.e. solutions that are multimodal in nature). All the discontinuous correlation curves as well as correlation curves exhibiting outlier sum of squares analysis (SOS) due to aberrant effects (dust, etc.) were excluded. Solution concentrations greater than 40 mM were not examined due to the high viscosity and the inherent error associated with estimating molecular weights in such solutions. All calculations were effectuated using both rod and globular shape approximations. Both approximations gave similar trends with minimal variations. The results presented herein are the rod shape approximations which are more adequate for AFGP8 molecules. Initial DLS experiments revealed the presence of a high molecular weight substance exceeding 10,000 kDa (Figure 4).

Despite careful efforts to ensure the samples were dust free, the high molecular weight substance was attributed to particulates. This was problematic in that a PMT saturation effect was observed when 100% of the laser power was used. However, by employing a 65% reduction in laser power the samples could be analyzed. Under these conditions an 8 kDa aggregate that was 1% of the signal intensity but corresponding to 99.9% of the sample mass was sporadically observed at higher AFGP8 concentrations. Given that the scattering of light in DLS is exponentially proportional to the size of the aggregate in solution, the detection of these small aggregates in the presence of particles in excess of 10,000 kDa was difficult using reduced laser power. However, filtration with a 100 kDa nanopore filter removed the particulates and permitted a higher laser power to be

used. Under these conditions, analysis of the 20 and 40 mM solutions (Figure 5) revealed the presence of aggregates consistent in size with at least two individual AFGP8 molecules.[24]

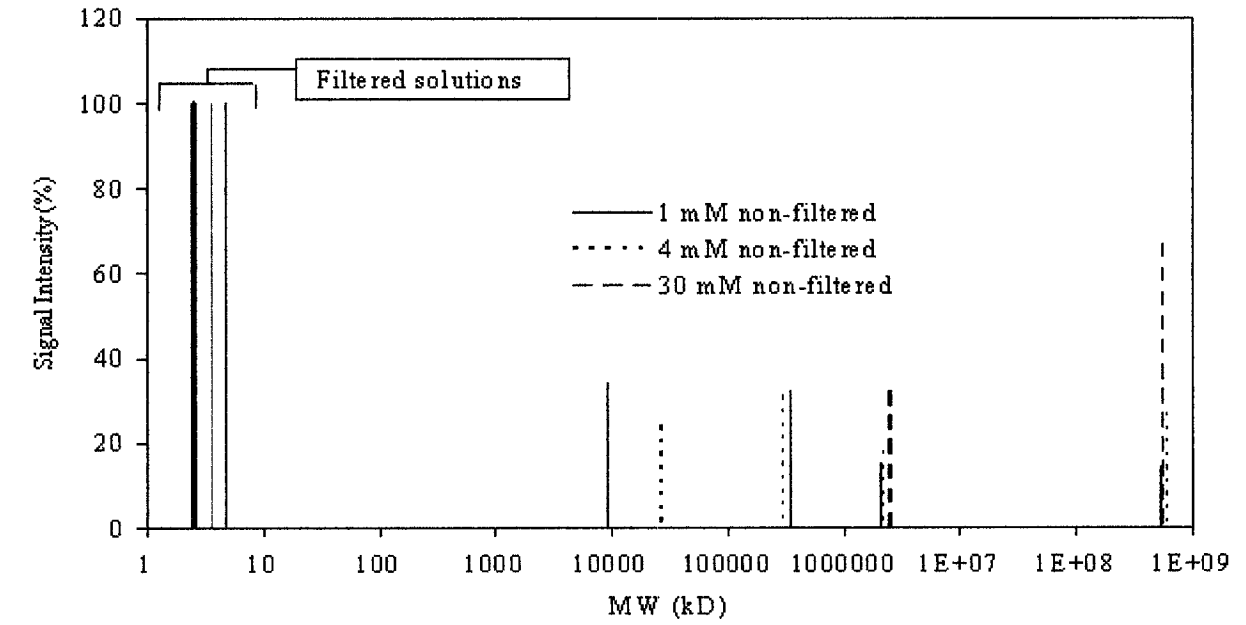


Figure 4: Dynamic light scattering (DLS) of unfiltered AFGP8 solutions in double distilled water

Figure 5 describes the average aggregate size for various solutions. Each solution was observed over 20 measurements and showed an overall molecular weight increase as a function of solution concentration. In contrast, 0.4 to 4 mM solutions of AFGP8 showed no evidence of aggregation. We investigated this aggregation process as a function of time and solution concentration (Figure 6 and Figure 8). These experiments confirm the propensity for the aggregates to increase in size when left suspended in solution. For instance, a freshly prepared 1 mM solution of AFGP8 initially showed no evidence of aggregation but analysis of a similar solution allowed to stand for 2 months revealed aggregates consistent with AFGP8 dimers (molecular weight approximately 5.0 kDa).

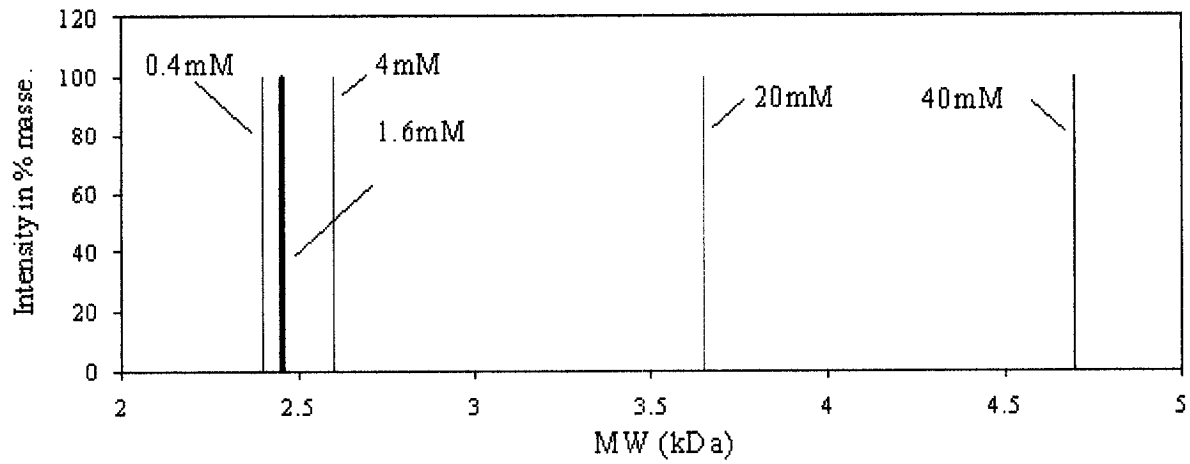


Figure 5: Dynamic light scattering (DLS) of filtered AFGP8 solutions in double distilled water *(each solution was measured 20 times)*

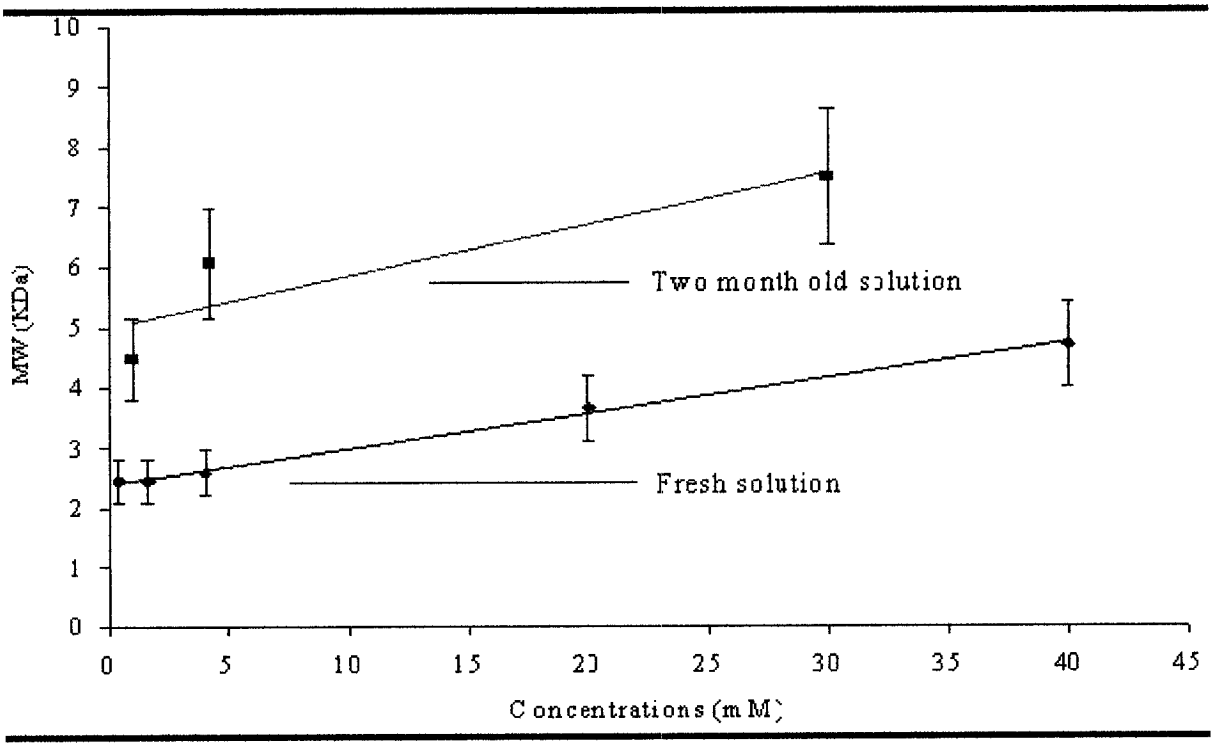


Figure 6: Aggregate size as a function of concentration and time. *All solutions were filtered with a 100 kDa Nanopore filter prior to analysis.*

It was also observed that the initial concentration of the AFGP8 solution had an influence on the size of aggregate formed. For instance, a freshly prepared 40 mM solution contained aggregates of two AFGP molecules (5.0 kDa) while a 1 mM solution contained only monomers. The low polydispersity and the continuity in the correlation curves

associated with a high sum of squares (SOS) value ruled out the presence of dust and suggested that we were observing a multi-modal system with low polydispersities. Figure 7 depicts the multimodality of the aggregates in a 4, 20 and 40 mM solution of AFGP8. This example is the result of 20 measurements without any molecular weight averaging. Each measurement represents a weighted average of eleven acquisitions, and thus twenty measurements were taken for a total of 220 acquisitions. The error bars in Figure 7 represent the mean standard deviation obtained from each measurement. The 40 mM solution of AFGP8 in water produces aggregates consistent with two individual AFGP8 molecules while the 20 mM solution contained predominantly individual AFGP8 molecules.

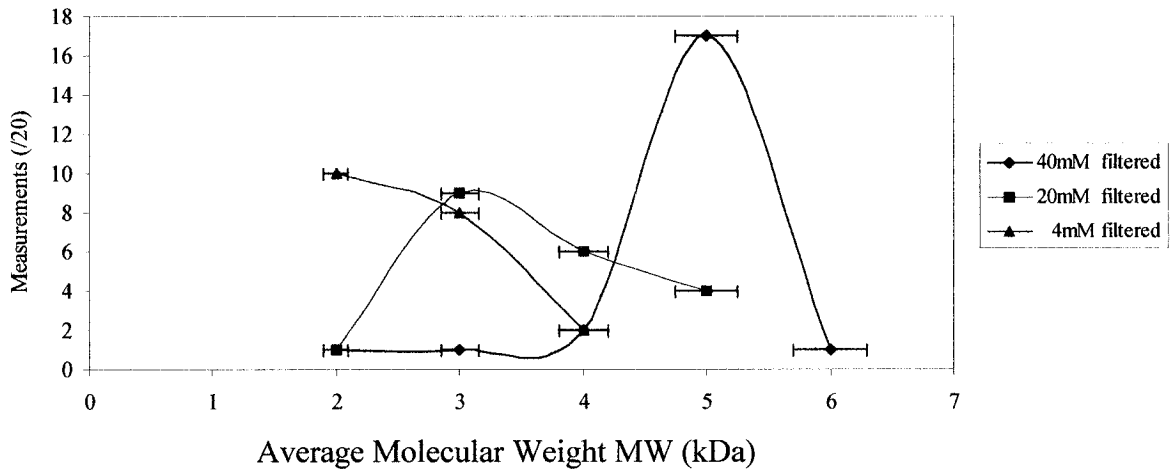


Figure 7: Multimodality of AFGP8 at different concentrations

Figure 8 depicts the multimodality of the aggregates in a 1, 4 and 30 mM solution of AFGP8 standing for 2 months. The 1 and 4 mM solutions produce aggregates consistent with both two and three individual AFGP8 molecules. The 30 mM solution clearly exhibits the presence of aggregates consistent with three and four AFGP8 molecules.

Our results confirm the existence of low molecular weight aggregates in solutions of AFGP8. More specifically, we see evidence of aggregation in fresh AFGP8 solutions that have a concentration of 20 mM or more. In two month old solutions the same trend as for the fresh solutions is observed, in that the more concentrated the solution, the higher the

aggregate molecular weight. If we compare the fresh 4 mM solution with the two month old solution we can conclude that the size of the aggregate is also dependent upon time. The size of the aggregates is different than previously estimated using AFM experiments.[20] The fact that the sizes of the aggregates observed is dependent upon concentration and time, that different size of aggregates are present (dimer, trimer, etc.) and that aggregation has been observed to be reversible suggest that there is a complex aggregation equilibrium dependent upon concentration and time.

Based upon the DLS results, we sought to investigate whether the aggregation significantly influenced solution conformation. In order to achieve this, we examined the solution conformation of AFGP8 as a function of concentration and temperature using circular dichroism (CD) spectroscopy.

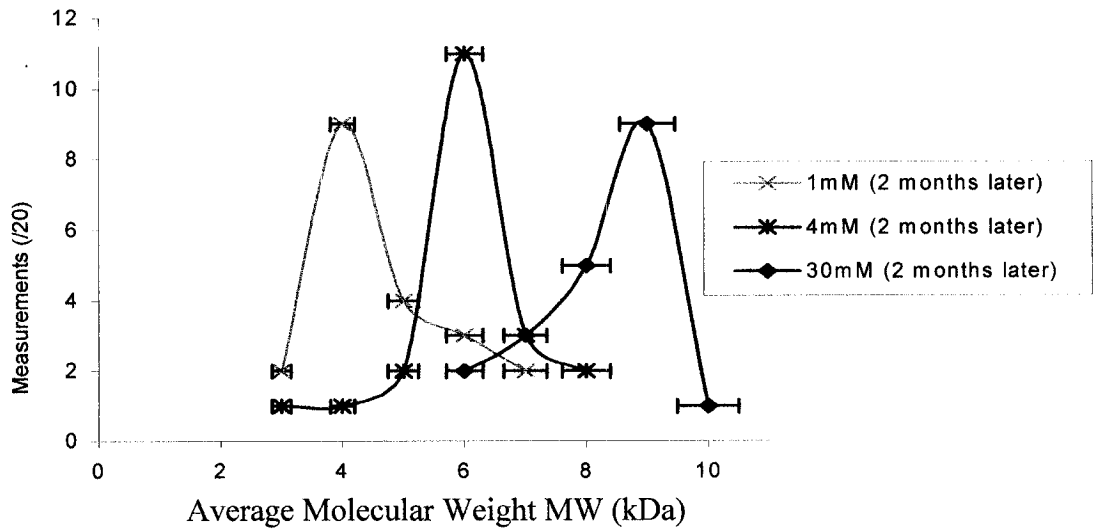


Figure 8: Multimodality of AFGP8 after 2 months

3.3 Circular dichroism experiments (CD)

3.3.1 Introduction to Circular Dichroism

Circular dichroism (CD) spectroscopy is an optical technique that provides information about the chirality of molecular structures.[25] This technique is also used to determine the secondary and tertiary structures of proteins. Circular dichroism measures the difference in absorption of left and right circularly polarized light which arise due to structural asymmetry. Chiral compounds are optically active and therefore a preferential absorption is observed fro either the left or the right circularly polarized light. The variation is directly related to the circular dichroism of the sample at that wavelength. Successive detection is performed at various wavelengths and leads to the generation of the full CD spectrum.

As for any other absorption phenomenon, the variation of CD is linked to the presence of particular chemical groups able to absorb the incident light at selected wavelength (chromophores). Moreover, this chromophore (amide bond, aromatic rings, etc.) must have the special property to absorb differently the left- and right circular polarized light (it has to be non-symmetric in its absorption behaviour). Chiral chromophores as well as chromophores linked to a chiral center or in a chiral environment behave exactly like this. However, there is another source of CD signal: the chiral disposition of chromophores in a 3D structure. For instance α helix has handedness, so its mirror image cannot be superposed on it. If a group of chromophores is disposed in a α helical arrangement a pronounced CD effect characteristic of α helix conformation would be observed. CD studies of protein folding are essentially dictated by the relative orientation and position of the amide linkage.

3.3.2 CD Experiments

All CD measurements were obtained using a Jasco 810 CD spectrometer and samples were dissolved in double distilled water without buffer at pH = 7.4. Deconvolution analysis was performed using CDPro with SELCON3 and IBASIS5. IBASIS5 was used as the set of reference proteins containing 37 proteins with α-helix, β-structure, polyproline II and

unordered conformations with optimal wavelength 185-240 nm.[26] This method has been shown to accurately estimate the secondary structure of a large number of proteins with an assorted secondary structure.[26] So far no set of reference containing glycoproteins are available for deconvolution analyses. Therefore, knowing that carbohydrates do not absorb between 190 and 300 nm and that previous conformational analyses suggest that AFGP8 folds in both random coil and polyproline type II conformations, IBASIS5 was the most reasonable database available.

At 21°C, 0.01 to 0.05 mM solutions of AFGP8 produced spectra consistent with random coil structure (Figure 9). Deconvolution of these spectra indicated that the major conformation present in solution was random coil (41%) with 13% α-helix, 20% β-sheet and smaller percentages of β-turn and polyproline type II helix.

At concentrations of 0.1 and 5.0 mM, the CD spectrum of AFGP8 changed dramatically and exhibited a large positive Cotton Effect at 195 nm. Estimation of the secondary structure using deconvolution software suggested this was the result of a 17% increase in α-helix content in the glycopeptide. The relative amount of random coil structure remained constant at 41% while the relative contribution for β-sheet conformation was reduced by 8%. A 10 mM solution of AFGP8 was approximately 85% random coil. This dynamic "mixture" of conformations was also reflected in the absence of a well-defined isodichroic point, consistent with a system possessing more than a two state conformational equilibrium.

At higher concentrations (20 to 100 mM), other changes in the CD spectra were evident (Figure 9). Deconvolution of these spectra revealed subtle changes in the local conformational ordering. The relative amount of random coil was between 40 to 48% with contributions from β-sheet structure at 35%. In the 70, 80, 90 and 100 mM AFGP8 solutions, the contribution of random coil secondary structure was approximately equal and β-sheet contributions were still the second most predominant contributor. All of these spectra, including those in Figure 9 reflected a small degree of polyproline II character ranging between 6 to 17%. A slight red shift was also observed with the highly concentrated AFGP8 solutions.

AFGP molecules protect fish at subzero temperatures, therefore their active conformation is present at these temperatures. Given the possibility that the conformation of AFGP8 might be different at lower temperatures, CD measurements were performed on the

same samples at -0.5°C (Figure 9 and Figure 10). The 0.5 and 1.0 mM solutions of AFGP8 possessed identical percentages of random coil (41%) and β-sheet (20%) secondary structures when compared to the same solutions at room temperature. α-Helix, polyproline II and β-turn secondary structures were equal contributors (approximately 13% each) to the remaining secondary structure. In contrast, the 5 mM AFGP8 solution possessed random coil as the predominant secondary structure (40%) but an 18% increase in the percentage of α-helix was observed. Interestingly, the ratio of random coil to α-helical content was identical to the 5 mM solution at 21°C with 30% total α-helix. The 10 mM AFGP8 solution exhibited a secondary structure of random coil (42%) with approximately 34% β-sheet.

At -0.5°C, the 20 to 100 mM solutions of AFGP8 exhibited varying degrees of random coil (20-55%) and β-sheet (37-45%) as the predominant secondary structures in solution. While these two conformations accounted for 80-90% of the solution conformations, it is likely that other conformations were present in solution. This is based upon the fact that a well-defined isodichroic point was not observed. As with the CD spectra at 21°C, a slight red shift was observed with increasing concentration.

Table 1: CD deconvolution of AFGP8 solutions

	≈ 41%	≈ 41%	≈ 40-48%
	≈ 13%	≈ 30%	≈ 0-8%
	≈ 20%	≈ 12%	≈ 35%
	≈ 11%	≈ 7%	≈ 6-17%
	≈ 14%	≈ 10%	≈ 5-12%

	≈ 41%	≈ 40%	≈ 42%	≈ 20-55%
	≈ 13%	≈ 31%	≈ 0%	≈ 0-8%
	≈ 20%	≈ 13%	≈ 34%	≈ 37-45%
	≈ 13%	≈ 5%	≈ 13%	≈ 5-12%
	≈ 13%	≈ 11%	≈ 11%	≈ 5-12%

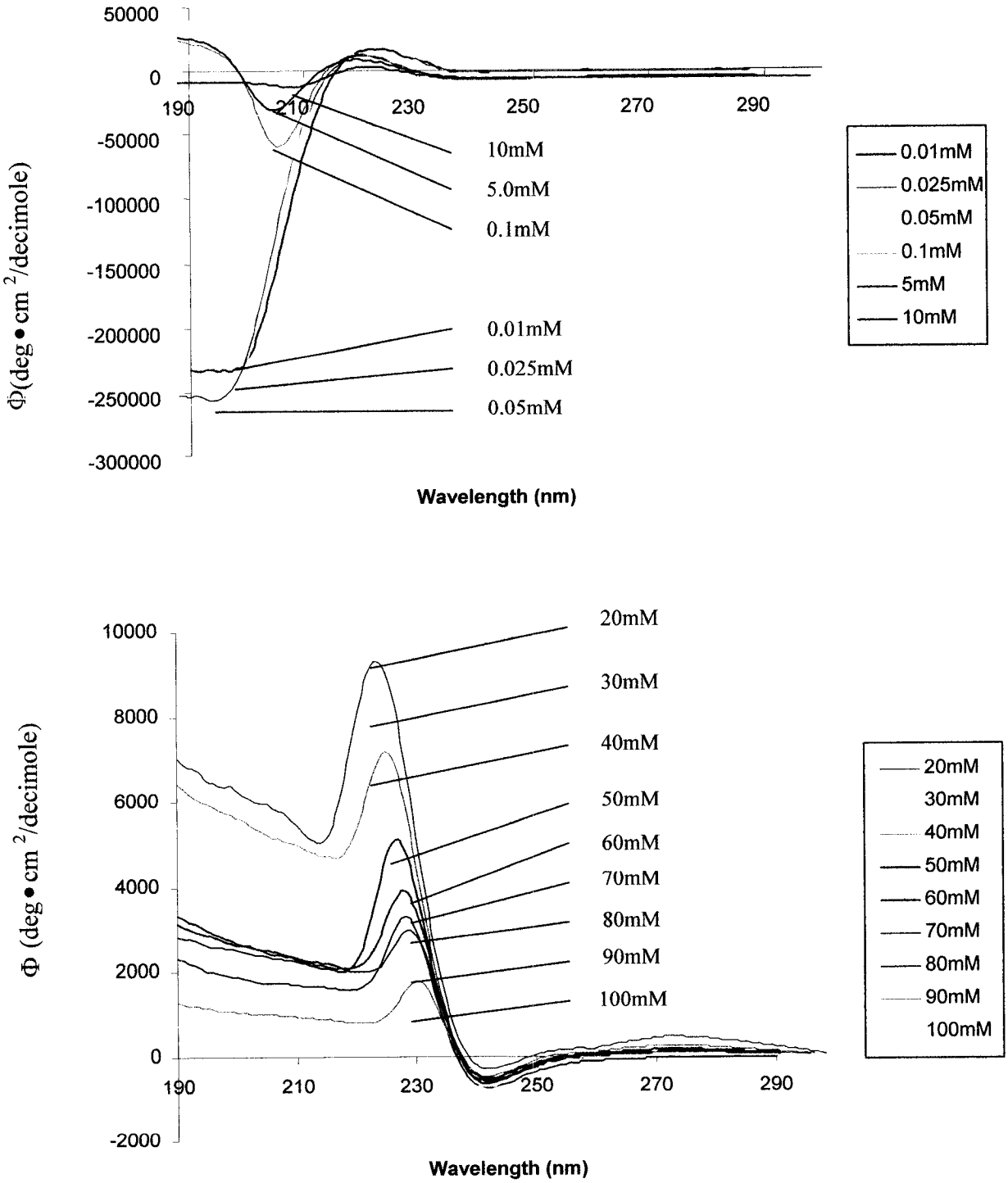


Figure 9: CD spectrum of AFGP8 in distilled H_2O as a function of concentration at 21 ^{o} C

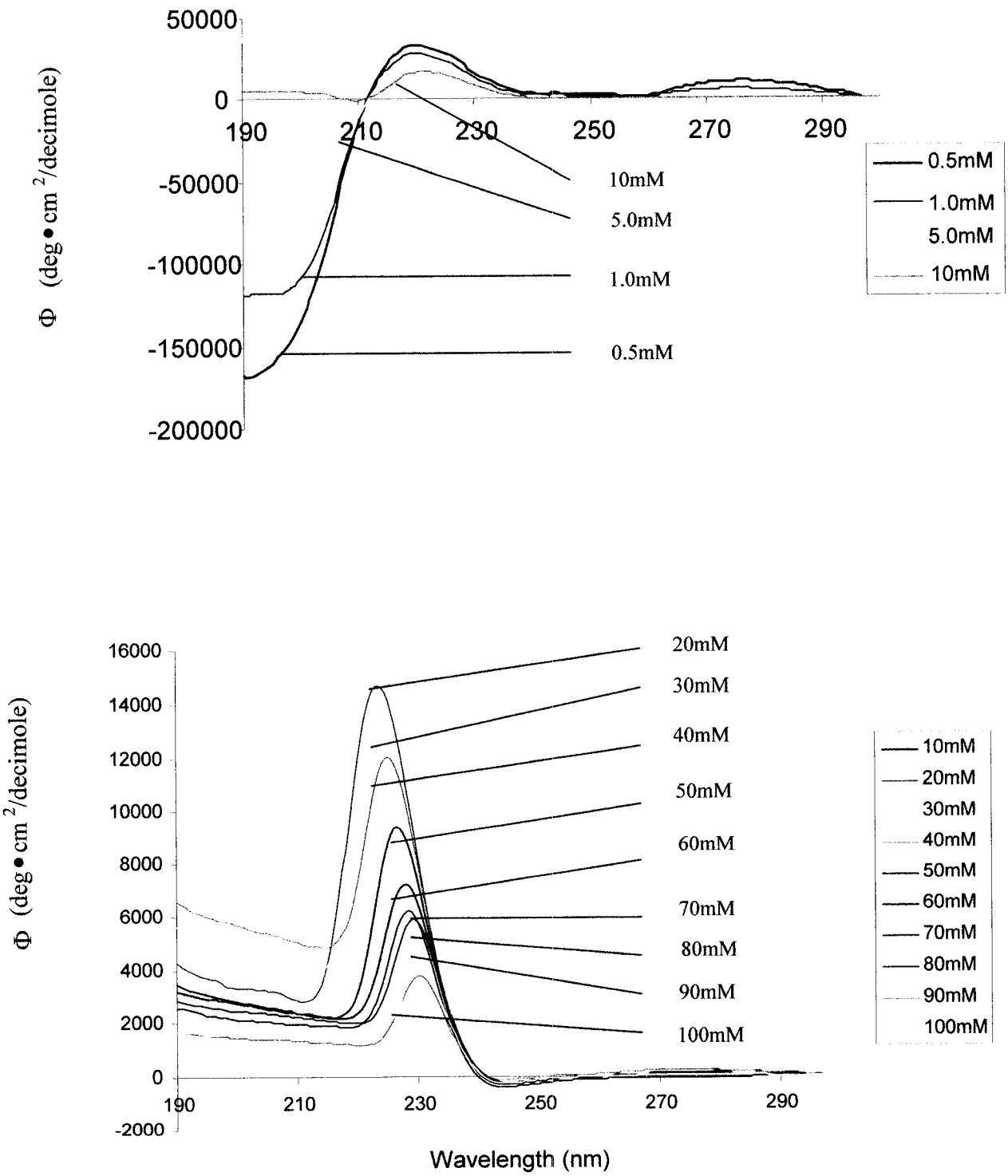


Figure 10: CD spectrum of AFGP8 in distilled H_2O as a function of concentration at -0.5° C

3.3.3 CD analysis and discussion

The solution conformation of AFGP8 at room temperature and 0°C is largely consistent with that of random coil structure, but other secondary structures appear to be present. No obvious conformational differences resulted from temperature changes. Recent reports describing the solution conformation of AFGP have suggested that it is consistent with a polyproline II structure.[5] However, our data is consistent with a random coil structure reported from earlier CD studies.[11] This apparent difference may arise from the fact that the CD spectrum of a random coil conformation is very similar to that of a polyproline type II[27] and that these two solution structures cannot be easily distinguished from each other.[28] Alternatively, the fact that a well-defined isodichroic point was not observed in our experiments indicates that other conformations are in equilibrium with the random coil structure, suggesting that the AFGP8 glycopeptide is highly flexible and assumes a number of solution conformations at different times (or in different molecules). This conclusion is supported by previous conformational studies of AFGP analogues.[6]

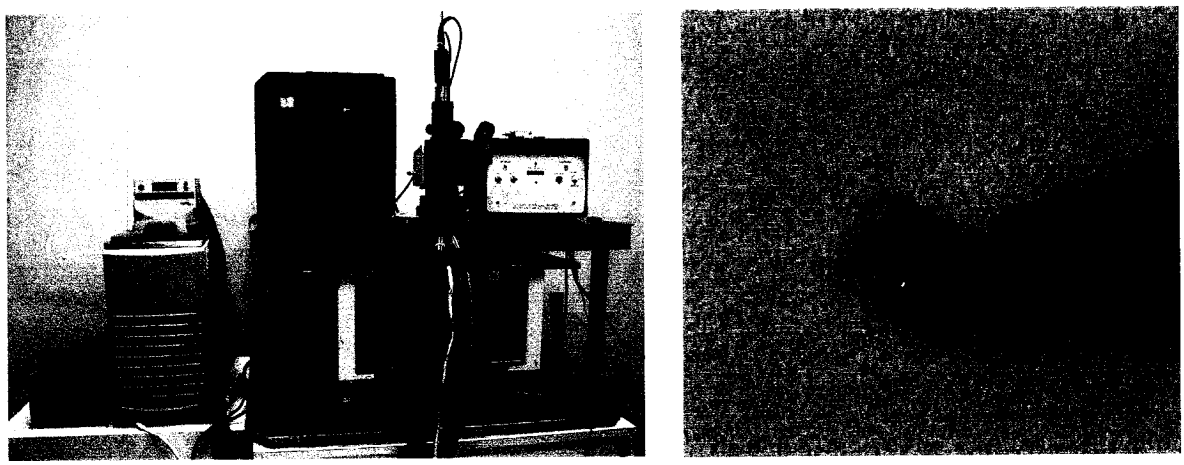
Recent NMR data suggests that AFGP8 possesses significant local order over the short fourteen amino acid sequence of an individual molecule but no long range order in solution.[19] This is consistent with our analysis as the interactions giving rise to the CD signals occur over relatively short distances (10-20 angstroms or 5-10 amino acid residues) and it would be expected that a number of different molecules may exhibit different conformational characteristics over a short period of time due to the inherent flexibility of the glycoprotein. The net result of this intermittent ordering would be an apparent mixture of solution conformations with random coil being the predominant secondary structure.[12] Given that it would be quite difficult for a protein to assume more than one secondary structure over a relatively short peptide sequence of fourteen residues, it is likely that individual AFGP8 molecules are adopting different secondary structures at different times. It is also possible that the different secondary structures arise from transient or stable interactions between monomers as is often found in coiled-coil proteins. Our findings are also in agreement with conformational studies performed using Raman spectroscopy[14] where the presence of an α-helix and β-sheet was observed, but the major conformation was considered to be random coil.

Another interesting feature in the CD analysis of AFGP8 is the prevalence of a shift for λ-max to decreased wavelength (red shift) in all sample concentrations between 20-100 mM. This trend was observed at room temperature and 0°C (Figure 9 and Figure 10). While the interpretation of a red shift is somewhat controversial, it is consistent with the formation of aggregates in solution,[29] and consequently correlates well with the results from our DLS experiments. It appears that AFGP8 aggregation does not facilitate significant changes in solution conformation, as the 20-100 mM AFGP solutions all possess random coil structure as the predominant conformation. However, the incremental conformational changes that are observed in these more concentrated solutions (i.e. slight increases in α-helix etc.) may be a direct result of the aggregation. These incremental changes were completely reversible and the fact that a well-defined isodichroic point did not exist suggested that more than two conformations are present in solution.

3.4 Relationship between aggregation and antifreeze activity of AFGP8:

3.4.1 Assessment of Thermal Hysteresis using nanoliter osmometry.

The most interesting aspect of AFGP8 aggregation is not that discrete aggregates are formed, but rather the effect these aggregates may have on antifreeze activity. The "golden standard" for assessment of antifreeze activity is thermal hysteresis using a Nanoliter Osmometer (Clifton Technical Physics) (Figure 11).



A **B**

Figure 11: ***A: Clifton Technical Physics Nanoliter Osmometer.*** *Instrument range 0 to 5,000 milliOsmos (0°C to 9.3°C) (1.0 Os=1.86°C). B: Sample Holder*

This technique has the advantage of allowing the observation of the ice habit as well as the growth patterns. As described in Section 1.3.1, only biological antifreezes have the unique ability to bind to specific surfaces of an ice crystal, becoming incorporated into the ice lattice and changing the ice crystal morphology and growth. To test our AFGP8 solutions, the AFGP8 is solubilized in distilled water. Then, a nanoliter droplet of the test solution is submerged in pre-cooled oil in the sample holder, after which the sample is flash-frozen at –40°C. At this step an ice sphere containing several thousand ice crystals is formed. The temperature is gradually increased until the sample is melted back to a single ice crystal, and its melting point is determined to the nearest one thousand of a degree Celsius. Once the melting point is determined, the temperature is lowered [10 mOsmol (0.0186°C) units every 15 seconds] to determine the freezing point. In the absence of biological antifreezes, the ice crystal begins to grow as soon as the temperature is lowered. The ice growth is rapid and continuous until the entire sample is frozen. The ice crystal is a rounded sphere shown in Figure 12A,[30] which grows without any observable edges or facets. The same is observed in the presence of substances that depress the freezing point in a colligative fashion, such as sodium chloride. With biological antifreezes, the crystal is no longer a rounded sphere, but contains defined sharp edges, as seen in Figure 12B.[30]

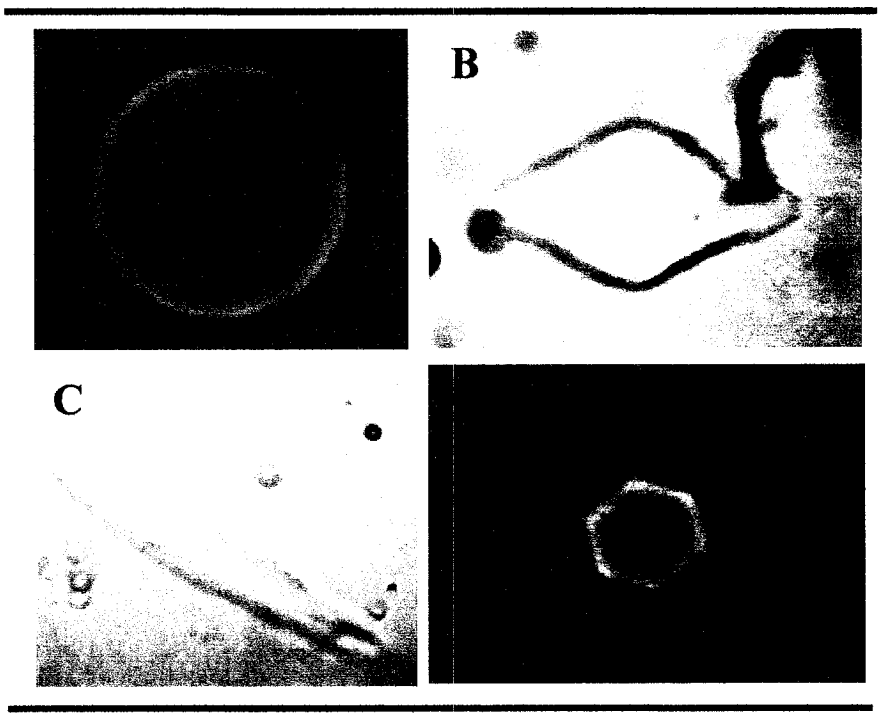


Figure 12: Nanoliter Osmometry images[30]
A) ice crystal observed with distilled water; B) ice crystal in the presence of AFGP8, side view and top view of a hexagonal bipyramidal crystal; C) ice crystal in the presence of AFGP8, ice spicule formation below freezing point.

In the presence of native AFGP8 the ice crystal is hexagonal bipyramidal (Figure 12B); the shape varies with different types of AFPs and AFGPs. As the temperature is lowered, the crystal maintains its shape and does not grow larger within the range of the thermal hysteresis gap. However, once the freezing point is reached, the ice crystal undergoes a "burst growth". Ice crystals in the presence of AFGP8 form sharp needle-like structures that grow very quickly once the temperature exceeds the freezing point as shown in Figure 12C.

3.4.2 Thermal Hysteresis (TH) measurements

The TH profile for AFGP8 isolated from *Gadus ogac* has been reported for concentrations ranging from 1-20 mg/mL (0.4 -7.5 mM).[21] In contrast, we have assessed TH activity over a much wider concentration range (0.05 to 60 mM) (Figure 13). All measurements were performed three times and the freezing point was recorded once "burst

crystal growth" occurred. While our TH values are slightly higher than previously reported, this was attributed to a small amount of AFGP7 present as a contaminant in the sample. Solutions of AFGP8 up to 10 mM exhibit a TH profile typical of a non-colligatively acting biological antifreeze. The plateau observed between 10 mM to 35 mM is also consistent with a non-colligative effect.

However, the TH profile increased sharply above 35 mM and continued to increase in a linear fashion indicating that AFGP8 solutions ranging in concentration from 40-60 mM are much more active than the dilute solutions. We believe the plateau in Figure 13 is representative of monomeric AFGP8 while the 40-60 mM solutions contain predominantly aggregated AFGP8. Interestingly, the ice crystal morphology observed with these AFGP8 solutions was indistinguishable from that of the 0.5-30 mM solutions. Given that our DLS experiments indicate that a 40 mM AFGP8 solution possesses a large amount of aggregates, we conclude that these aggregates are responsible for the observed increase in activity.

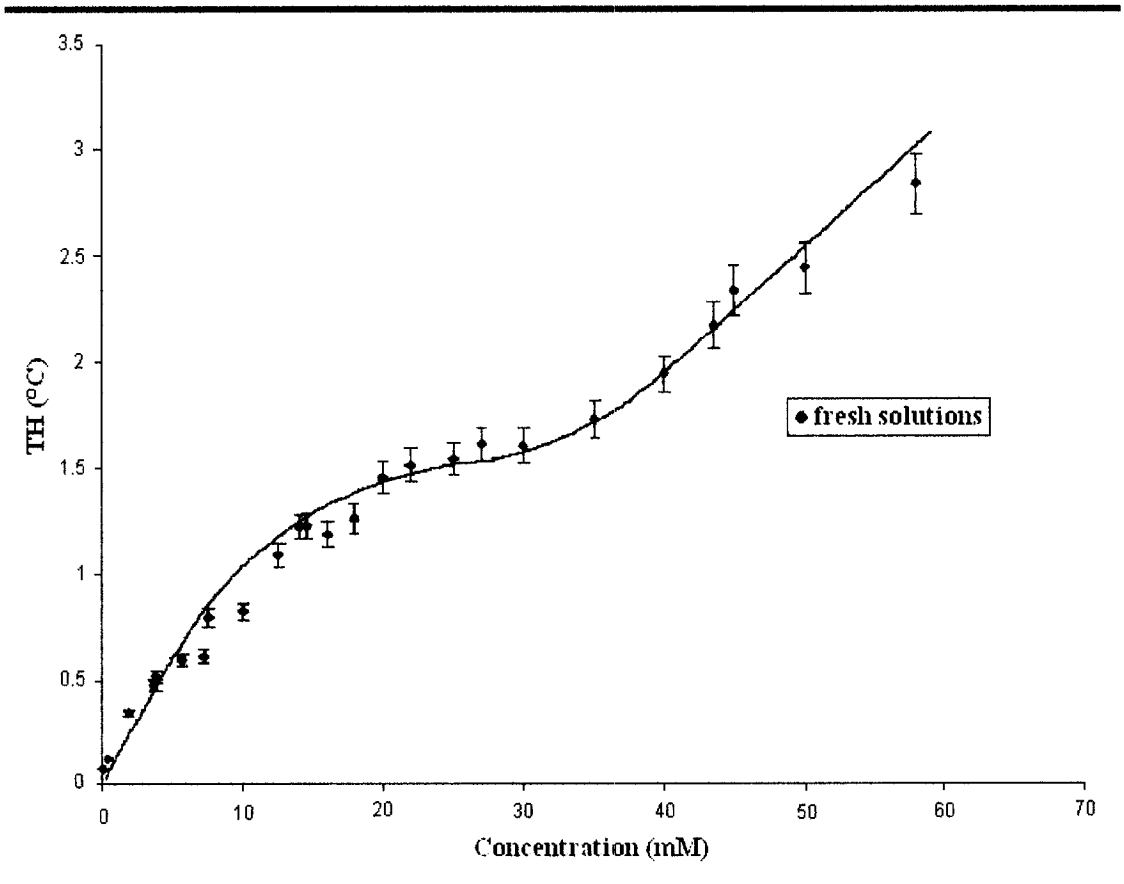


Figure 13: Thermal hysteresis activity as a function of AFGP8 concentration

A plausible explanation for the increase in TH activity may be attributed to AFGP8 binding to the ice surface as an aggregate and not a single AFGP8 molecule. The increased size of the AFGP8 aggregate (relative to a single AFGP8 molecule) would result in a decreased surface area of the exposed ice front between adjacent aggregates. As a consequence, the surface curvature in areas between adjacent aggregates would increase and the amount of energy required to incorporate additional water molecules into the ice lattice would also increase dramatically.[1] A similar effect has been reported with AFP-fusion proteins.[31] Our results are consistent with the fact that lower molecular weight fractions of AFGPs have been shown to work in a cooperative manner with larger fractions, (AFGP1-5) resulting in a 2-8 fold potentiation of thermal hysteresis activity.[32,33] However, this is the first example where lower molecular weight AFGP fractions have demonstrated such a cooperative effect in the absence of higher molecular weight fractions or other proteins. The biological significance of the increased TH activity as a function of aggregation is not understood at this time as aggregation starts to appear in solution at concentrations equal to 60 mg/mL *in vitro*. To date, such high concentrations of AFGP have not been reported *in vivo*. However, it should be noted that the smaller fractions of AFGP such as AFGP8 are typically found in much higher concentrations than the larger fractions in certain fish blood.

We have demonstrated that AFGP8 forms aggregates in solution consisting of at least two individual AFGP8 molecules. These aggregates increase in size as a function of time. The formation of aggregates appears to occur at solution concentrations greater than 20 mM in freshly prepared solution and 1 mM in solution allowed to stand for 2 months. While these results confirm the propensity of AFGP8 to form aggregates in solution, the size of the aggregates observed using DLS are much smaller than was previously observed using atomic force microscopy.[20] This may be explained by the fact that AFM resolution in the X and Y dimensions is an order of magnitude better than its resolution in the Z-dimension. Consequently, the size of the aggregates as determined using AFM would have been overestimated.

Our analyses of AFGP secondary structure in aqueous solution using CD spectroscopy suggest that random coil is the major conformation at room temperature in 0.01 to 100 mM solutions. However, large amounts of β-sheet and α-helix character are also present in the more concentrated solutions (i.e. 10-100 mM). A distinct red shift at concentrations between 20-100 mM is consistent with aggregate formation. Analysis of the

same solutions at -0.5 °C revealed no significant differences in solution conformation relative to those analyzed at room temperature. The largest amount of random coil structure was seen at concentrations between 0.01 to 5 mM irrespective of temperature.

In contrast to AFGP1-5, the propensity for AFGP8 to aggregate has not been reported. This ability may be unique to the lower molecular weight AFGPs. A correlation of AFGP8 concentration and thermal hysteresis activity indicates that aggregated AFGP8 solutions are significantly more active than the monomeric species. While the cooperative functioning of AFGPs has been demonstrated in a solution containing both AFGP8 and AFGP1-5, this is the first example in which AFGP8 demonstrates a cooperative effect in the absence of higher molecular weight fractions. The physiological implications of this are not clear as AFGP8 aggregation appears to occur at concentrations much higher than those reported in deep sea Teleost fish. The interactions of the dimeric species with the ice lattice have not been addressed and will be the topic of future work.

[1] Wilson, P. W. *Cryo-Letters* **1993**, 14, 31.

[2] Knight, C. A. *Nature* **2000**, 406, 249.

[3] Harding, M. M.; Anderberg, P. I.; Haymet, A. D. J. *Eur. J. Biochem.* **2003**, 270, 1381.

[4] Knight, C.A.; Driggers, E.; DeVries, A.L. *Biophys. J.* **1993**, 64, 252.

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

[6] Filira, F.; Scolaro, B. B.; Foffani, M. T.; Mammi, S.; Peggion, R.; Rocchi, R. *Int. J. Of Biol. Macromol.* **1990**, 12, 41.

[7] Franks F.; Morris E. R. *Biochem. Biophys. Acta.* **1978**, 540, 346.

[8] Rao, B. N.; Bush, C. A. *Biopolymers* **1987**, 26, 1227.

[9] Tachibana, Y.; Matsubara, N.; Nakajima, F.; Tsuda, T.; Tsude, S.; Monde, K.; Nishimura, S.-I. *Tetrahedron* **2002**, 58, 10213.

[10] DeVries, A. L.; Komatsu, S. K.; Feeney, R. E. *J. Biol. Chem.* **1970**, 245, 2901.

[11] Raymond, J. A.; Radding, W.; DeVries, A. L. *Biopolymers* **1977**, 16, 2575.

[12] Bush, C. A.; Feeney, R. E.; Osuga, D. T.; Ralapati, S.; Yeh, Y. *Int. J. Peptide Protein Res.* **1981**, 17, 125.

[13] Ahmed, A. I.; Feeney, R. E.; Osuga, D. T,; Yeh, Y. *J. Biol. Chem.* **1975**, 250, 3344.

[14] Yomimatsu, Y.; Scherer, J. R.; Yin Y., Feeney, R. E. *J. Biol. Chem.* **1976**, 251, 2290.

[15] Bush C. A.; Feeney R. E. *Int. J. Peptide Protein Res.* **1986**, 28, 386.

[16] Berman, E.; Allerhand, A.; DeVries, A. L. *J. Biol. Chem.* **1980**, 255, 4407.

[17] Bush, C. A.; Ralapati, S.; Matson, G. M.; Yamasaki, R. B.; Osuga, D. T.; Yeh, Y.; Feeney, R. E. *Arch. of Biochem. and Biophys.* **1984**, 232, 624.

[18] Lane, A. N.; Hays, L. M.; Feeney, R. E.; Crowe, L. M.; Crowe, J. H. *Protein Science* **1998**, 7, 1555.

[19] Lane, A. N., Hays, L. M., Tsvetkova, N., Feeney, R. E., Crowe, L. M., Crowe, J. H. *Biophys. J.* **2000**, 78, 3195.

[20] Sarno, D.; Murphy, A. V.; DiVirgilio, E. S.; Jones, W. E., Jr.; Ben, R. N. *Langmuir* **2003**, 19, 4740.

[21] Wu, Y.; Banoub, J.; Goddard, S. V.; Kao, M. H.; Fletcher, G. L. *Comparative Biochemistry and Physiology Part B* **2001**, 128, 265.

[22] "Measurement of Proteins in Solution", Application Note from ProteinSolutions DynaPro Dynamic Light Scattering Instruments.

The description of the equations were excerpted from:

http://mxp.physics.umn.edu/s05/Projects/s05lightscattering/introduction.html

[23] Berne, B.J.; Pecora, R. *Dynamic Light Scattering: With Applications to Chemistry, Biology, and Physics,* Dover Publications, **2000**.

[24] A similar result was observed when AFGP4 polydispersity was studied using quasi-elastic light scattering (QELS or DLS). Other AFGP fractions failed to show the same polydispersity. See reference 16.

[25] Fasman G.D. Circular Dichroism and the Conformational Analysis of Biomolecules., Editor. Plenun Press, New York pp **1996**, 635.

[26] a) Greenfield, N. J. *Analytical Biochemistry* **1996**, 235, 1. b) Sreerama, N.; Venyaminov, S. Y.; Woody, R. W. *Analytical Biochemistry* **2000**, 287, 243. c) Sreerama, N.; Woody, R. W. *Analytical Biochemistry* **2000**, 287, 252.

[27] Steim, J. M.; Fleischer, S. *Proc. Natl. Acad. Sci* **1967**, 58, 1292.

[28] Davies, P. L.; Hew, C. L. *FASEB J.* **1990**, 4, 2460.

[29] Chakrabartty, A.; Yang, D. S. C.; Hew, C.L. *J. Biol. Chem.* **1989**, 264, 11307.

[30] Murphy A.V. Ph.D. Thesis, State University of Binghamton, Binghamton, **2004**.

[31] Deluca, C. I.; Comley, R.; Davies, P. L. *Biophys. J.* **1998**, 74, 1502.

[32] Osuga, D. T.; Ward, F. C.; Yeh, Y.; Feeney, R. E. *J. Biol. Chem.* **1978**, 253, 6669.

[33] Mulvihill, D. M.; Geoghegan, K. F.; Yeh, Y.; DeRemer, K.; Osuga, D. T.; Ward, F. C.; Feeney, R. E. *J. Biol. Chem.* **1980**, 255, 659.

Chapter

SYNTHESIS AND ACTIVITY ASSESSMENT OF *O*-LINKED AFGP8 ANALOGUES

Bouvet, Vincent R., Ben, Robert N.*ACS Symposium Series. **2004**, 896, 151-156.

4 SYNTHESIS AND ACTIVITY ASSESSMENT OF *O*-LINKED AFGP8 ANALOGUES

4.1 Formation of *O*-linked glycosidic bonds: a general introduction

Since the mid-1970's, carbohydrates have created a growing interest because they are commonly found in plants, animals and bacteria. To investigate their significance in different biological mechanisms, several techniques to form a glycosidic linkage have been published.[1] A glycosidic linkage is generally formed using a glycosyl donor (α or β configuration at C-1) and a glycosyl acceptor (Figure 1). Generally, the stereoselectivity of this reaction is high and no changes in the conformation of the glysosyl donor are observed.

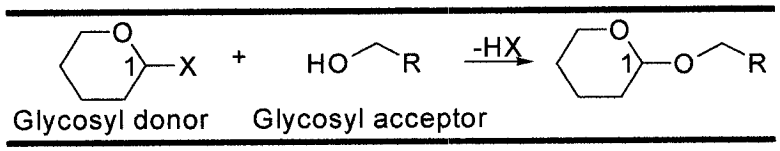


Figure 1: Formation of glycosidic linkages

Most of the methods published for the formation of the glycosidic linkage utilize a glycosyl donor that is a precursor of an intermediate oxacarbenium ion (Figure 2).

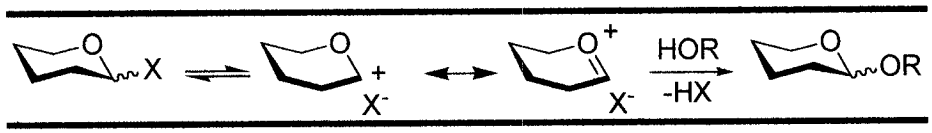


Figure 2: Oxocarbenium formation

Several factors can influence the formation of a glycosidic linkage. First, shielding of one face of the oxacarbenium can occur because of the presence of solvents or the counter-ion. Secondly, the substituent at the C-2 position is important for the selectivity of the glycosylation. For instance, if the protecting group is an ester, an anchimeric effect can occur and the formation of a β-glycosidic linkage is favored if C-2 is equatorial (Figure 3).

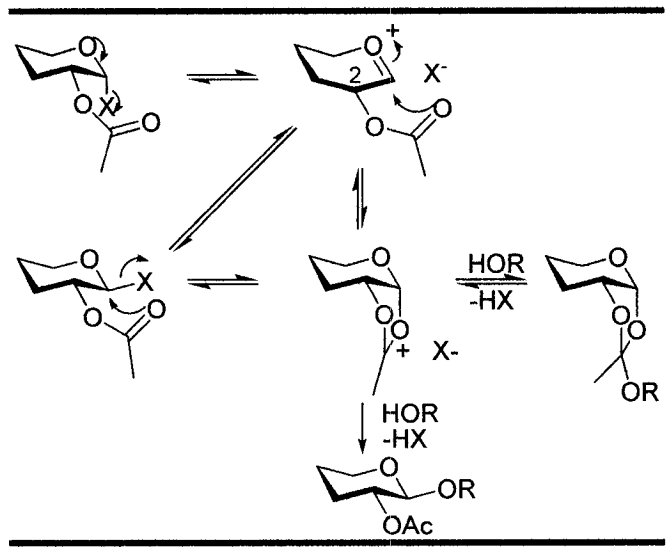


Figure 3: Example of anchimeric effect[1]

One of the first syntheses of glycosidic linkages was reported by Koenigs and Knorr in 1901.[2] Acetobromoglucose was treated with alcohols in presence of silver(I) carbonate and a new glycosidic linkage was formed (Figure 4). Since this experiment, several improvements were published. These improvements include using co-solvents, desiccants and metal promoters. The heterogeneous process is believed to be bimolecular resulting in inversion of the stereochemistry at the anomeric center (Figure 4). For the homogeneous process, an unimolecular heterolysis of the C1-Br bond results in an anomeric carbocation which is then stabilized by a C-2 anchimeric effect. This results in the formation of a β linkage when C-2 is equatorial. When there is no ester group at the C-2 position, no anchimeric effect is observed and therefore, the halide glycosyl donor favors the formation of the α anomer. This is presumably due to the thermodynamic stabilization resulting from the anomeric effect (see chapter 1).

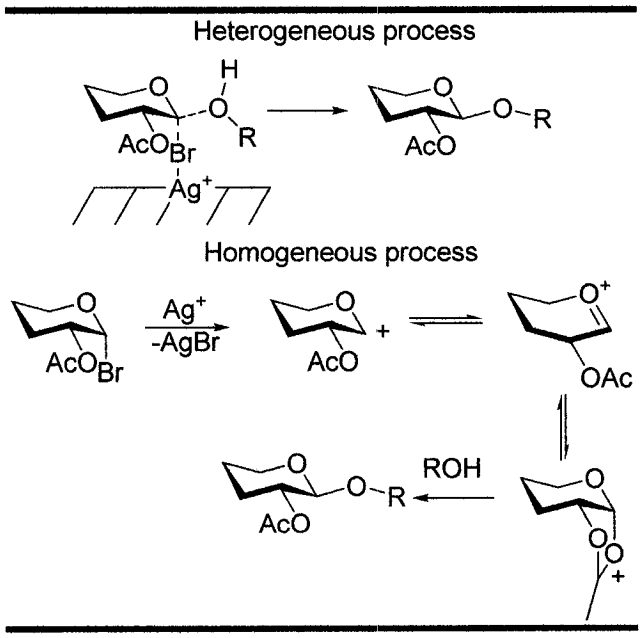


Figure 4: Mechanism of the Koenigs and Knorr reaction[2]

Paulsen[3] first noticed the difference in stability between the tetra-*O*-benzyl-α-D-glucopyranosyl bromide and its acetate counterpart. The benzyl ethers are electronically passive groups that favor the formation of positive charge at the anomeric position ("armed"). However, the acetate groups are electron withdrawing and therefore disfavor the formation of this charge ("disarmed"). The implication of this is that the "armed" glycosyl donor is several folds more reactive than the "disarmed" donor and thus it is possible to selectively couple the "armed" glycosyl donor in presence of the disarmed donor. This armed/disarmed concept was applied to additional glycosyl donors (Figure 5).

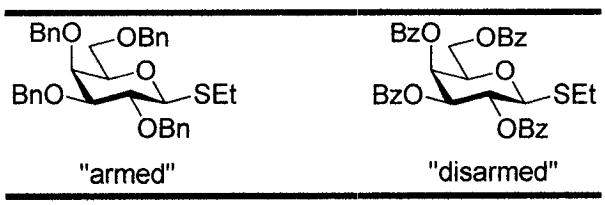


Figure 5: Armed and disarmed concept

Since 1980, several new types of glycosyl donors that improve the stereoselectivity of the anomeric position appeared in the literature. Stable glycoside fluorides[4,5] with promoters as well as thioglycosides,[6,7] glycosyl sulfoxides,[8] seleno- and telluro-glycosides[9,10] have been used (Figure 6).

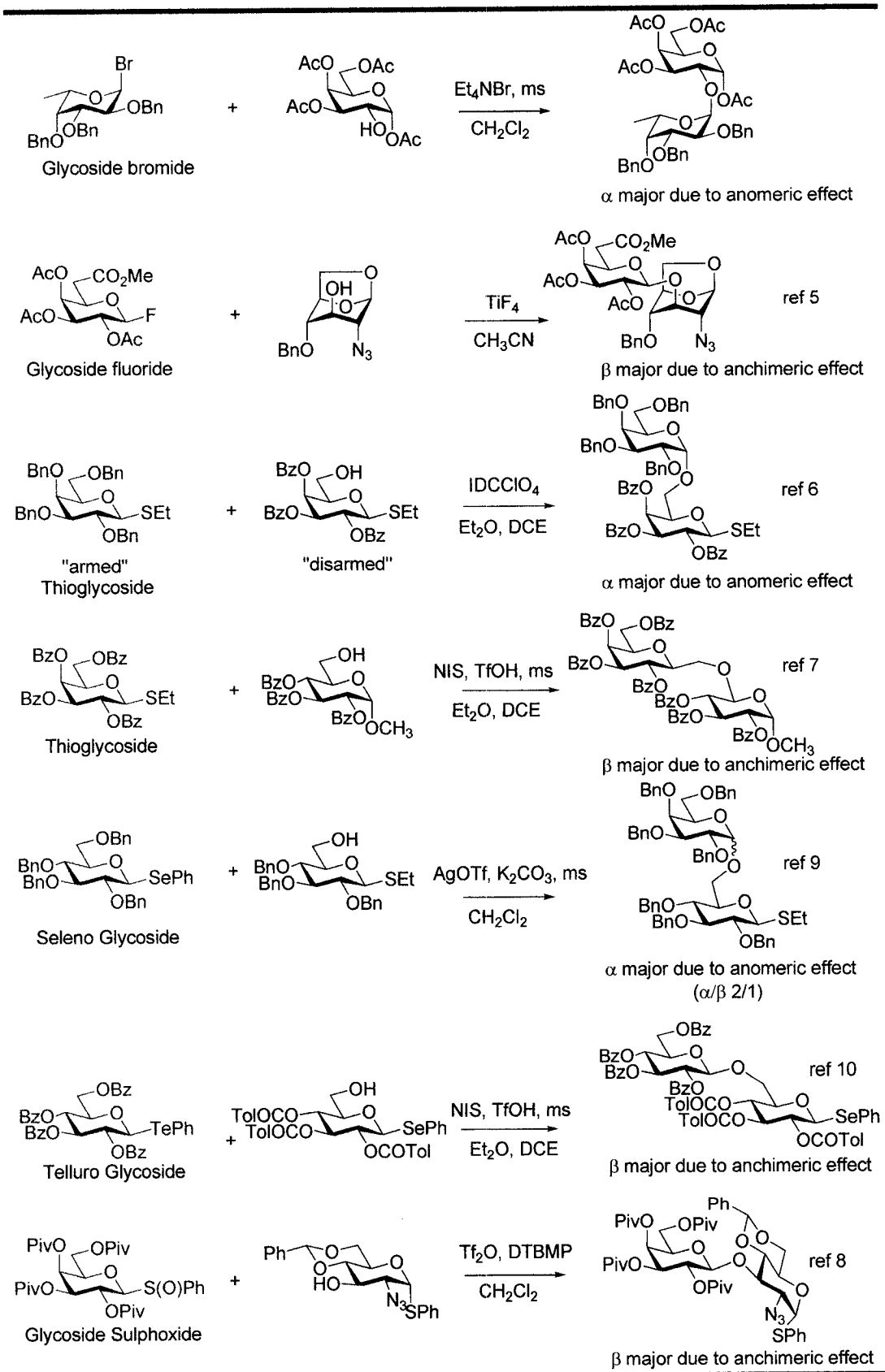


Figure 6: Example of glycosyl donors[4-10]

These glycosyl donors can be activated using different conditions, and are therefore compatible with each other; consequently, one can couple a glycosyl donor without activating another. In the late 1980s, Fraser-Reid *et al.*[11] discovered and optimized the utilization of 4-pentenyl glycoside as a glycosyl donor. This technique gained popularity due to the relative stability of the pentenyl group and its easy activation (Figure 7).

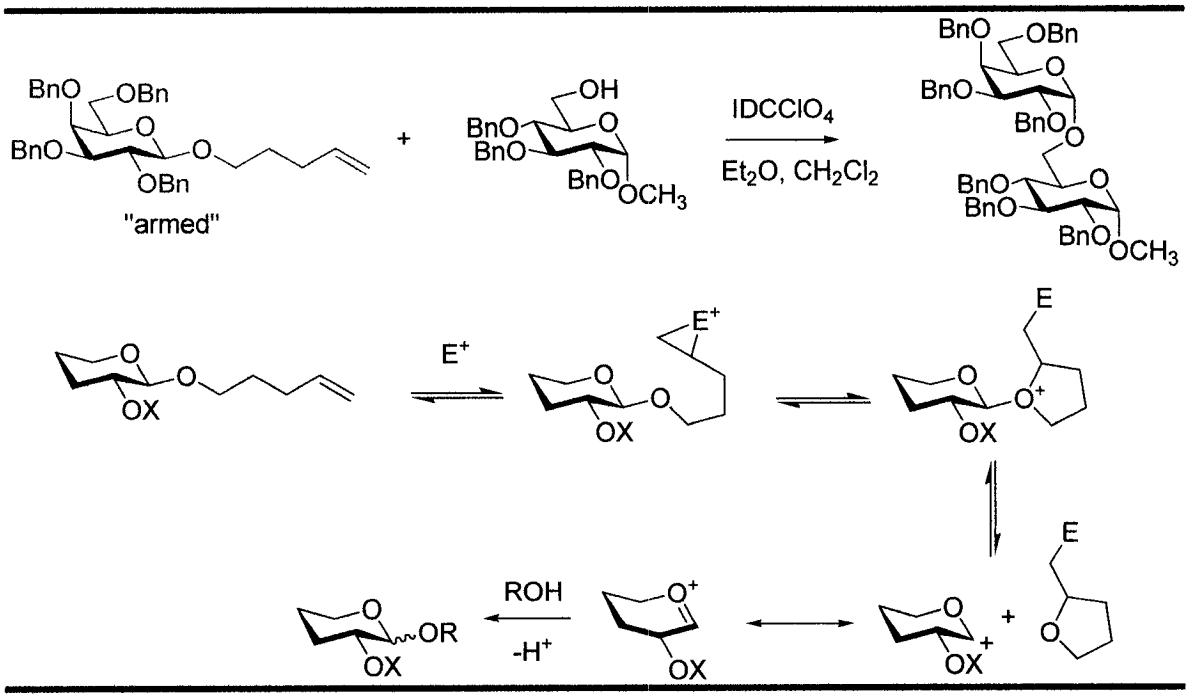


Figure 7: 4-pentenyl glycosides donors[11]

The last general method to activate carbohydrates that will be discussed is a trichloroimidate (TCA) method. This method was first used by Sinay[12] and further developed by Schmidt[13] and is now one of the most common techniques for the formation of the glycosidic linkage. This is mainly due to the activation which is heavy metal-free, and the fact that they have the ability to favor one anomer. The treatment of the glycosyl donor with trichloroacetonitrile and different bases gives rise to anomerically pure trichloroimidate (Figure 8). This selectivity on forming different anomers was attributed to a kinetic/thermodynamic process, where stronger bases and a longer reaction time favor the formation of the more stable α anomer. Once the trichloroimidate is obtained, several factors can be controlled to selectively obtain either the α or β glycosidic linkage.

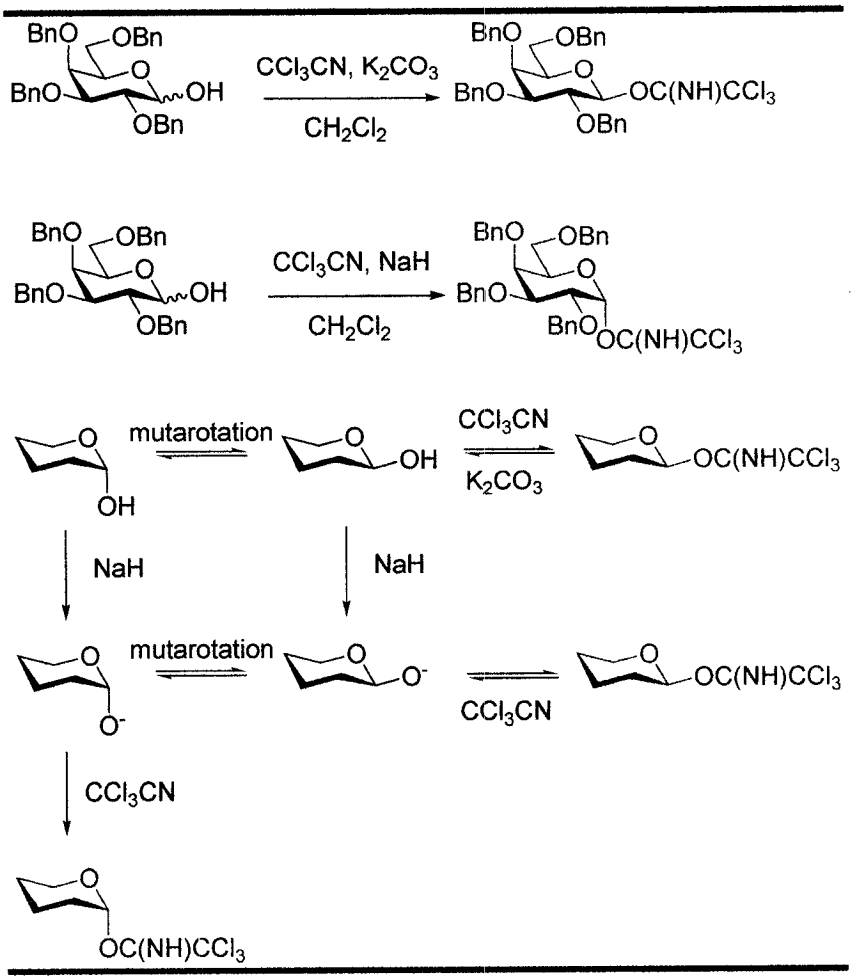


Figure 8: Formation of the trichloroimidate[1]

For TCA donors containing a non-participating group at C-2, treatment with a relatively mild Lewis acid like BF_3-OEt_2 generally results in an S_N2 type displacement. For glycosylation promoted by TMSOTf the reaction proceeds through the formation of the oxacarbenium ion. In dichloromethane or ether the stable α-D-glycoside is formed; however in acetonitrile at low temperature, the rapid formation of the "α-nitrilium" is favored leading to the formation of the β-D-anomer (Figure 9). For TCA donors possessing a participating group at C-2, the expected anomer resulting from the anchimeric effect is generally the major product observed.

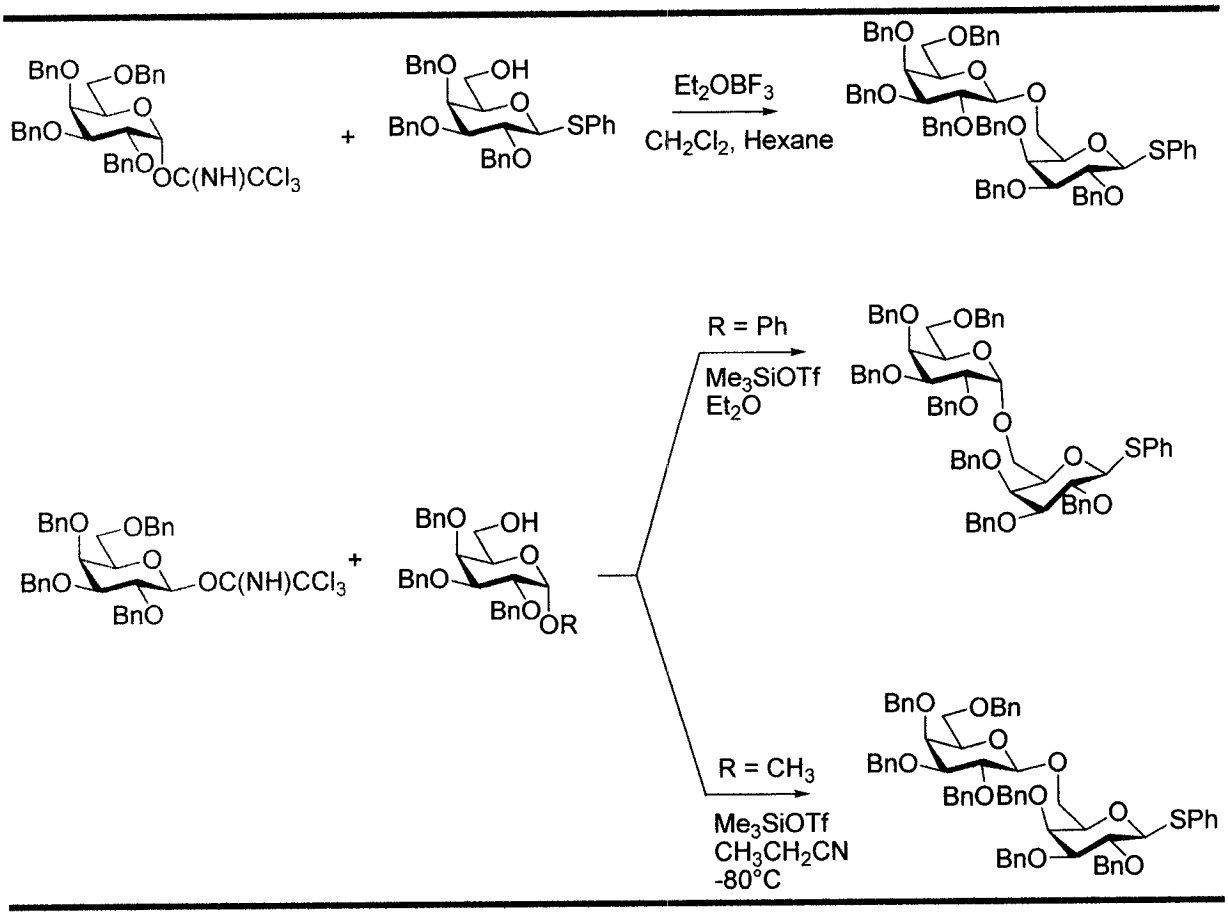


Figure 9: Stereoselectivity of the trichloroimidate method[1]

4.2 Synthesis of *O*-linked AFGP analogues

As mentioned in the previous chapter, in order to carry out a SAR study on the *O*-linked AFGP8, we choosed to synthesize the eight analogues shown in Figure 10.

4.2.1 O-Linked AFGP analogues

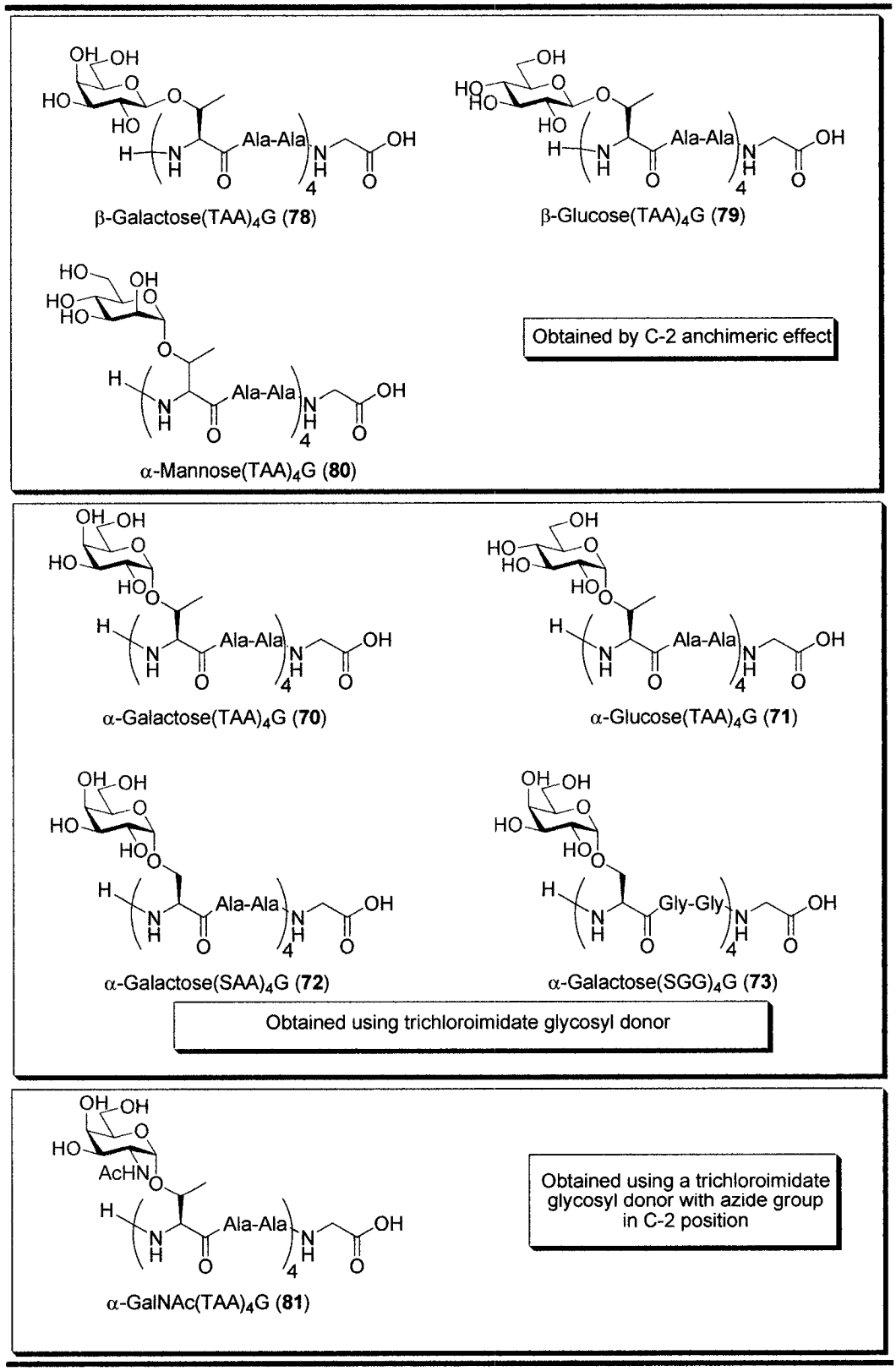


Figure 10: AFGP8 analogues for a SAR study on O-linked system

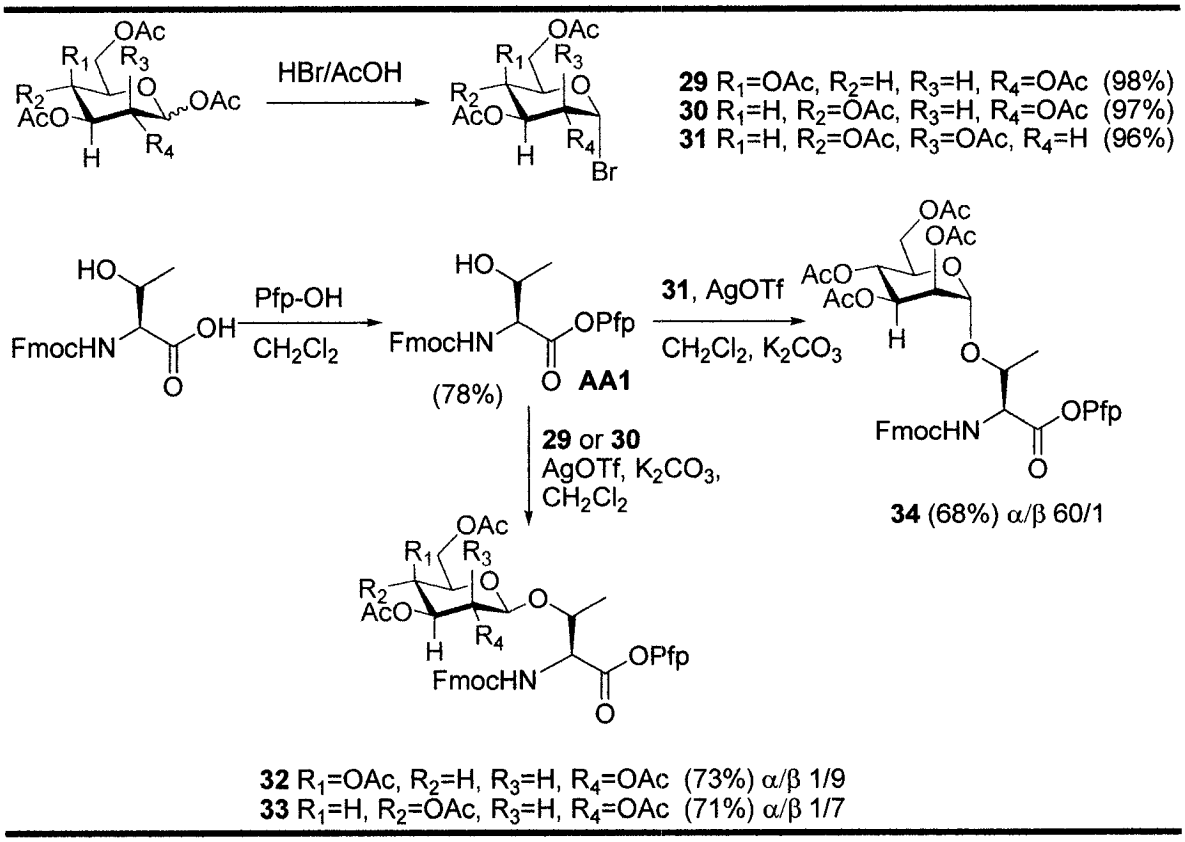
The synthesis of these analogues is performed using three different strategies. To obtain the β-galactose analogue **75**, the β-glucose analogue **76** and the α-mannose analogue **77**, we will use the anchimeric effect from the participating protecting group in the C-2 position. α-Galactose analogues **67**, **69** and α-glucose analogues **78** can be obtained using trichloroimidate glycosyl donors under thermodynamic conditions.

The synthesis of the α-acetyl-galactosamine is more labour intensive; indeed, in this strategy the presence of the *N*-acetyl participating group on the C-2 position makes accessing the α-anomer more challenging. Consequently, to obtain this analogue, the presence of an inert and convertible functionality is necessary. One of the most widely used chemical groups for this purpose is the azide functionality. When this azide functionality is introduced, a regular thermodynamic process involving the trichloroimidate glycosyl donor should give access to the desired analogue.

4.2.2 Synthesis of β-galactose, β-glucose and α-mannose building blocks 38, 39 and 40

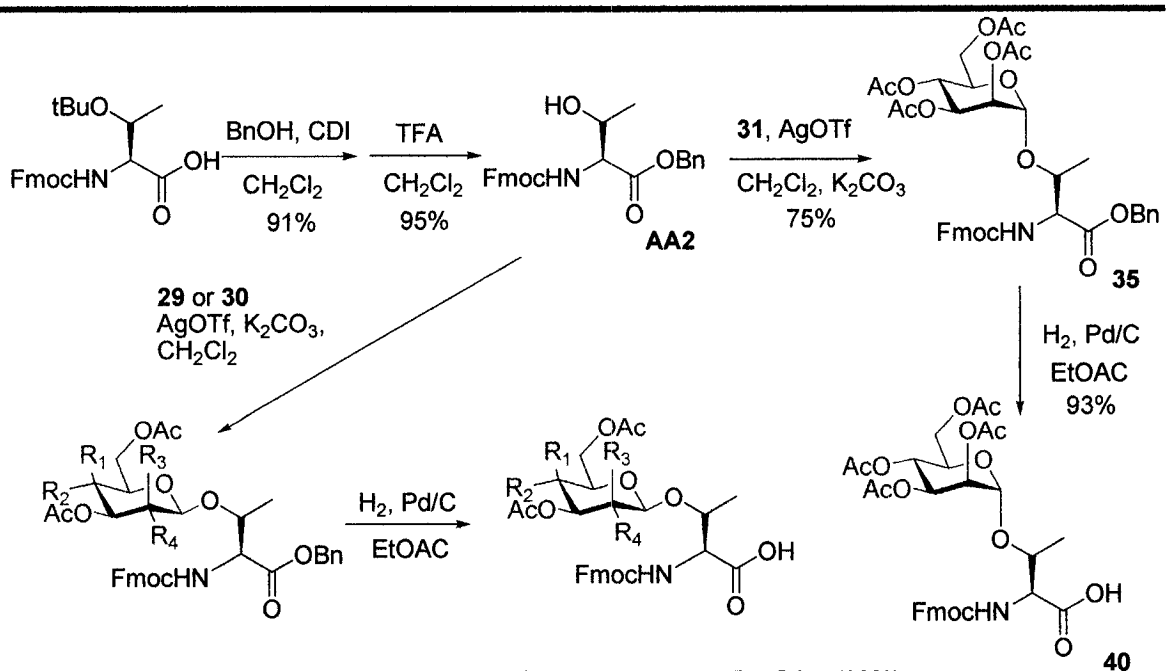
In order to obtain these analogues, we needed to prepare the glycosyl bromides **29**, **30** and **31**. Toward this goal galactose, glucose and mannose were treated with acetic anhydride, pyridine and DMAP to produce an α/β mixture of the *per*-acetatylated glycoside. This mixture then underwent a bromination using 30% HBr in acetic acid to afford pure α-bromoglycoside. The acid functionality of the amino acid was then protected using pentafluorophenol.[14,15] The use of pentafluorophenol affords two advantages. It is sufficiently stable to survive a glycoside coupling and it is reactive enough to play the role of an activated ester during the subsequent solid phase synthesis. The amino acid **AA1** was then coupled[16] to the different glycosyl donors to obtain compounds **32**, **33** and **34** (Scheme 1).

Scheme 1: Synthesis of β-glucose, β-galactose and α-mannose building blocks using pentafluorophenol protected amino acids



The galactose building block was obtained as a 1/9 (α/β) mixture, the glucose derivative was obtained as a 1/7 (α/β) mixture and finally the mannose building block was obtained with a 60/1 (α/β) ratio. All compounds were further purified by HPLC to afford pure building blocks suitable for solid phase synthesis. Although this is a short and high yielding synthesis, the instability of these compounds (i.e. pentafluorophenyl ester) does not make this strategy amenable for large scale preparation. Indeed, due to the high sensitivity of the pentafluorophenyl carbonate, the purification, storage and the glycoside coupling of large quantities of these compounds was tedious and a great deal of hydrolyzed by-product was observed. Therefore we decided to use benzyl ethers as a protecting group instead (Scheme 2).

Scheme 2: Synthesis of β-glucose, β-galactose and α-mannose building blocks using benzyl alcohol protected amino acids



36 R_1=OAc, R_2=H, R_3=H, R_4=OAc (75%) **38** R_1=OAc, R_2=H, R_3=H, R_4=OAc (93%)
37 R_1=H, R_2=OAc, R_3=H, R_4=OAc (82%) **39** R_1=H, R_2=OAc, R_3=H, R_4=OAc (96%)

In this new approach several additional steps were necessitated namely the deprotection of the *t*-butyl ether on the amino acid, debenzylation of the benzyl ester before the solid phase synthesis and pre-activation of the building block in order to perform the solid phase coupling. Despite these extra steps, the preparation of the building blocks has been greatly simplified: while the overall yields are comparable, making this approach amenable for scale up. These three extra steps afforded the desired compounds with greater than 90% yields and made this strategy amenable for the preparation of large scale and storage. The coupling of **AA2**[17] with the different glycosyl bromide affords the different analogues with the same α/β ratio than for the compounds **32**, **33** and **34**.[16] The final building blocks **38**, **39** and **40** are also easily prepared using regular hydrogenation conditions.

Table 1 summarizes the different yields and number of steps required in order to obtain suitable building block to synthesize the AFGP8 analogues.

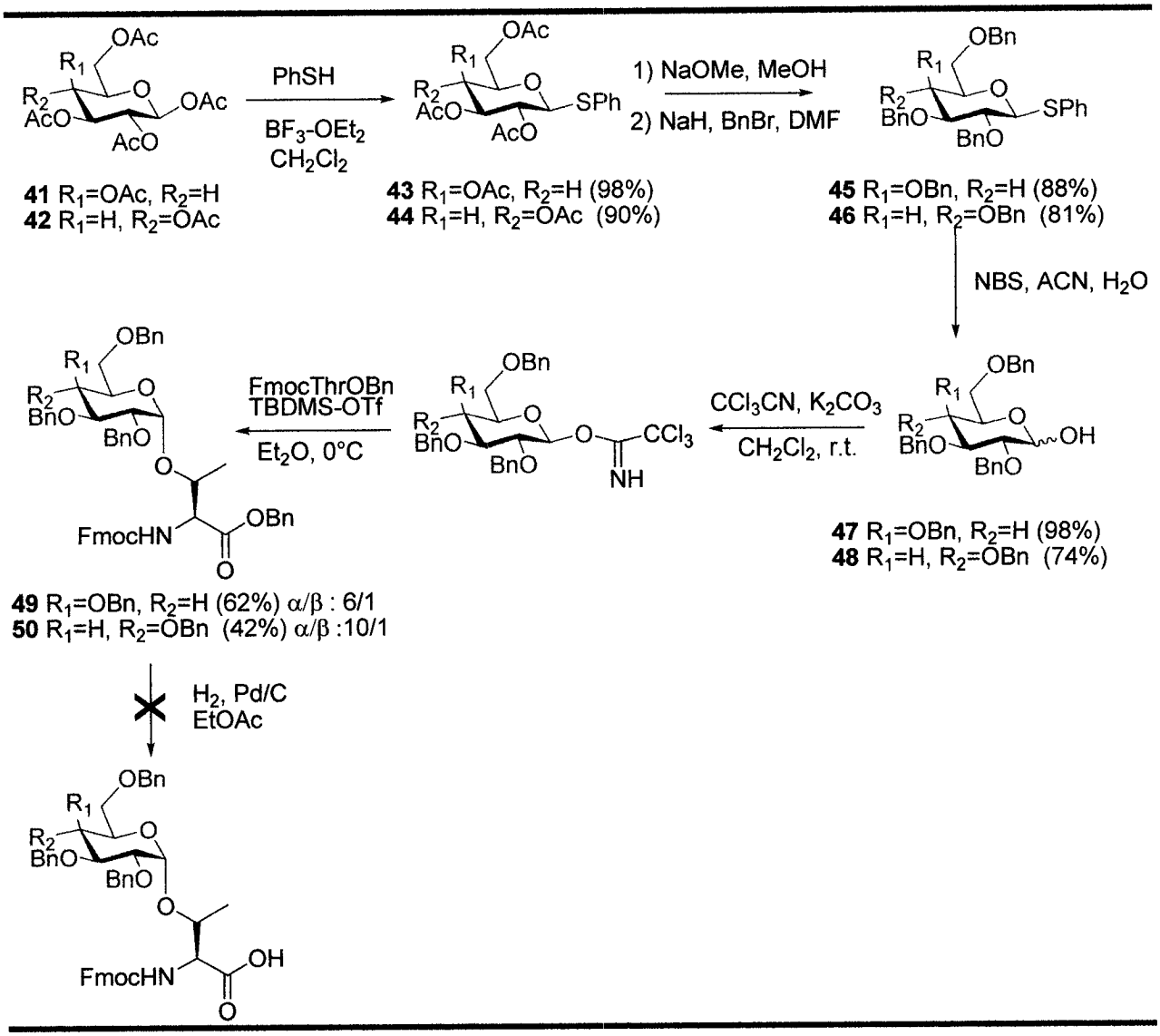
Table 1: Yields of the different approaches to obtain the building blocks

***O*-Glycosylated Threonine building block**	*Overall yields* **(Using pentafluorophenol ester) (3 steps)**	*Overall yields* **(Using benzyl ester) (5 steps)**
β-Galactose	**32 : 64%**	**38 : 61%**
β-Glucose	**33 : 60%**	**39 : 66%**
α-Mannose	**34 : 65%**	**40 : 68%**

4.2.3 Synthesis of α-galactose and α-glucose building blocks 51, 52 and 53

The use of a benzyl ester protected amino acid greatly simplified the synthesis of the previous analogues. Therefore in order to prepare the α-galactose and α-glucose building blocks we decided to use the same strategy. For the synthesis of the α building blocks, it was necessary to prepare a glycosyl donor without a participating group on the C-2 position. Benzyl ethers are frequently employed as protecting groups for this purpose. Utilizing benzyl ester as a protecting group should allow a selective deprotection to give the free amino acid. First the anomeric position was protected using thiophenol and afforded selectively the β-anomers of **43** and **44** starting from the pure β compounds **41** and **42** (Scheme 3).

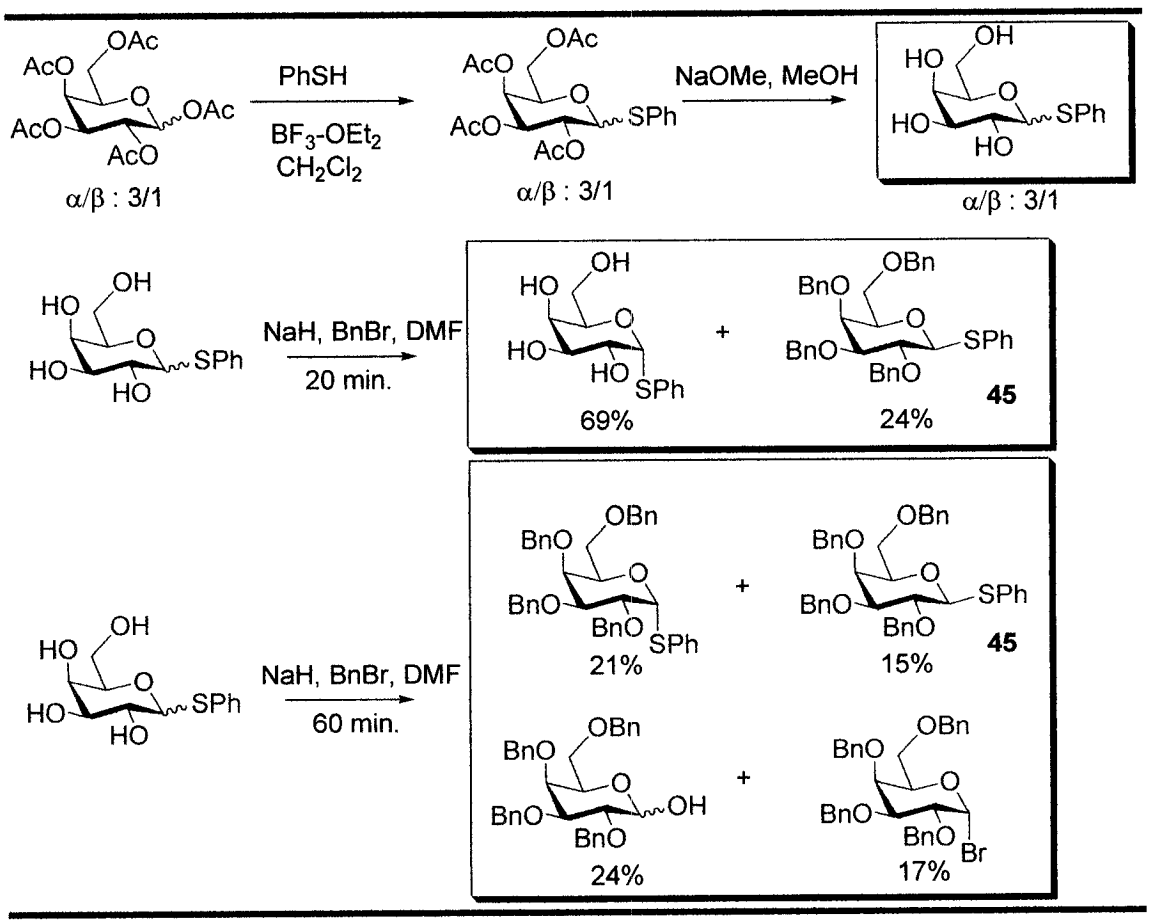
Scheme 3: Synthesis of α-glucose and α-galactose building blocks using benzyl ester amino acid



In this synthesis, it was easier to start with pure β-pentaacetylated glycoside. This was because during the benzylation, the α-thiophenol anomer reacted much slowly than the β anomer and degradation by-products appear if the reaction is left for a long time (Scheme 4). All the reactions were therefore executed on pure β-anomer compounds. The acetate protecting groups are then exchanged for benzyl ether groups to produce **45** and **46**. Once the non-participating protecting groups are installed, the anomeric position was deprotected to obtain an α/β mixture of **47** and **48**. The compounds were then ready to undergo the formation of the β-trichloroimidate, which allow the selective coupling of the benzylated

threonine.[18] The absence of C-2 participating group gives preferentially the α anomer (6/1 (α/β) for galactose and 10/1 for glucose), while using thermodynamic conditions. Compounds **49** and **50** were then obtained. The last necessary step prior to the solid phase synthesis is the debenzylation of compounds **49** and **50**. Several techniques have been published that describe selective deprotection of benzyl esters in the presence of benzyl ethers.[19] First, we decided to see whether regular catalytic hydrogenolysis would be selective and preserve the C-terminal benzyl ester over benzyl ether.

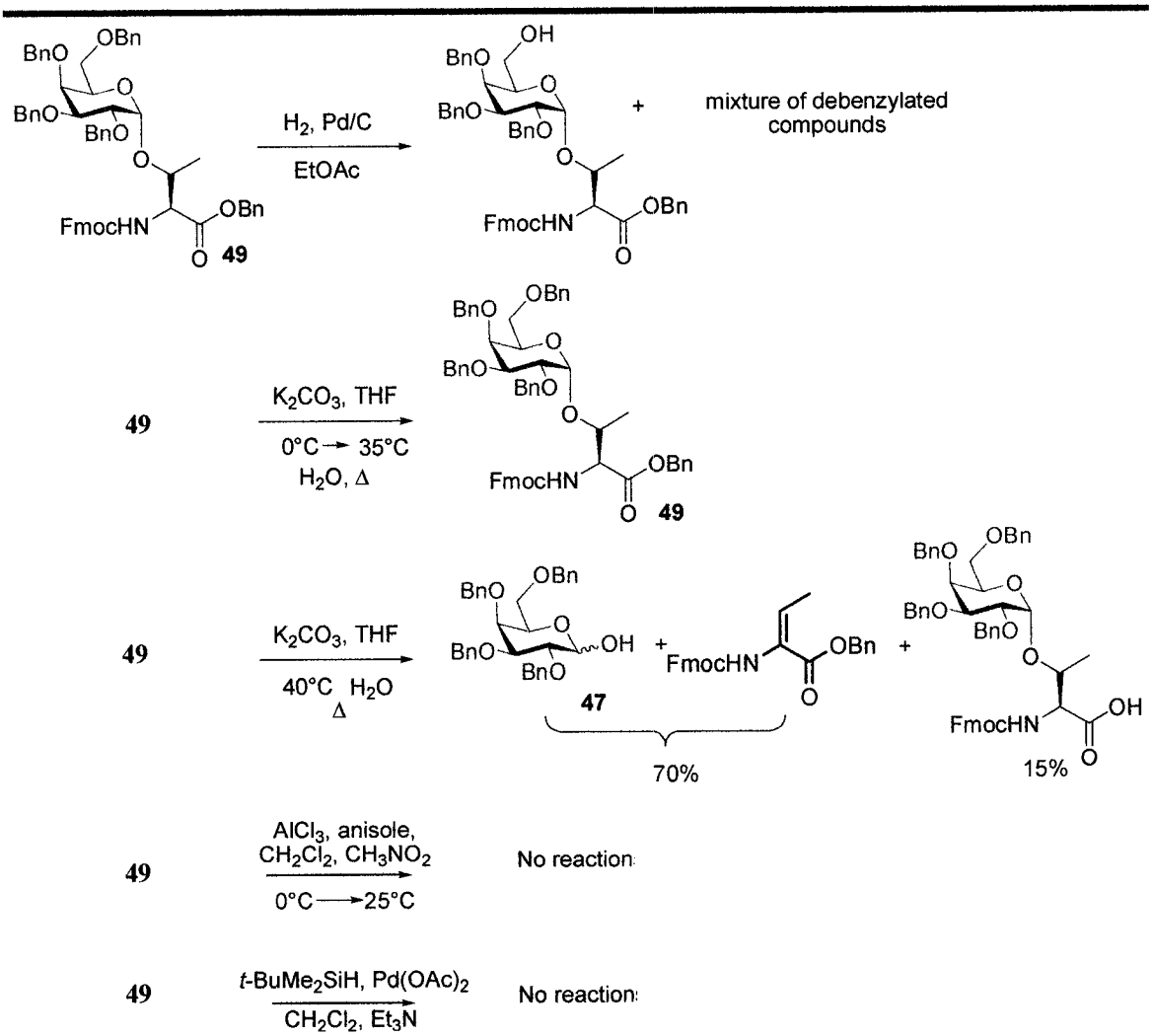
Scheme 4: Benzylation of the α/β mixture of thioglucose and thiogalactose



Unfortunately, the catalytic hydrogenolysis removed the C-6 benzyl ether group first, followed by the deprotection the other benzyl groups to afford a complex mixture of debenzylated products. Consequently this technique was not suitable for our purpose. We then tried to hydrolyze the benzyl ester selectively using potassium carbonate; however the glycosidic linkage appeared to be more susceptible to hydrolysis than the benzyl ester and

only a small amount of the desired compound was formed (Scheme 5). One of these promising techniques was to convert the benzyl ester into a silyl ester; however, although the conversion occurred with reasonable yields, the silyl ester was surprisingly difficult to hydrolyze and our substrate degraded before hydrolysis. All of these techniques having failed on our substrate, we decided to returne to our first strategy using the pentafluorophenol as protecting group (Scheme 6).

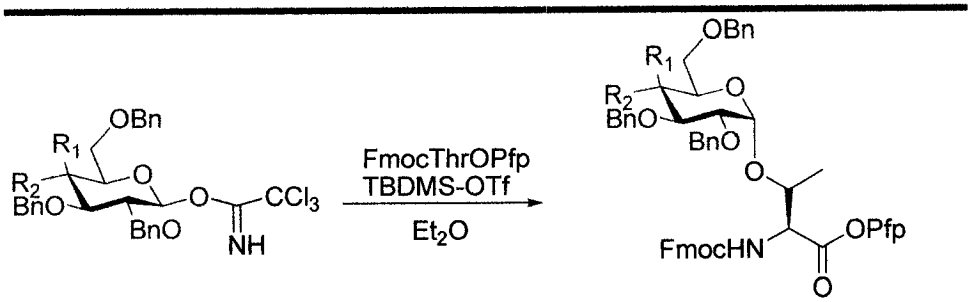
Scheme 5: Attempts at selective deprotection[19]



The coupling afforded the galactose building block **51** with 82% yield and with an α/β mixture of 6/1. It also afforded the glucose building block **52** with 49% yield and a 10/1 α/β ratio.

These building blocks were further purified to obtain pure α anomer building blocks using HPLC purification.

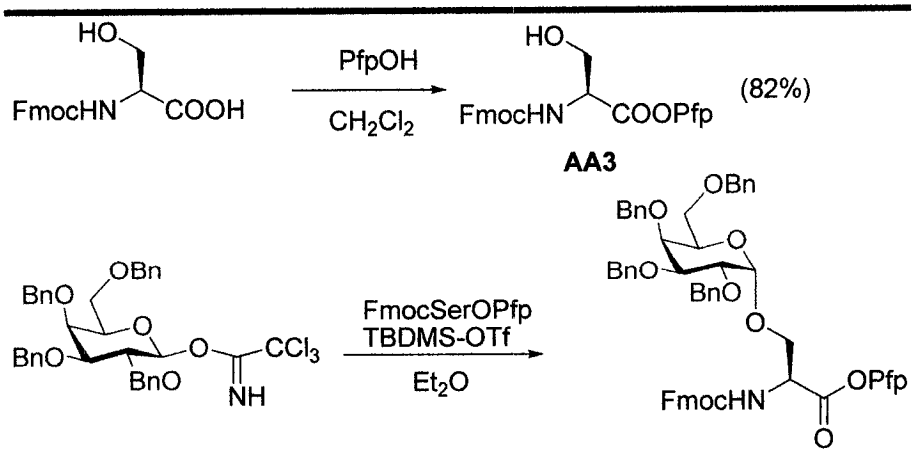
Scheme 6: Synthesis of α-glucose, α-galactose building blocks using pentafluorophenol protected amino acid



51 R_1=OBn, R_2=H (82%) α/β : 6/1
52 R_1=H, R_2=OBn (49%) α/β :10/1

The next analogue that needed to be synthesized was the α-galactose building block that is linked to serine. In order to obtain this analogue the same procedure as for the synthesis of **49** was followed (Scheme 7). The serine pentafluorophenol ester **AA3** was obtained from commercially available FmocSerOH in 82% yield. Compound **53** was obtained by coupling the previously synthesized trichloroimidate of the perbenzylated galactose with the protected serine. Compound **53** was obtained with 65% yield as an α/β mixture of 5/1. Further HLPC purification was necessary in order to obtain the pure building block.

Scheme 7: Synthesis of α-galactose building blocks linked to serine



53 (65%) α/β : 5/1

Table 2 describes the different yields and number of steps performed in order to obtain suitable building blocks to synthesize the AFGP8 analogues.

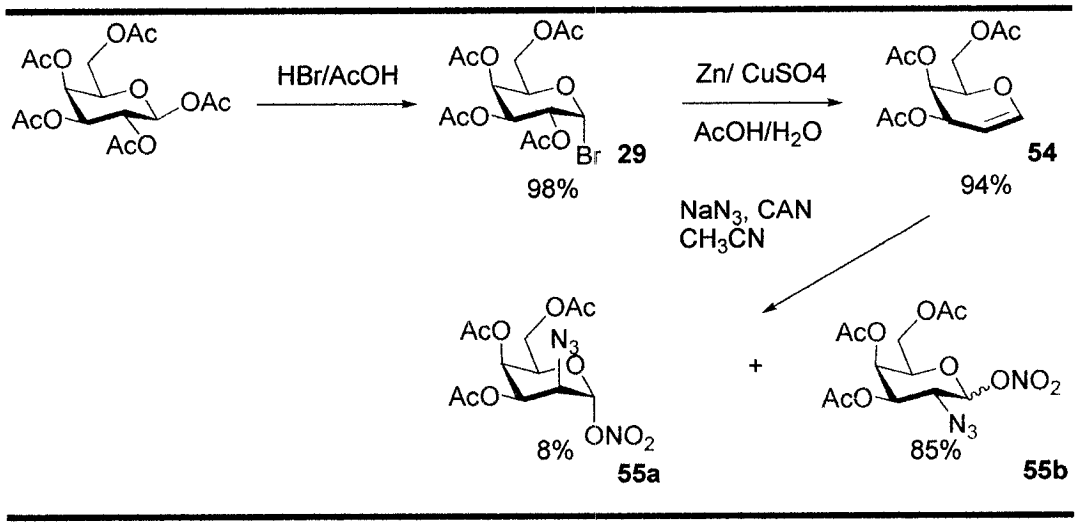
Table 2: Summary of the synthesis of compounds 51, 52 and 53

***O*-Glycosylated Threonine building block**	***Overall yields*** **(Using pentafluorophenol ester)** **(6 steps)**
α-Galactose (Threonine)	**51 : 64%**
α –Glucose (Threonine)	**52 : 60%**
α -Galactose (Serine)	**53 : 65%**

4.2.4 Synthesis of α-GalNAc building block 65

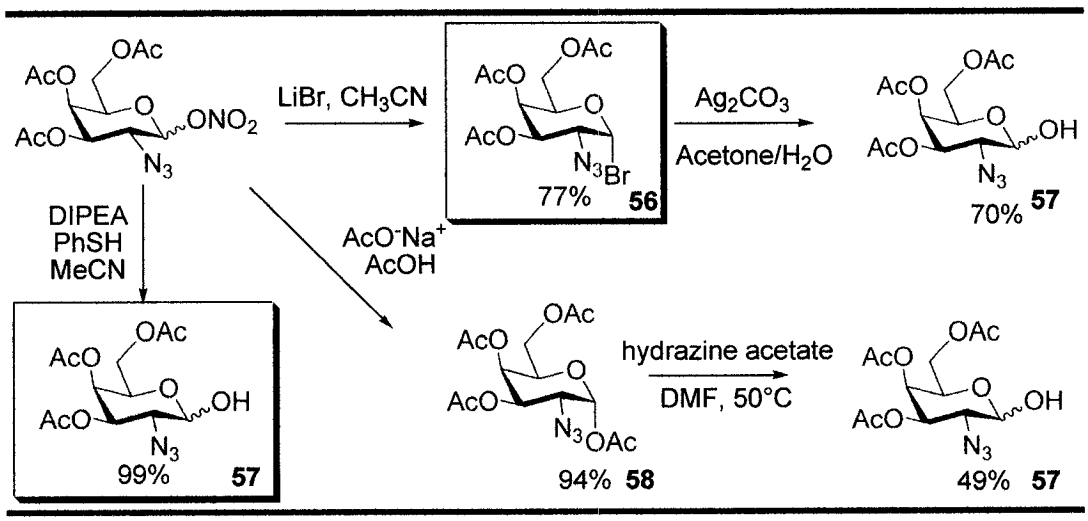
The last building block that we desired to synthesize is the α-*N*-acetyl-galactosamine linked to threonine. As previously mentioned, we needed to introduce a non-participating functional group in the C-2 position and this group needed to be convertible to an amine functionality later in the synthesis. In our case an azide functionality was the most appropriate group for this purpose. The bromide derivative **29** was synthesized and subsequently reduced to obtain galactal **54**. Working on a 2 gram scale or larger for the formation of the galactal makes the solubilization more difficult and the hydrolysed by-product appears. Once the tri-*O*-acetyl-D-galactal **54** was purified, it underwent an azidonitration (Scheme 8).[20]

Scheme 8: Bromination of galactose and azidonitration of galactal[20]



This reaction generates three different compounds, specifically the α and β galactose **55b** and the α-talose derivatives **55a**. The talose can be purified from the α/β mixture of azido galactose nitrate. This mixture can be converted to the bromide **56** or the galactol **57**.

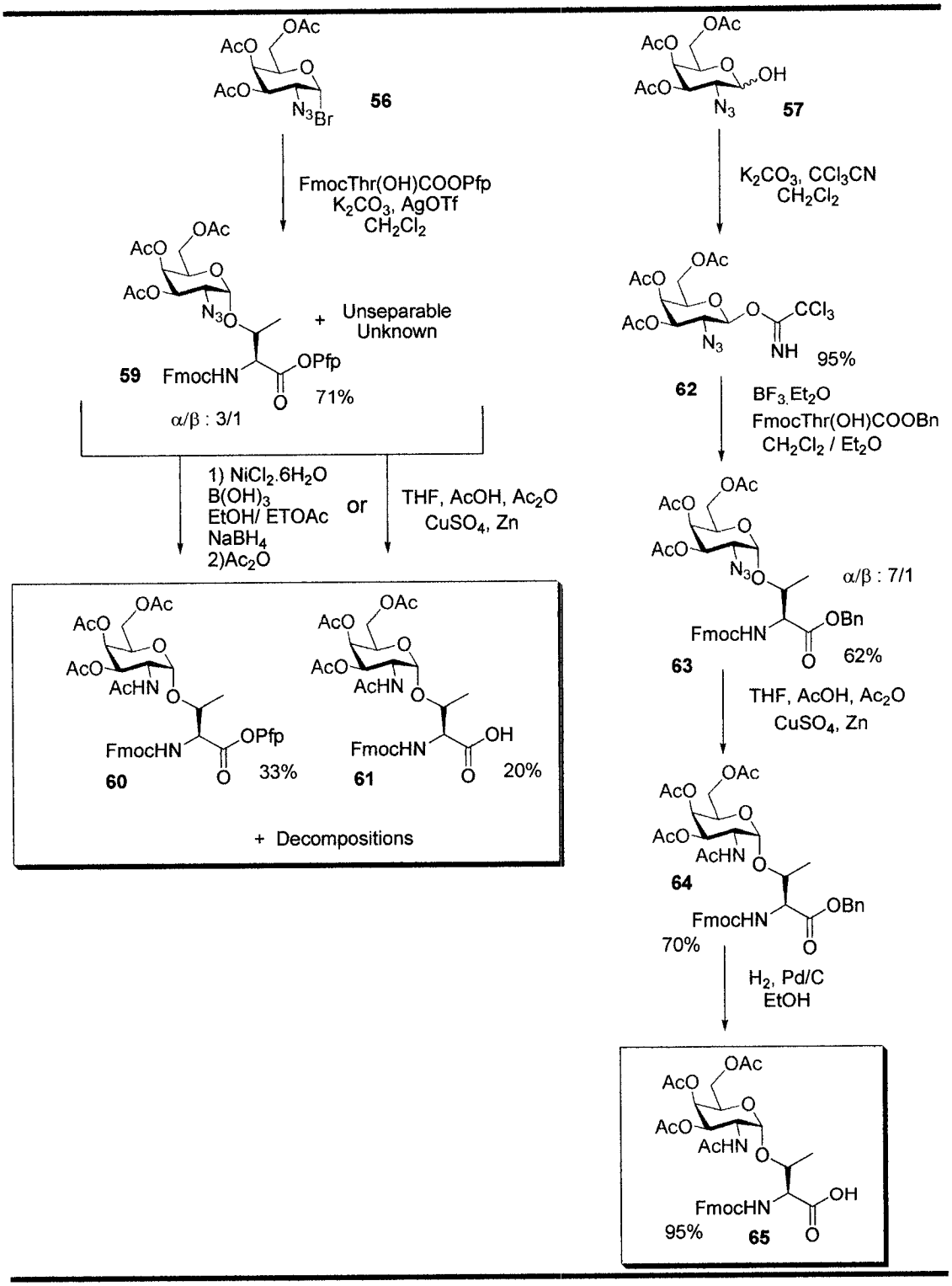
Scheme 9: Formation of the bromide and hydroxyl intermediate[20,21]



To afford the bromo glycoside **56**, the product of the azidonitration was treated with lithium bromide in acetonitrile. To obtain the hydroxyl glycoside **57**, several strategies were tested and are presented in Scheme 9, but the treatment of the nitro compound with thiophenol was the fastest and most efficient method.[21] Once compounds **56** and **57** were

obtained, we were ready to investigate the Koenig/Knorr and the trichlorimidate couplings (Scheme 10). We decided to first investigate the coupling of the bromo glycoside with the pentafluorophenol ester of the threonine. This coupling afforded three different compounds, the α-anomer of compound **59**, its β anomer (α/β ratio (3/1)) and an unknown compound with the same R_f as **59**. The two diastereomers were separated, however we were unable to separate **59** from the unknown compound. We then decided to continue the sequence with the product mixture. We were expecting that the reduction and the acetylation reactions would help us to separate these two compounds; however this was not the case. Although we tested several reduction and acetylation procedures, all of them were unsuccessful, giving several degradation products. The two best procedures are presented in Scheme 10.[16,22] The first procedure consists on a nickel chloride, sodium borohydride reduction of the azide function to afford the free amine followed by an acetylation; the second procedure involves a one step reduction-acetylation using copper sulfate, zinc, acetic acid and acetic anhydride. Although this procedure usually gives excellent yields with benzyl ester protected amino acid, these conditions for the reduction and the acetylation induced partial hydrolysis of the pentafluorophenol ester and afforded the hydrolysed compound **61**. After these problems we decided to investigate the trichloroimidate strategy using the benzyl ester threonine. This strategy afforded an improved α/β ratio (7/1) and no other by-products were observed.

Scheme 10: Different strategies to obtain the desired GalNAc building block.



Compound **63** was also stable to the reduction, acetylation and debenzylation conditions and the desired building block **65** was obtained with satisfactory yields and further purified by recrystallization in Et_2O.

Table 3 summarizes the number of steps and the overall yield to obtain the building block **65**.

Table 3: Overall summary of the synthesis of the building block

***O*-Glycosylated Threonine building block**	***Overall yields* (Using benzyl ester) (8 steps)**
α-*N*-acetyl-Galactosamine (Threonine)	**65 : 31%**

4.2.5 Solid Phase Peptide Synthesis

The use of 9-fluorenylmethoxycarbonyl (Fmoc) as an α-amino protecting group for SPPS was developed into a powerful strategy for the preparation of biologically active polypeptides and proteins.[23,24] The resin acts as the solid support for sequential coupling of *N*-protected amino acids. An acid-labile ester bond ties the *C*-terminal amino acid to the resin. This linkage is stable to reaction conditions for deprotecting the α-amino group and amide bond condensation, but can be cleaved by trifluoroacetic acid (TFA) to free the synthesized polypeptides from the solid resin. Amino acids to be assembled are temporarily protected with a base labile Fmoc group on the α-amino position, which is usually cleaved with 20% piperidine in *N*,*N*-dimethyl formamide (DMF) solution. After deprotecting the α-amino group, the next amino acid is coupled to the resin-bound peptide. In order to improve the efficiency of coupling reactions and avoid racemization, α-carboxyl groups are activated, either to form *N*-hydroxybenzotriazole (HOBt) *in situ* with activators such as 2-(1H-Benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HBTU) or to pre-form activated esters (OPfp), symmetrical anhydrides or acid halides. A new amide bond is formed between the activated amino acid and the resin-bound amino acid under basic conditions on the solid resin. The Fmoc deprotection and amide bond condensation are

repeated until the desired polypeptide is assembled. After coupling the *N*-terminal amino acid, the last Fmoc group is deprotected and the polypeptide is cleaved from the resin to afford the final product. Analogues are then ready for physical and biological testing (for detailed solid phase procedure see the experimental section).

The selection of resin is important for the success of SPPS. A good solid resin should have enough active sites for polypeptide propagation, a functional linker compatible with coupling and cleavage conditions used in the Fmoc strategy, and good swelling capability to make the functional sites in the polymer matrix accessible to the reagents in solution phase. Based on this, we selected the glycine preloaded Wang resin, which has been widely used as a solid support compatible with the Fmoc SPPS strategy. Wang resins are made of cross-linked polystyrene (PS) beads, which are functionalized with an acid labile *p*-benzyloxyl benzylalcohol handle to load the first amino acid at the C-terminus through an ester bond. TFA cleavage of the ester bond is facilitated by the resonance stabilization of the resulting carbocation by *p*-benzyloxy phenyl group. The Wang resin used in this study has loading capacities of 0.6-0.65 mmol/g and a good swelling capacity of three-fold in volume in DMF. *O*-Linked AFGP analogues **70-73** were prepared by manual sequential assembly of building blocks **51**, **52** and **53** with *N*-Fmoc alanine or glycine (Scheme 11). A 20% piperidine in DMF solution was used to deprotect the Fmoc group at α-amino positions. HBTU was not necessary for the couplings, indeed the pentafluorophenol acted as an *in situ* activating reagent for amide bond coupling in the presence of diisopropylethylamine (DIPEA). The coupling reactions were repeated twelve times with either *N*-Fmoc alanine, glycine or the building blocks **51**, **52**, or **53** to afford the resin-bound polypeptides. To ensure the completion of reactions, each Fmoc deprotection and amino acid coupling step was monitored by both Ninhydrin (Kaiser) and 2,4,6-trinitrobenzenesulfonic acid (TNBS) tests to detect the existence of free amino groups (Figure 11). A positive test reveals the presence of free amines and therefore that the *N*-Fmoc protecting group in not present (either the coupling is not completed or the *N*-Fmoc deprotection is proceeding). A negative test means that the *N*-Fmoc protecting group is present and that therefore there is no free amine present (either the coupling is completed or the deprotection failed).

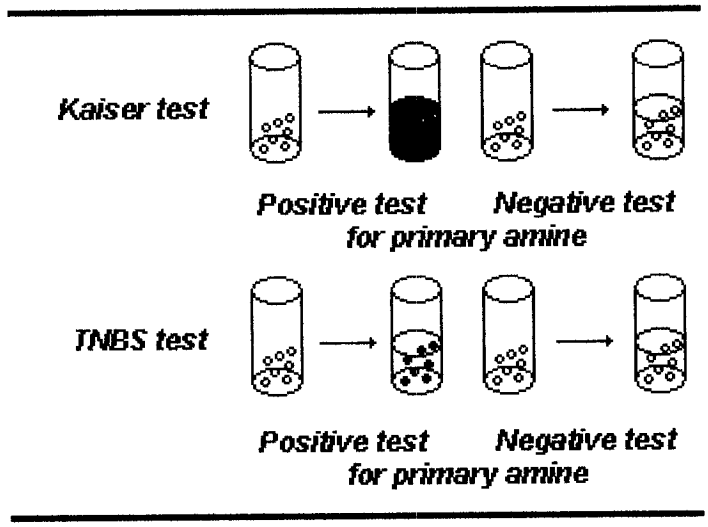
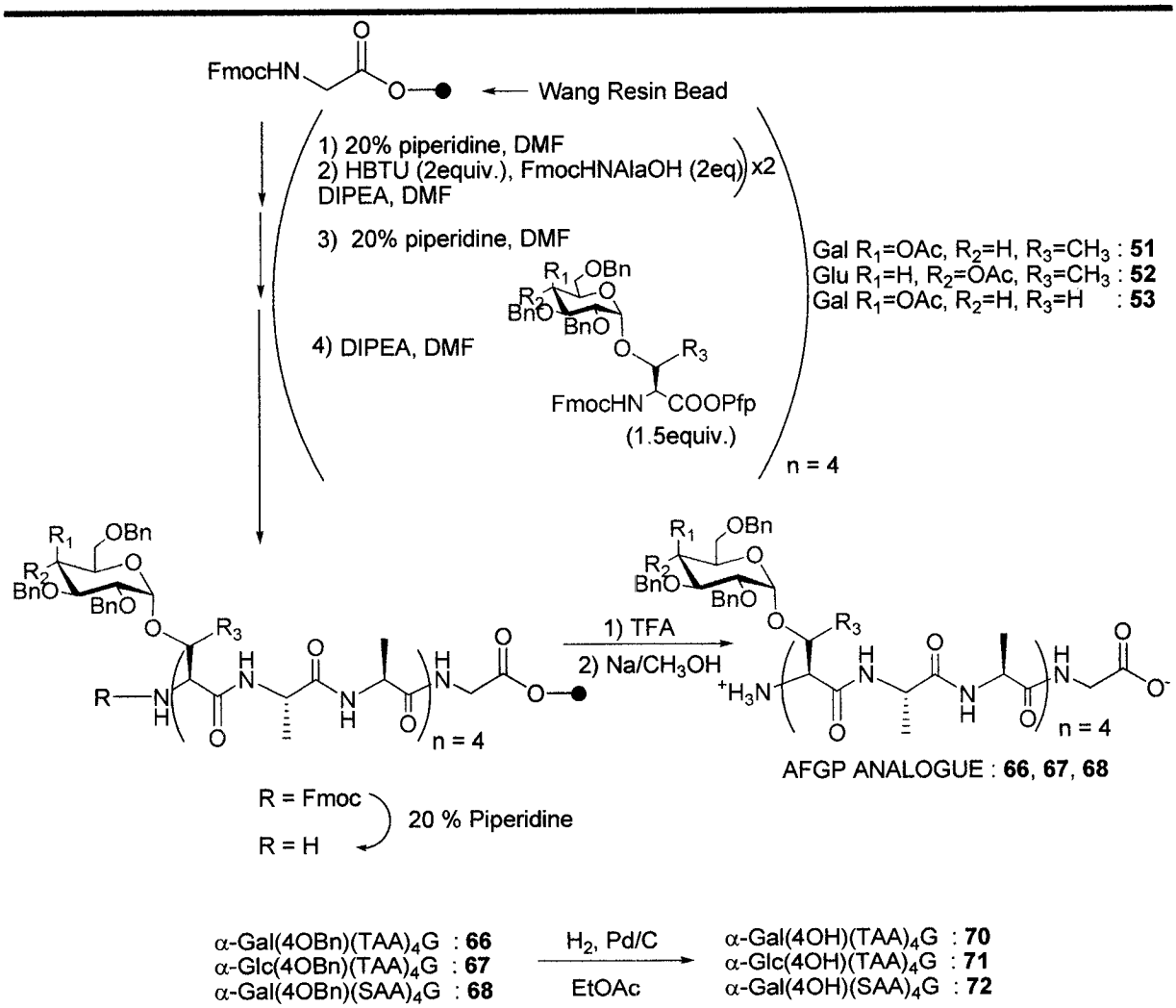


Figure 11: Colorimetric tests for primary amine detection on solid phase resin

After the last building blocks were assembled, the *N*-Fmoc groups were removed and the polypeptides were cleaved from the resin with 50% TFA in dichloromethane solution to furnish benzyl protected analogues **66-69.** Before further modifications these analogues were purified by reverse phase HPLC. Finally, polypeptides **66-69** were treated with H_2 and 5% Pd/C to afford final *O*-linked AFGP analogues **70-73** (Figure 10). The polypeptides were then precipitated out from reaction mixtures with diethyl ether, re-dissolved in water and lyophilized in order to remove any traces of TFA. Purification of the AFGP analogues was performed with reverse phase HPLC and the identity of the analogues were verified with mass spectrometry. Compound **69** never precipitated from the mother solution and therefore underwent the debenzylation as a crude mixture to afford compound **73**. HLPC tracing as well as mass spectrometry were not satisfactory for this compound. Therefore the compound would need to be synthesized again. For this synthesis the building block **53** was not freshly prepared. The compound was stored under argon at 5°C for several months and even though no degradation products were observed by TLC, it is still possible that some were present and perturbed the synthesis.

Scheme 11: Solid phase synthesis using pentafluorophenol ester



O-Linked AFGP analogues **78-81** (Figure 10) were prepared by manual sequential assembly of building blocks **38-40** and **65** with *N*-Fmoc alanine (Scheme 12). A 20% piperidine in DMF solution was used to deprotect the Fmoc group at α-amino positions. Here, HBTU was necessary for the couplings to generate the *in situ* activating reagent for amide bond coupling in the presence of diisopropylethylamine (DIPEA). The coupling reactions were repeated twelve times with either *N*-Fmoc alanine or the building blocks **38-40**, or **65** to afford the resin-bound polypeptides. After the last building blocks were assembled, the *N*-Fmoc groups were removed and the polypeptides were cleaved from the resin with 50% TFA in dichloromethane solution to furnish acetate protected analogues **74-**

77 (for detailed solid phase procedure see experimental section). Polypeptides were then precipitated out from reaction mixtures with diethyl ether, re-dissolved in water and lyophilized in order to eliminate any traces of TFA. Before further modifications these analogues were purified by reverse phase HPLC. To afford final *O*-linked AFGP analogues **78-81**, polypeptides **74-77** were treated with a basic solution of sodium methoxide (pH > 10) with one drop of water in order to avoid the methylation of the *C*-terminal acid, followed by neutralization with Amberlite IR-120H resin. Purification of our AFGP analogues was performed with reverse phase HPLC and the identity of the analogues were verified using mass spectrometry.

Scheme 12: Solid phase synthesis using benzyl ester

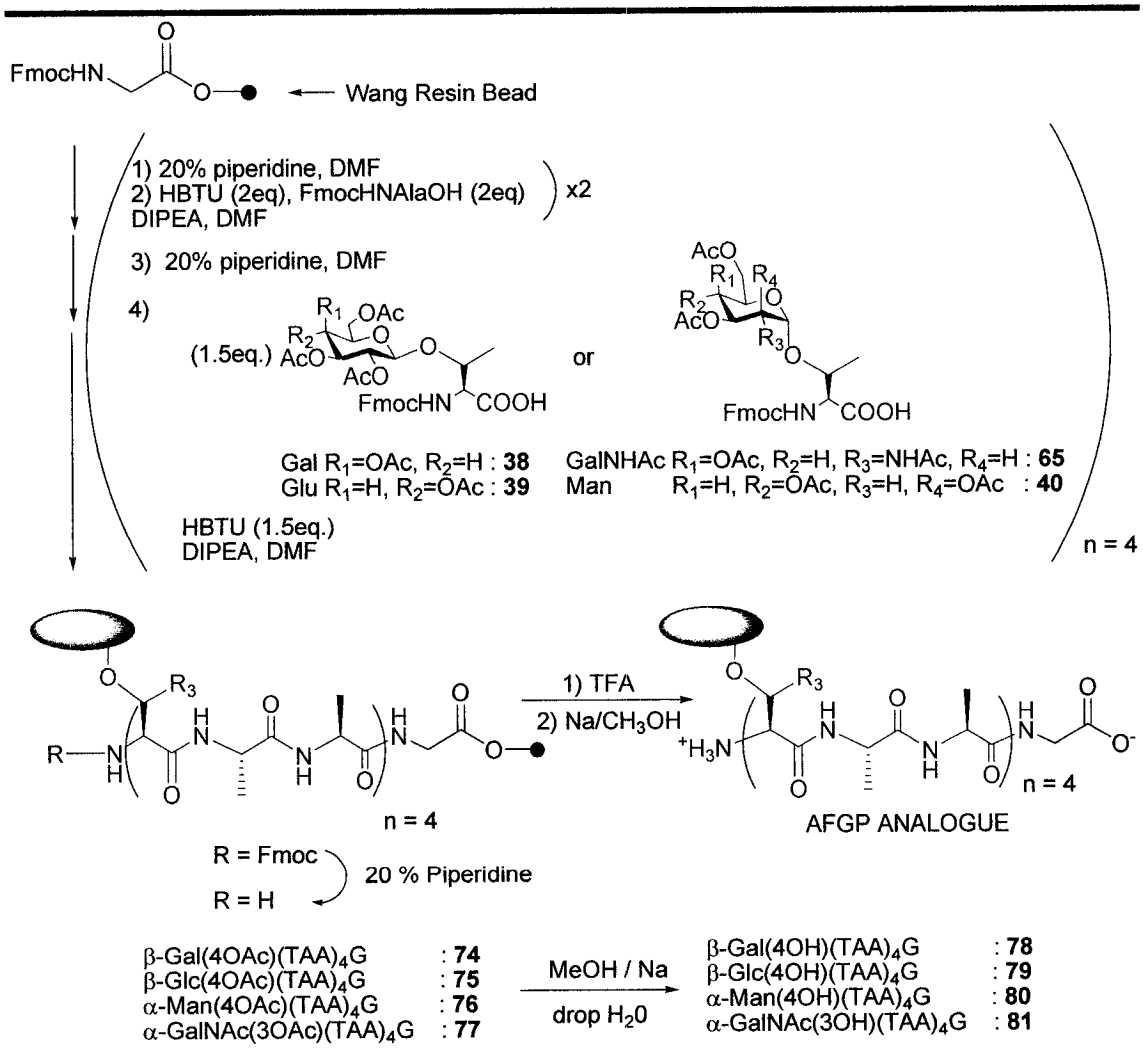


Table 4 summarizes the yield of the solid phase synthesis and the overall yields of all our *O*-linked AFGP8 analogues.

Table 4: Overall yields of all syntheses of the different O-linked analogues

***O*-Glycosylated Threonine Derivatives**	*Yields* **Solid Phase synthesis**	*Overall yields* **(From pyranose pentaacetate)**
β-Galactose-$(TAA)_4G$: **78**	**69%**	**42%**
β-Glucose-$(TAA)_4G$: **79**	**72%**	**47%**
α-Glucose-$(TAA)_4G$: **71**	**66%**	**16%**
α-Galactose-$(TAA)_4G$: **70**	**58%**	**33%**
α-Mannose-$(TAA)_4G$: **80**	**65%**	**44%**
α-GalNAc-(TAA)4G: **81**	**80%**	**25%**
α-Galactose-$(SAA)_4G$: **72**	**49%**	**22%**

4.3 Physical properties and biological activities of *O*-linked AFGP8 analogues

Thermal hysteresis (TH) and recrystallization inhibition (RI) are the two distinct properties that distinguish biological antifreezes from colligatively acting cryoprotectants. TH is a non-colligative phenomenon that results in a freezing point depression below that of the equilibrium melting point. Alternatively, RI shows protecting effects to cells by prohibiting small ice crystals from growing into large harmful ice grains. Both TH and RI activities are believed to be caused by the direct binding of antifreeze molecules to ice

crystals. Conformational studies of biological antifreezes are important for us to understand whether antifreeze molecules need to adopt certain conformations in order to form effective interactions with ice lattices. Circular dichroism (CD) has been widely used to measure the conformation of biological macromolecules. Therefore, in this study, CD spectroscopy was used to characterize the conformations of *O*-linked AFGP analogues.

4.3.1 Assessing Thermal Hysteresis (TH) and Dynamic Ice-Shaping (DIS) Activities

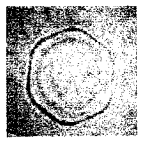
4.3.1.1 The thermal hysteresis (TH) assay

Nanoliter osmometry is the golden standard used to assess thermal hysteresis (Section 3.4.1).[25,26] In addition to the TH activity, which is the quantitative indicator for the effectiveness of freezing point depression,[27,28] biological antifreezes can also modify the morphology of ice crystals and change the dynamics of the ice growing process.[29] A single ice crystal is used as the seed crystal in the nanoliter osmometry assay. This is important for studying the morphological and dynamic changes of the seed crystal at temperatures around the TH gap. During the cooling process of a colligative solution, solutes are excluded from the ice crystal and remain in the solution phase.[30,31] Therefore, the ice crystal grows independently without interaction with the colligative solutes and forms a circular disk. Dynamically, the growth of the ice crystal initiates from the beginning of the cooling process and proceeds gradually until the entire solution is frozen. In contrast, biological antifreezes can selectively bind to a specific prism or secondary prism faces of the ice crystal and force ice to grow only on the basal planes, which results in formation of the characteristic bipyramidal structures. This unique ability is known as dynamic ice-shaping (DIS).[32]

4.3.1.2 Testing O-linked AFGP analogues for thermal hysteresis (TH) and dynamic ice shaping (DIS) activities

In these studies, TH and dynamic ice-shaping activities of *O*-linked AFGP analogues **78-81** and **70-72** were studied using a Clifton nanoliter osmometer and the results are shown in Table 5.

Table 5: Dynamic ice shaping and thermal hysteresis of our O-linked analogues

Analogues		Ice shaping	TH
Thr-Ala-Ala α (Gal-GalNAc) 4-mer	**AFGP8**		-0.12°C
Thr-Ala-Ala β (Gal) 4-mer	: **78**	No ice shaping	No TH
Thr-Ala-Ala β (Glc) 4-mer	: **79**	No ice shaping	No TH
Thr-Ala-Ala α (Man) 4-mer	: **80**		No TH
Thr-Ala-Ala α (Gal) 4-mer	: **70**		No TH
Thr-Ala-Ala α (Glc) 4-mer	: **71**	No ice shaping	No TH
Thr-Ala-Ala α (GalNAc) 4-mer	: **81**	No ice shaping	No TH
Ser-Ala-Ala α (Gal) 4-mer	: **72**	No ice shaping	No TH

a) Study of the importance of the (1, 3)-β-Gal-GalNAc disaccharide unit versus the monosaccharide moiety

Toward this goal, the data from the α-GalNAc analogue **81**, which can be considered as a monosaccharide structural mimic of (1, 3)-β-Gal-GalNAc (AFGP8), were analysed. This analogue did not display any TH or DIS activity while AFGP8 displayed a TH gap of 0.12°C. This result confirms that the terminal galactose is essential for TH activity. This

might be due to the fact that the number of hydroxyl groups present on the molecules can result in increased favourable interactions with the ice lattice. The interactions of the glycopeptide analogues and the ice surface are regulated by strong hydrophilic/hydrophobic interactions; consequently it is expected that any changes in the number of hydroxyl groups on the glycopeptide analogues will influence any activities that would involve interaction with the ice lattice.

b) Determining the effect of stereochemistry of the glycosidic linkage on TH activity (i.e. α or β linkages)

Once it was determined that the terminal galactose was essential for TH activity, it was necessary to verify whether an α or β anomeric configuration was necessary for our monosaccharide analogues. In the native system two different linkages are present, an α-linkage between the carbohydrate moiety and a β-linkage between the two saccharides. As a result, the α-galactose and β-galactose analogues **70** and **78** were tested. It was important to test these galactose analogues because of the nature of the native disaccharide. Indeed compound **70** α-linked to the amino acid, mimicked the α-linkage between the disaccharide and the amino acid of AFGP8, while compound **78** mimicked the interglycosidic β-linkage with its β-linkage with the amino acid. None of the analogues exhibit any TH activity; however the α-galactose analogue **70** demonstrates some limited DIS activity. The fact that we do not observe any TH activity for any of our analogues and that only the α-analogues **70** and **80** exhibit limited DIS activity, suggests that the disaccharide moiety is a key structural feature for TH activity in our 4-mer analogues.

The fact that α-analogues galactose **70** and mannose **80** display some DIS activity implies that they interact with the ice lattice with a higher degree than any other of tested analogues. Several β-analogues were tested and none of them demonstrate any degree of ice interaction as observed by DIS. Therefore we can conclude that α linkages between the carbohydrate and the peptide side chain is more favourable for generating interaction with the ice lattice. This is especially confirmed when the DIS activity of the α-galactose analogue **70** is directly compared to its β-counterpart **78**. The α-linkages as compared to the β-linkages make the carbohydrates sit "flat" on the peptide backbone (Figure 12). This relative orientation of the carbohydrate one to another might favour ice interactions and explain why β-analogues lack any antifreeze activities.

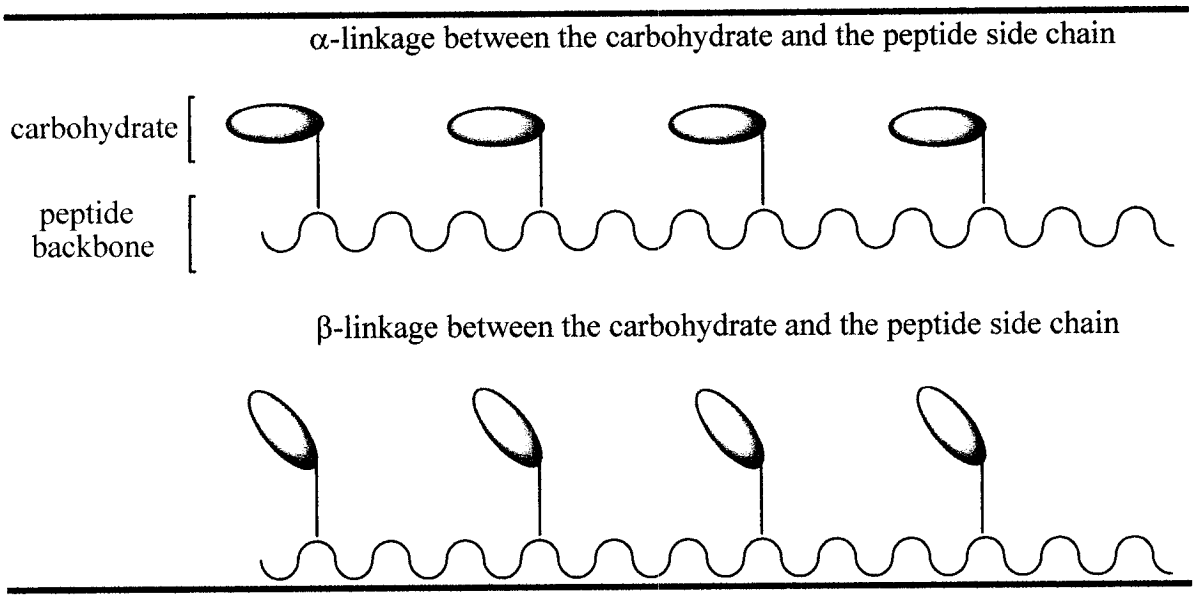


Figure 12: orientation of the carbohydrate depending of the anomeric linkage

c) *Influence of relative stereochemistry of the hydroxyl groups in the pyranose ring of the hydroxyl groups on TH activity*

Interestingly, the α-galactose **70** and the α-mannose **80** analogues display a noticeable DIS activity while the α-glucose analogue **71** does not. At first glance, there is no an obvious reason for this, however the result suggests that the relative stereochemistry of the hydroxyl groups of the carbohydrate moiety is important for the interaction of the AFGP with the ice lattice. It has been reported that the degree of hydratation of a solvated carbohydrate depends on the stereochemistry of the hydroxyl groups in the carbohydrate due to their ability to "fit" into the three-dimensional hydrogen-bonded structure of water. Engberts *et al.*[33] postulated that galactose can disturb the hydration layer more than glucose or mannose. It is possible that the ability to interact with the ice lattice is somehow correlated with the ability of the carbohydrate moiety to disturb the hydration layer.

d) *Comparing the length of the glycopeptide analogues and their TH activity*

One plausible explanation for the lack of TH activity as well as for the very weak display of DIS activity of our analogues is that we might be at the threshold for TH detection with our 4-mer analogues. This hypothesis is confirmed by our 4-mer α-GalNAc analogue **81** which does not display any TH or DIS activity whereas the 11-mer α-GalNAc analogue **86** synthesized by Nishimura[34] is highly potent (Table 6). Consequently this information

proves that length of the analogues influences TH activity. This trend has also been observed in several of our *C*-linked analogues while only $[LGG(Gal)]_6$ and $[LGG(Gal)]_9$ exhibited TH activity where $[LGG(Gal)]_3$ does not (Graph 5 in Chapter 1).

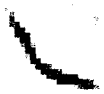
e) *The role of the N-acetyl group (C-2 position) on TH activity*

No TH activity was observed for either the α-GalNAc 4-mer analogue **81** and or the α-Gal 4-mer analogues **70**. Moreover no DIS activity was observed for the α-GalNAc analogue **81**. Consequently, it would seem that the presence of an *N*-acetyl in the C-2 position is detrimental for ice interaction in our 4-mer monosaccharide analogues. This information at first seems contradictory to the results obtained by Nishimura with his 11-mer analogues. Nishimuras' α-GalNAc 11-mer analogue **86** is as active as AFGP6 for TH activity and α-Gal 11-mer analogue **88** only exhibit weak DIS activity. These data suggest that for his 11-mer analogues the C-2 *N*-acetyl function is essential for TH activity. An explanation for these contradictory results can be explained by the fact that we are probably at the threshold for detection of TH and DIS activity with our 4-mer analogues. Therefore the *N*-acetyl functionality in the C-2 position is essential for TH activity in 11-mer analogues and irrelevant for our 4-mer analogues.

f) *The role of the γ-methyl (on the threonine residue) on TH activity*

Neither of our 4-mer analogues α-Gal(thr) **70** nor α-Gal(ser) **72** exhibit any TH activity, however α-Gal(thr) **70** displays a weak DIS activity which is not observed in α-Gal(ser) **72**. This indicates that the γ-methyl on the threonine is somehow important for ice interaction in our monosaccharide 4-mer analogues. The importance of the γ-methyl on the threonine on TH activity was also confirmed in Nishimura's study.[34] Indeed, when the threonine of analogue **82** (TAA [α (βGal(1-3)GalNAc]11-mer) is substituted by serine in the analogue **85** (SAA [α (βGal(1-3)GalNAc]11-mer) (Table 6) in the native system all the TH and the DIS activity is lost. Therefore the γ-methyl on the threonine is essential for DIS and TH activity on 4-mer and 11-mer analogues as well as on monosaccharide and disaccharide analogues.

Table 6: Dynamic ice shaping and thermal hysteresis of Nishimura's O-linked analogues[34]

Analogues		Ice shaping	TH
Thr-Ala-Ala α (βGal(1-3)GalNAc) 11-mer :	**82**		-0.2°C
Thr-Ala-Ala β (Gal-GalNAc) 11-mer :	**83**		No TH
Thr-Ala-Ala α (Gal-Gal) 11-mer :	**84**	No ice shaping	No TH
Ser-Ala-Ala α (Gal-GalNAc) 11-mer :	**85**	No ice shaping	No TH
Thr-Ala-Ala α (GalNAc) 11-mer :	**86**		-0.09°C
Thr-Ala-Ala α (βGal(1-4)GalNAc) 11-mer :	**87**		-0.03°C
Thr-Ala-Ala α (Gal) 11-mer :	**88**		No TH

4.3.2 Assessing Recrystallization Inhibition (RI) Activity

4.3.2.1 Recrystallization inhibition (RI) assay

Recrystallization inhibition (RI) is another property of biological antifreezes. RI protects organisms in partially frozen environments using a different mechanism from TH. Ice recrystallization transforms small ice grains to larger grains that can physically deform or rupture cell membranes, causing mechanical injury to cells and tissues.[35] As described

previously, recrystallization is the result of the boundary migration of ice grains with different crystallographic orientations. The driving force of this spontaneous process is to lower the overall interfacial energy of the system by decreasing the overall number of ice grains. The mechanism by which biological antifreezes inhibit ice recrystallization has been thoroughly studied by Knight.[36,37,38] Biological antifreezes can effectively bind to the ice surface at the ice solution interfaces between the neighbouring grain boundaries and immobilize the migration of ice grains, thus preventing them from merging with each other to form larger ice crystals.

The RI activities of *O*-linked AFGP analogues were measured using a recrystallization inhibition assay developed by Myers and Horwath.[39,40] A "splat cooling" method was used to produce mono-layer ice samples for the RI activity study.[37-41] In this assay, a 10 μL analogue solution in PBS was dropped onto a polished aluminum block which was pre-cooled to -80°C on dry ice. With this method, the analogue solutions formed a thin ice wafer composed of extremely tiny ice grains. The wafer was then moved onto a microscope cover glass and quickly transferred into a temperature-controlled Pelletier unit, which was dried with continuous airflow to prevent moisture from condensing on the ice wafer. The sample was then annealed at -6°C for 30 minutes and the size of the ice grains was photographed through a microscope. Ice recrystallization occurs during this annealing process. The size of the ice grains is dependent on the sample's recrystallization inhibition ability. For each sample, six photographs were taken of different areas of the ice wafer. Ice grains were regarded as being two-dimensional[36] and the size of the ice grains was calculated by taking the average surface area of the ten largest ice grains in each photograph (Mean Largest Grain Size (MLGS)). The surface of each ice grain was approximated as an elliptical area (Figure 13). The major and minor elliptical axes were defined by the two largest dimensions orthogonally across the ice grain surface. The surface area of each ice grain was calculated based on the formula: $A = \pi ab$, in which A represented area; a and b represented the radii of the major and minor elliptical axes.

Figure 13: Illustration of calculating surface areas of ice grains.

Although biological antifreezes depress freezing points in a non-colligative manner, it is known that the RI effects of biological antifreezes are still somewhat concentration-dependent.[39] Figure 14 shows the results of the RI assay of native AFGP8 with different concentrations in PBS. Given this, the RI activity of AFGP8 dramatically increases as a function of concentration. Compared to PBS, AFGP8 demonstrated obvious RI activity at concentrations as low as 10^{-8} M. When the concentration of AFGP8 was increased to be greater than 10^{-5} M, the ice grains became smaller than the thickness of the ice wafer (about 20 μm)[36] even after 30 minutes annealing, which caused difficulty in the resolution of the ice grains on photographs.

The non-specific RI effect of regular proteins is a result of their ability to inhibit the mobility of the solid/liquid boundary. This non-specific RI activity is usually weak and does not increase dramatically as the concentration increases. Therefore, we can differentiate these regular proteins from RI-active proteins by using three different concentrations to assess RI activity. Studies on first generation AFGP analogues, including control experiments with glycosylated bovine serum albumin (BSA) and peptide backbone controls, have demonstrated that the RI effects observed for our analogues are due to antifreeze protein-specific activity. This study used PBS solution as a control for the mean largest grain size (MLGS) of ice crystals formed. The MLGS for ice crystals observed with PBS solution is ~0.0126 ± 0.002 mm^2 (SEM = standard error of the mean).

The three following concentrations, $5.5X10^{-6}$ M, $5.5X10^{-8}$ M and $5.5X10^{-11}$ M, were selected for measuring the RI activity of each of our *O*-linked AFGP analogues and compared to that of native AFGP8.[41] For all our analogues the SEM was calculated to

determine if the difference in RI activities between our analogues and the controls were legitimate.

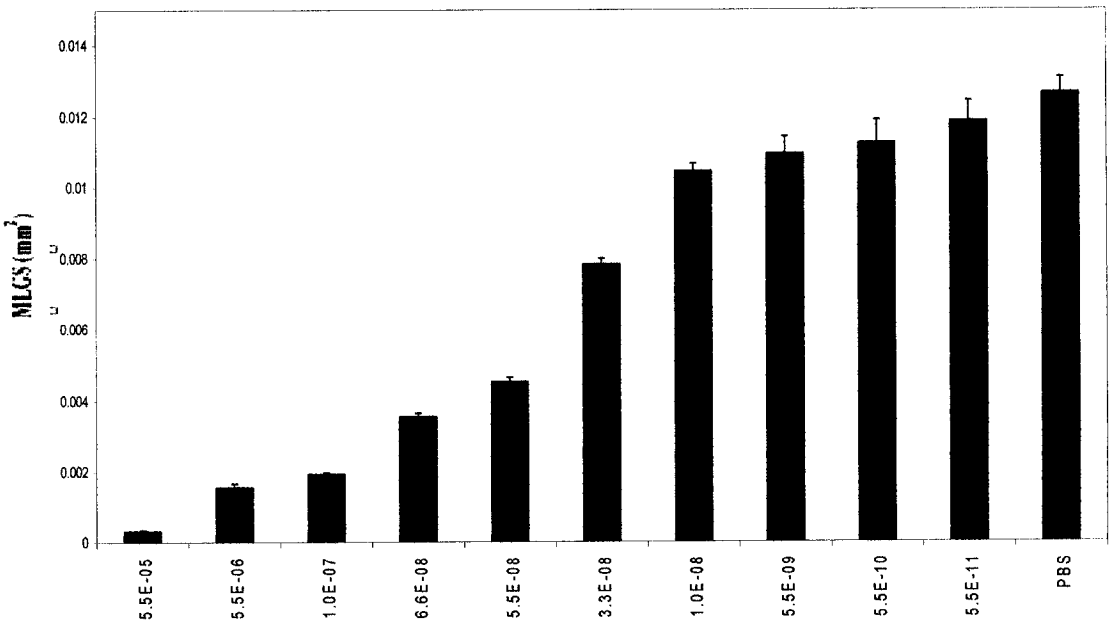


Figure 14: Recrystallization inhibition assay of AFGP8.[37]

It is worth noting that the effective concentrations for imparting RI activity (10^{-8} M, or $2.6X10^{-5}$ mg/mL for AFGP8) is much lower than that required for imparting TH activity (usually 10 mg/mL). This is probably due to the fact that the essential criteria for each activity are different. Some of the preliminary data suggest TH and RI activity are not associated and that a compound can exhibit one property without the other. Therefore, the fact that RI activity can be observed at lower concentrations than TH may be due to a difference in mechanism.

4.3.2.2 Recrystallization inhibition (RI) results for O-linked AFGP analogues

The RI assay results of our *O*-linked galactosyl AFGP analogues **70-72** and **78-81** are shown in Figure 15 and Table 7.

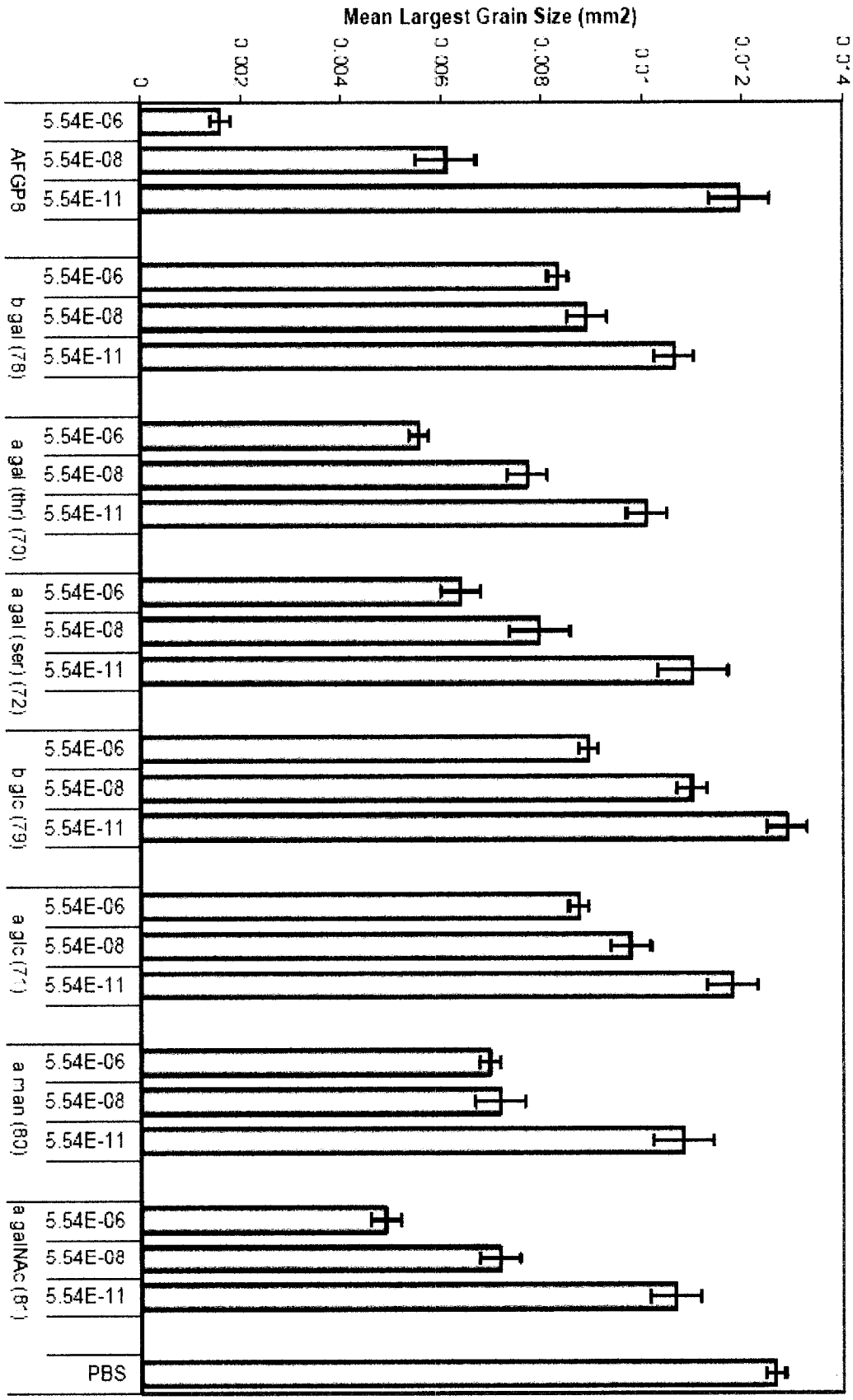


Figure 15: RI data for O-linked analogues

Table 7: MLGS of AFGP8, our analogues and PBS (surface area unit: mm^2).

Conc. (M)	AFGP8 ± SEM	α-Gal (thr) 70	α-Gal (ser) 72
5.54E-06	**0.0016±3.9e^{-5}**	**0.0056±0.0002**	**0.0064±0.0004**
5.54E-08	0.0061±0.0006	0.0077±0.0004	0.0080±0.0006
5.54E-11	0.0119±0.0006	0.0101±0.0004	0.0110±0.0007

Conc. (M)	α-Glc 71	α-Man 80	α-GalNAc 81
5.54E-06	**0.0087±0.0002**	**0.0070±0.0002**	**0.0049±0.0003**
5.54E-08	0.0098±0.0004	0.0072±0.0005	0.0071±0.0004
5.54E-11	0.0118±0.0005	0.0108±0.0006	0.0107±0.0005

Conc. (M)	β-Gal 78	β-Glc 79	PBS
5.54E-06	**0.0083±0.0002**	**0.0089±0.0002**	**0.0127±0.0002**
5.54E-08	0.0089±0.0004	0.0110±0.0003	
5.54E-11	0.0106±0.0004	0.0129±0.0004	

Although no TH and only weak DIS activities were detected, all of our analogues demonstrated at least a minimal RI activity compared to PBS solution, with the α-GalNAc analogue **81** being the most effective analogue. At a concentration of 5.5X10^{-6} M, analogue **81** exhibited an average ice grain size of 0.0049 mm^2, which corresponded to 1/3 of the activity of AFGP8 when it showed neither TH nor DIS activity. For all these analogues, the RI activity increased in a non colligative manner ruling out any non-specific RI effects. When we observe the highest concentration (5.5x10^{-6}M) several trends can be detected. At this point in the discussion, our goals and the progress made will be discussed.

a) Study of the importance of the (1,3)-β-Gal-GalNAc disaccharide unit versus monosaccharide moiety

Similar to the TH analysis, the RI activity of the α-GalNAc analogue **81** was compared to the activity of (1, 3)-β-Gal-GalNAc of AFGP8. This analogue lost 2/3 of the RI activity as compare to the native form. This information seems to confirm that the terminal galactose is also essential for RI activity. As a result, the α-galactose and β-galactose

analogues **70** and **78** were also tested. Both exhibited RI activity, however the α-galactose analogue **70** demonstrates a more significant RI activity while the β-analogue exhibits only limited activity. This suggests that the disaccharide moiety is an important structural feature for RI activity in our 4-mer analogues. However, it is not essential and some RI activity can be observed even with monosaccharide analogues.

b) Determining the effect of stereochemistry of the glycosidic linkage on RI activity (i.e. α or β linkages)

The fact that the α-analogues **70, 72, 80** and **81** are 20% to 70% more active than any of the β-analogues **78** and **79**, indicates that α linkages to the peptide backbone are more favourable to generate interaction with the ice lattice and RI activity. This trend is clearly demonstrated by the difference of activity between α-Gal **70** and β-Gal **78**, where the α-analogue is almost 50% more active than its β-counterpart.

c) Influence of relative stereochemistry of the hydroxyl groups in the pyranose ring on RI activity

Interestingly, neither α-glucose analogue **71** nor β-glucose analogue **79** exhibits any noteworthy RI activity. Compared to these analogues, α-galactose **70** and α-mannose **80** are significantly more active (from 20% to 50%). Moreover even while comparing β-galactose to β-glucose, we can notice a slight increase of 7% in RI activity. These two observations suggest that the relative stereochemistry of the hydroxyl groups of the carbohydrate moiety is as important for RI activity as it was for the interaction with the ice lattice observed during the DIS testing.

d) The role of the N-acetyl group (C-2 position) on RI activity

The *N*-acetyl at the C-2 position on our *O*-linked analogues seems to have some influence on RI activity. Indeed, α-GalNAc analogue **81** is 12% more active than its galactose counterpart **70**.

This trend was contradictory to what was observed in the previously reported *C*-serine analogues.[42] Indeed the *N*-acetyl galactosamine *C*-serine analogues at a concentration of $5.5X10^{-6}$ M exhibit an RI activity of about 31% of analogue *C*-[SGG(Gal)]$_4$ (Figure 16).[42] Moreover *C*-[SGG(GalNAc)]$_4$ displayed dynamic ice shaping while *C*-[SGG(Gal)]$_4$ did not. There are two possible explanations for this. First, all the information that we will learn about the necessary structural features for RI and TH activities on our *O*-linked analogues

might not be transposable to our *C*-linked analogues. Secondly, because *C*-[SGG(GalNAc)]$_4$ only displayed DIS activity while *C*-[SGG(Gal)]$_4$ only displayed RI activity, we can confirm that DIS, RI and TH activities do not imply a similar degree of interaction with the ice lattice (i.e. concentration threshold, necessary structural features, etc.) and it is therefore possible to custom tailor an AFGP analogues for one activity or another.

e) *The role of the γ-methyl (on the threonine residue) on RI activity*

As mentioned previously, the serine analogue **72** was synthesized in order to determine whether the hydrophobicity of the γ-methyl group would have any influence on antifreeze activity. For the RI activity, it seems that this influence is notable, in fact a loss of 15% activity is observed as compared to its threonine counterpart **70** (Figure 16). Therefore the γ-methyl group of the threonine possess a small but beneficial influence on the RI activity.

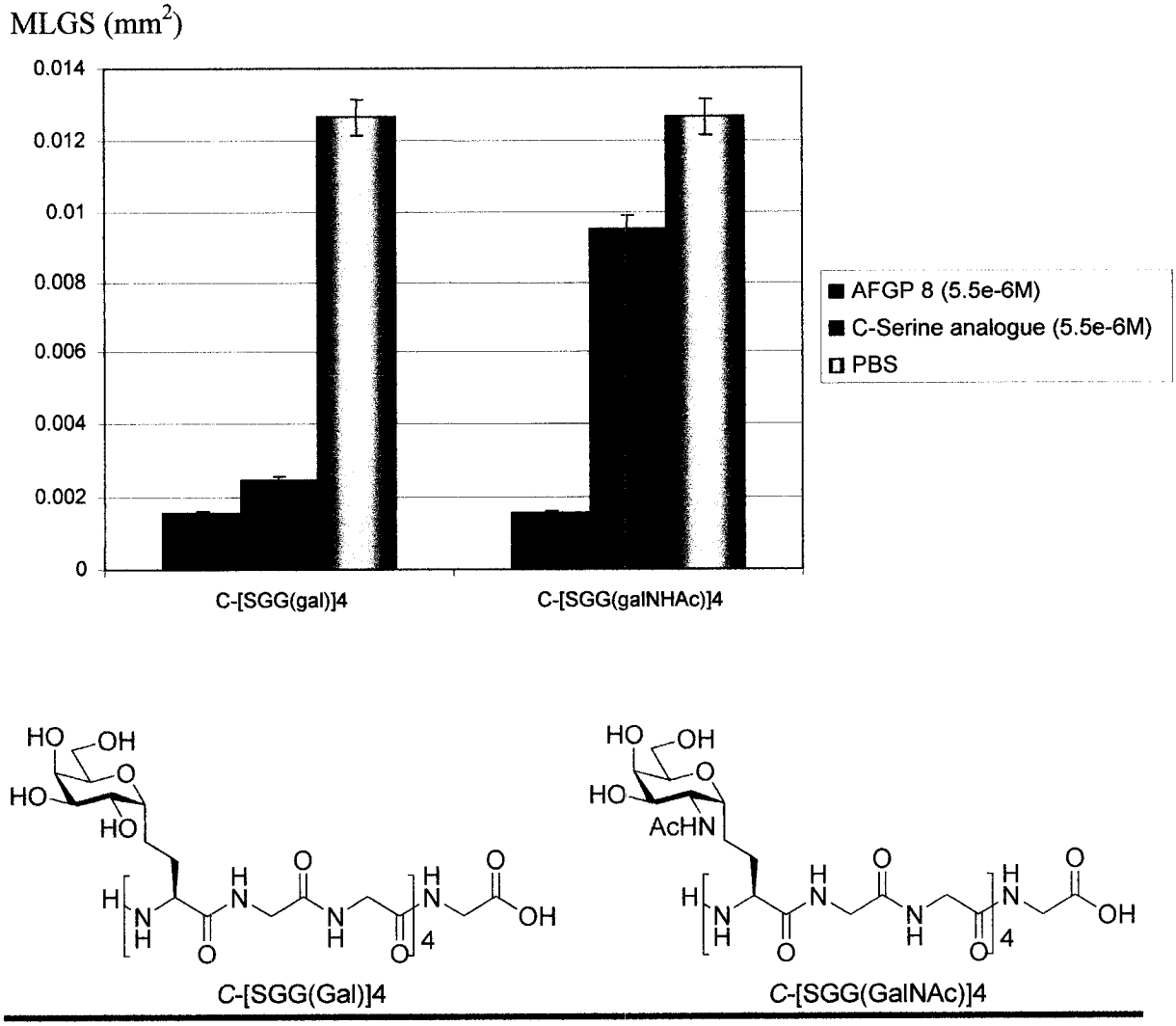


Figure 16: RI activity for C-Serine analogues[42]

4.3.3 General conclusions on TH and RI activities

From this study several conclusions can be made regarding the relative role ofdifferent AFGP structural features in TH and RI activity. Furthermore, the results of the research presented in this chapter strongly suggest that future generations of AFGP analogues could selectively display RI activity similar to the native AFGPs without exhibiting any TH or DIS activities. Design of future analogues will be based on the following conclusions extracted from the research presented in this chapter:

- TH and DIS activities are more sensitive to the length of the peptide backbone of the glycopolymer than RI activity. Therefore a shorter analogue (i.e. 4-mer) could display no TH and DIS activities while still exhibiting RI activity.
- The terminal galactose is essential for TH activity but not for RI activity.
- The α-linkage and the stereochemistry of the galactose hydroxyl groups are key elements for ice lattice interaction and therefore essential feature for both RI and TH activity.
- the presence of a γ-methyl group on the threonine, and the C-2 *N*-acetyl functional group on the galactose can improve both TH and RI activity.

4.3.4 Conformational Study of our analogues using Circular Dichroism (CD)

In order to understand whether the different antifreeze activities demonstrated in *O*-linked AFGP analogues are related to their solution conformations, we performed a CD spectroscopy study with these analogues. The secondary structural elements of proteins and polypeptides, such as *α*-helix, *β*-sheet, *β*-turn, polyproline type II helix and random coil can give characteristic absorption bands at the far-UV region (190-250 nm). Therefore, CD spectroscopy has been used as a valuable tool to characterize the secondary structures of folded proteins and polypeptides. In chapter 3, we described a concentration and temperature study of the conformation of AFGP8. Our findings are also in agreement with conformational studies performed using Raman spectroscopy[43] where the presence of an α-helix and a β-sheet was observed, but the major conformation was considered to be random coil.

The CD spectra of *O*-linked AFGP analogues and AFGP8 were measured for a concentration of 43μM in double distilled water at 25°C (Figure 17). In order to understand whether the conformations of AFGPs and analogues are related to antifreeze-specific activities, it is ideal to perform the CD spectroscopy study at temperatures below 0°C and with ice existing in the system. However, in the past quasi-elastic light scattering and CD studies proved that no obvious conformational changes were observed for AFGPs at different temperatures.[44,45]

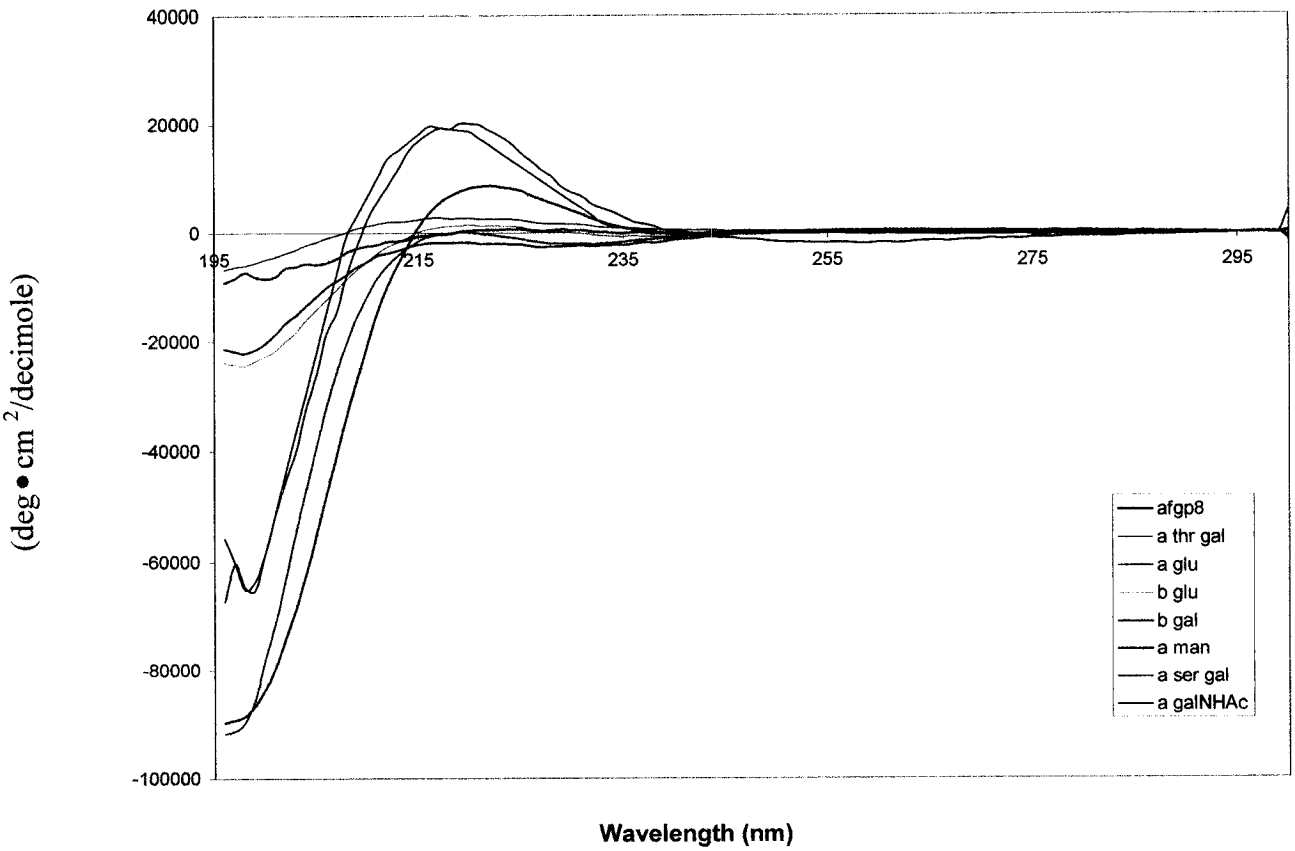


Figure 17: Circular dichroism of our O-linked analogues

All the spectra were deconvoluted using the same software (SELCON3) and parameters (IBASIS5) as for chapter 3. Obviously, it is not practical to understand the exact conformation of AFGP8 and *O*-linked AFGP analogues based on the CD deconvolution study alone. However, if performing all the deconvolution calculations with the same set of reference proteins, useful information still can be gained on the conformational differences among the *O*-linked AFGP analogues. The deconvolution analysis results of *O*-linked AFGP analogues and AFGP8 are listed in Table 8.

Table 8: CD deconvolution of O-linked analogues. *The analogues are in order from the most active to the least active for RI activity.*

Compound	α-helix	β-sheet	β-turn	PPII	Random coil	Sum
AFGP8	0.101	0.212	0.136	0.108	0.434	0.990
α-GalNAc 81	0.155	0.197	0.171	0.145	0.311	0.978
α-Gal(thr) 70	0.281	0.201	0.156	0.102	0.273	1.013
α-Gal(ser) 72	0.133	0.192	0.144	0.115	0.378	0.962
α-Man 80	0.256	0.182	0.145	0.104	0.284	0.970
β-Gal 78	0.068	0.122	0.106	0.100	0.608	1.005
α-Glc 71	0.030	-0.025	0.126	0.209	0.704	1.043
β-Glc 79	0.032	0.122	0.101	0.072	0.657	0.983

The major conformation observed in the most active compounds [α-GalNAc (**81**), α-Gal(thr) (**70**), α-Gal(ser) (**72**) and α-Man (**80**)] is the random coil (ranging from 27 to 33%). All the other conformations are equally observed with a minimum amount of α-helix (14%). A positive band at wavelengths between 216-220 nm are observed in all of our α-galactosyl analogues [α-GalNAc, α-Gal(thr), and α-Gal(ser)], which resemble the same trend observed in AFGP8. The overall CD spectra of analogue α-GalNAc and α-Gal(ser) share a very similar shape, which is also verified by the similar deconvolution data. These results, taken together with the fact that our *O*-linked analogues exhibit a relatively limited RI activity compared to AFGP8, indicate that conformation may not be the only determining factor for RI activity. One interesting comment for dynamic ice shaping activity is the fact that only α-Gal(thr) and α-Man exhibit any dynamic ice shaping activity, whereas α-GalNAc, α-Gal(thr), α-Gal(ser) and α-Man possess similar proportions of conformations. These results indicate here again that the overall population of conformations of our analogues is not the sole factor in order to obtain dynamic ice shaping activity. While comparing α-GalNAc with AFGP8, we observed that the ratio of different conformations is relatively identical. This is suggesting that the terminal galactose present in AFGP8 does not imply any essential conformational modifications, therefore the difference of RI activity between the two

compounds is probably due to the difference in carbohydrate concentration on a single glycopeptide.

Another interesting piece of information that can be extracted from these spectra is the fact that α-Gal and α-GalNAc analogues possess very similar conformations. Therefore, the potential intramolecular hydrogen bonding of the *N*-acetamide does not seem to have much influence on the overall distribution of the conformation of the population. However, this compound is slightly more active for RI activity than its α-Gal analogue. It is therefore possible that this hydrogen bonding intervenes in the ice adsorption mechanism.

α-Gal(thr) and α-Gal(ser) also possess very similar conformations, suggesting that the γ-methyl of the threonine does not cause significant conformational changes even though it slightly improves the RI activity of the analogue.

The three less active compounds β-Gal (**78**), α-Glc (**71**) and β-Glc (**79**) analogues are also the ones with the least α-helix conformation. It is therefore possible that a minimum amount of α-helix or other rigid conformations might favour the RI activity. This preliminary CD spectroscopic study indicates that the conformation of our *O*-linked AFGP analogues is an important, though not the only, factor for antifreeze-specific activities.

[1] Stick, R.V. *Carbohydrates: The sweet molecules of life*, Academic Press **2001**.

[2] Koenigs, W.; Knorr, E. *Ber. Dtsch. Chem. Ges.* **1901**, 34, 957.

[3] Paulsen, H. Angew. *Chem. Int. Ed. Engl.* **1982**, 13, 15.

[4] Mukaiyama, T.; Murai, Y.; Shoda, S. *Chem. Lett.* **1981**, 431.

[5] Takeuchi, K.; Mukaiyama, T. *Chem. Lett.* **1998**, 555.

[6] Ferrier, R.J.; Hay, R.W.; Vethaviyasar N. *Carbohydr. Res.* **1973**, 27, 55.

[7] Veeneman, G.H.; Van Boom, J.H. *Tetrahedron Lett.* **1990**, 31, 275.

[8] Kahne, D.; Walker, S.; Cheng, Y.; Van Engen, D. *J. Am. Chem. Soc.* **1989**, 111, 6881.

[9] Green, L.; Hinzen, B.; Ince, S.J. Langer, P.; Ley, S.V.; Warriner, S.L. *Synlett*, **1998**, 440.

[10] Stick, R.V.; Tilbrook, D.M.G.; Williams, S.J. *Aust. J. Chem.* **1997**, 50, 233.

[11] Fraser-Reid, B.; Madsen, R. *Oligosaccharide synthesis by n-pentenyl glycosides, in Preparative Carbohydrate Chemistry,* Hanessian, S. ed., Marcel Dekker, New-York, **1997**, 339.

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

[13] Schmidt, R.R. *Modern Methods in Carbohydrates Synthesis*, Khan, S.H.; O'Neill, R.A. eds, Harwood Academic, Netherlands, **1996**, 20.

[14] Schon, I.; Kisfaludy, L *Synthesis* **1986**, 303.

[15] Schon, I.; Kisfaludy, L *Synthesis* **1983**, 325.

[16] Frische, K.; Meldal, M.; Werdelin, O.; Mouritsen, S.; Jensen, T.; Galli-Stampino, L.; Bock, K. *J. Pept. Sci.*, **1996**, 2, 212.

Garegg, P.J.; Norberg, T. *Acta Chem. Scand.* B **1979**, 110.

[17] Carvalho, I.; Scheuerl, S.L.; Kartha, R.; Field, R.A. *Carbohydr. Res.* **2003**, 338, 1039.

[18] Lacombe, J.M.; Pavia A.A. *J.Org. Chem.* **1983**, 48, 2557.

[19] Greene, T.W.; Wuts, P.G.M. *Protective groups in organic chemistry,* John Wiley and Sons, Inc. New-York **1999**.

[20] Lemieux, R.U.; Ratcliffe R.M. *Can. J. Chem.* **1979**, 57, 1244.

[21] llen, J. R.; Harris, C. R.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2001**, 123, 1890.

[22] Kunz, H.; Birnbach, S. *Angew. Chem. Int. Ed.* **1986**, 98, 354.

[23] Carpino, L.A.; Han, G.Y. *J. Org. Chem.* **1972**, 37, 3404.

[24] Fields, G.B.; Noble, R.L. *Int. J. Pept. Protein Res.* **1990**, 35, 161.

[25] Houston, M.E.; Chao, H., Hodges; R.S.; Sykes, B.D.; Kay, C.M.; Sonnichsen, F.D.; Loewen, M.C.; Davies, P.L. *J. Biol. Chem.* **1998**, 273, 11714.

[26] Fletcher, G.L; Hew, C.L.; Davies, P.L. *Annu. Rev. Physiol.* **2001**, 63, 359.

[27] Feeney, R.E.; Burcham, T.S.; Yeh, Y. *Annu. Rev. Biophys. Biophys. Chem.* **1986**, 15, 59.

[28] Yeh, Y.; Feeney, R.E. *Chem. Rev.* **1996**, 96, 601.

[29] Scholander, P.F.; Maggert, J.E. *Cryobiology* **1971**, 8, 371.

[30] Knight, C.A.; DeVries, A.L.; Oolman, L.D. *Nature* **1984**, 308, 295.

[31] Raymond, J.A.; DeVries, A. L. *Proc. Natl. Acad. Sci. U.S.A.* **1977**, 74, 2589.

[32] Davies, P.L.; Hew, C.L. *FASEB J.* **1990**, 4, 2460.

[33] Galema, S.A.; Howard, E.; Engberts, J.B.F.N.; Grigera, J.R. *Carbohydr. Res.* **1994**, 265, 215.

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

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

[35] Meryman, H.T. *Annu. Rev. Biophys. Bioeng.* **1974**, 3, 341.

[36] Knight, C.L.; Wen, D.Y.; Laursen, R.A. *Cryobiology* **1995**, 32, 23.

[37] Knight, C.L. *Nature* **2000**, 406, 249.

[38] Knight, C.A.; Hallett, J.; DeVries, A.L. *Cryobiology* **1988**, 25, 55.

[39] Myers, K.L. "Quantification of recrystallization inhibition: an assay for antifreeze protein activity" M.A. Thesis, Binghamton University (S. U. N. Y.) **1999**.

[40] Horwath, K.L.; Easton, C.M.; Poggioli Jr., G.J.; Myers, K.; Schnorr, L. *Eur. J. Entomol.* **1996**, 93, 419.

[41] Enaide, A.; Purushotham, M.; Ben, R.N. *Cell Biochemistry and Biophysics* **2003** 38, 115.

[42] Liu S. Ph.D. Thesis, University of Ottawa, Ottawa, **2006**.

[43] Yomimatsu, Y.; Scherer, J.R.; Yin Y., Feeney, R.E. *J. Biol. Chem.* **1976**, 251, 2290.

[44] Ahmed, A.I., Feeney, R.E., Osuga, D.T., Yeh, Y. *J. Biol. Chem.* **1975**, 250, 3344.

[45] Bouvet, V. R., Lorello, G. R., Ben, R. N. *Biomacromol.* **2006**, 7, 565.

Chapter 5

SYNTHESES AND ACTIVITY ASSESSMENT OF *C*-LINKED ANALOGUES OF AFGP8

Bouvet, V. R.; Ben, R. N. *Journal of Organic Chemistry* **2006**, 71, 3619.

5 SYNTHESES AND ACTIVITY ASSESSMENT OF *C*-LINKED ANALOGUES OF AFGP8

5.1 Introduction to the formation of *C*-linked carbohydrates

C-glycoside synthesis has been reviewed by Postema,[1] Levy,[2] Sinay,[3] Beau[4] and Nicotra.[5] A general overview of the most common synthons and intermediates in *C*-glycosylation is shown in Figure 1.

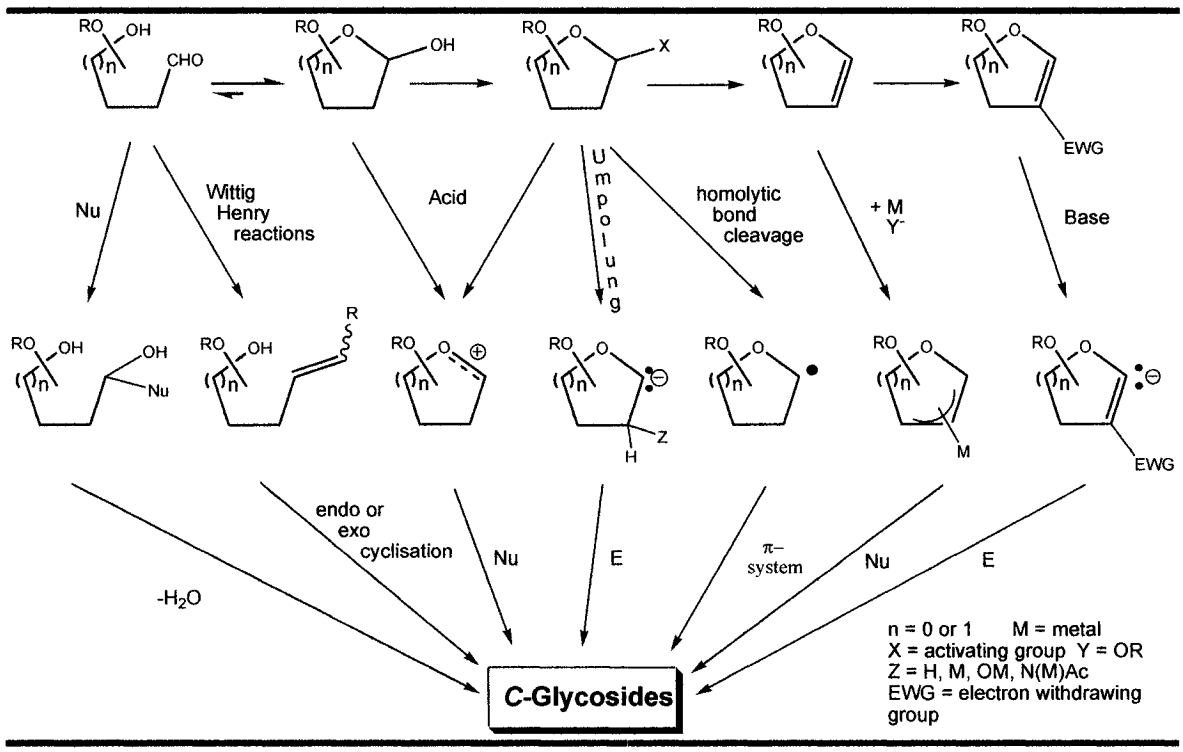


Figure 1: Common synthons and intermediates in the C-glycosylation[6]

These methods all involve the use of glycosides and chiral amino acid equivalents to stereoselectively form the *C*-linked glycosidic bond at the anomeric center. Toward this end, different intermediates (Figure 2), such as glycals,[7] glycosyl pyridyl sulfones,[8] and glycosyl trichloroacetimidates,[9] have been used as the glycoside component for stereoselective glycosylation.

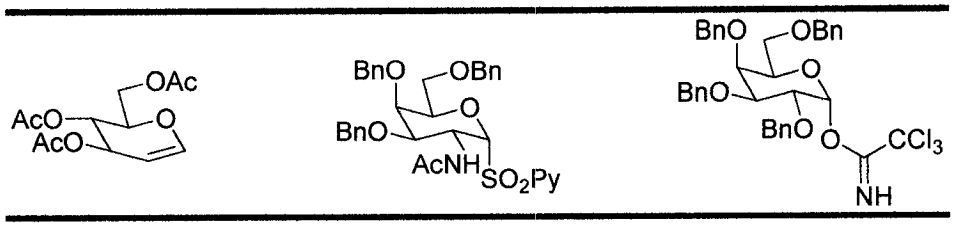


Figure 2: Glycosides used in stereoselective glycosylation.[7-9]

Although these strategies, along with many similar published methods,[10] have been employed to prepare *C*-glycosyl amino acids, problems such as low chemical yields and/or low stereoselectivity still exist and limit their application. For example, the harsh reaction conditions for generating enolates in the chiral alkylation strategy restrict the scope of protecting group selection for the carbohydrate moieties. In addition, after glycosylation reaction, additional multi-step manipulations are often required to convert chiral amino acid precursors to final products.

5.2. Synthesis of *C*-galactosyl building blocks for AFGP analogues

As outlined in the previous chapters, the SAR study of previously prepared *C*-linked AFGP analogues demonstrated that D-galactose possessed the optimal hydroxyl group configuration for antifreeze-specific activities (TH, DIS and RI). In this section our goal is to synthesize $[LGG(Gal)]_4$ and $[LAA(Gal)]_4$ (Figure 3).

The synthesis and study of $[LGG(Gal)]_4$ will allow us to elucidate the effect of:

-polymer length on RI and TH activity as compared to previously prepared $[LGG(Gal)]_3$, $[LGG(Gal)]_6$ and $[LGG(Gal)]_9$.

-the length of the peptide side chain, and allow us to investigate which length of the peptide side chain is favorable for RI activity, by verifying whether the glycosylated ornithine $[OGG(Gal)]_4$ or lysine $[LGG(Gal)]_4$ analogue gives the optimal RI activity.

-peptide sequence on RI and TH activities (*i.e.* influence of the proline insertion in the peptide backbone when compared to previously prepared $[LPG](Gal)]_4$). Comparing the two analogues will allow us to determine if the substitution of the proline, which incorporates some rigidity in the peptide backbone $[LPG(Gal)]_4$, has

any influence on the RI and TH activity of the *C*-linked analogues $[LGG(Gal)]_4$. Proline was chosen because of its natural occurrence in the native AFGP8.

$[LAA(Gal)_4]$ should give us information on the significance of the peptide sequence. The peptide backbone of the native system is mainly composed of alanine, whereas our *C*-linked analogues contain glycine. Investigating the influence of alanine substitution on our *C*-linked analogues may elucidate the role of the alanine and its influence on RI and TH activity in our *C*-linked analogues. This can be performed by comparing $[LAA(Gal)]_4$ to $[LGG(Gal)]_4$.

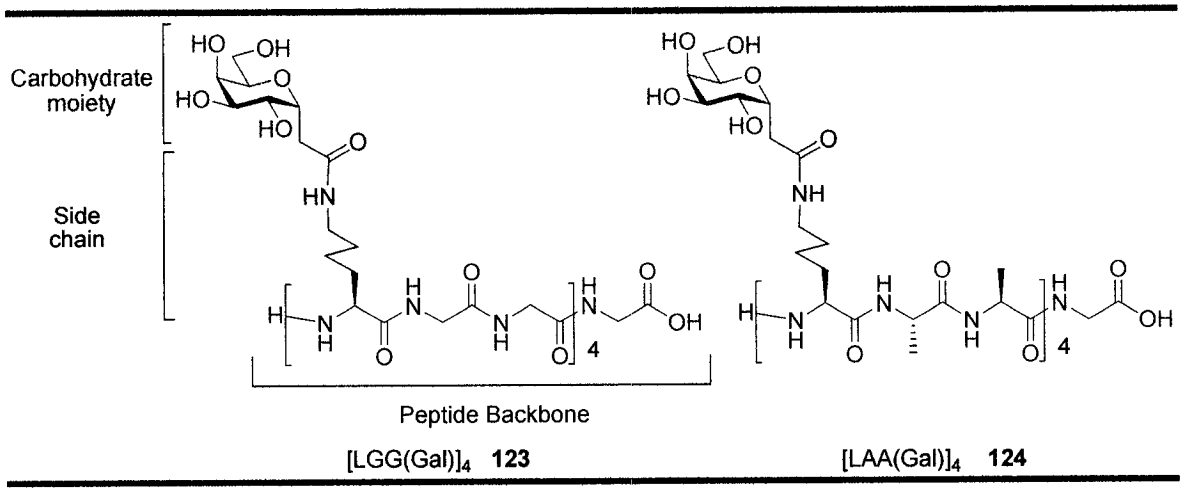


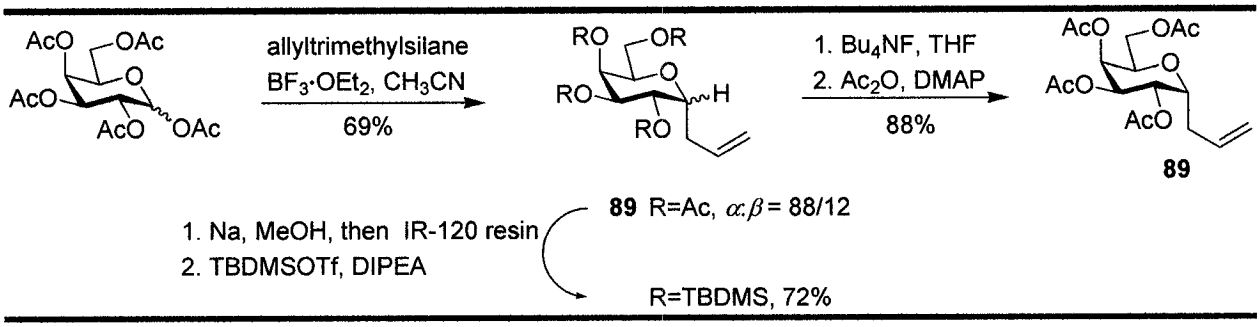
Figure 3: C-linked analogues containing lysine in the peptide backbone

5.2.1 Preparation of C-galactosyl alkenes as C-linked building blocks.

In order to obtain these analogues it was necessary for us to prepare a *C*-linked precursor, which in our case was *C*-allylated galactose **89**. *C*-Allylated glycosides have been previously used as important synthons for the preparation of *C*-linked glycoconjugates.[11] These compounds are usually prepared using one of two strategies: Lewis acid-catalyzed nucleophilic displacement on activated glycosides,[12] or radical-mediated addition of allyl-containing agents to anomeric glycosyl radical precursors.[13] The latter is also known as the Keck-type allylation,[14] but due to the high cost of the required allylphenylsulfone and the stannane, we decided to first attempt a nucleophilic displacement. Peracetylated bromogalactose **29** was treated with allyltrimethylsilane and a catalytic amount of BF_3-OEt_2 to afford an 88/12 α/β mixture of **89** with a combined yield of 69% (Scheme 1). One

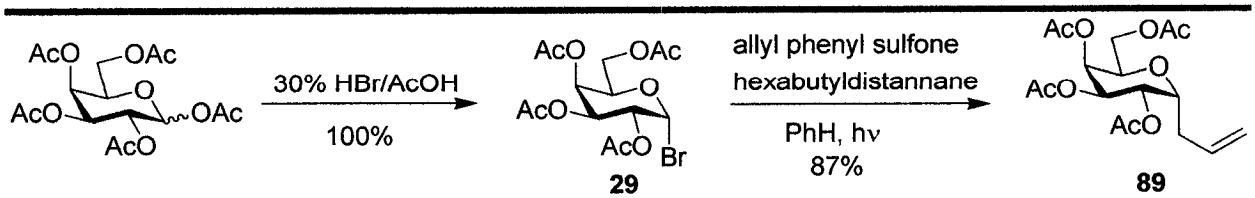
problem associated with this strategy is that the α/β mixture was inseparable by column chromatography.[15] The formation of the *β* anomer is due to neighboring-group participation.[16] In order to obtain pure *α-C*-allylated galactose **89**, the acetate protecting groups were converted to the less polar *tert*-butyldimethylsilyl ether (OTBDMS) groups after the allylation. This change of protecting group permitted the separation of the α/β anomers. Once pure anomers were obtained the silyl groups were converted back to the acetate functionality. While these extra steps permitted the isolation of pure **89**, the long length of the synthesis prompted us to pursue a radical-mediated approach.

Scheme 1: Preparation of C-allylated glycosides by Lewis acid-catalyzed allylation[14]



The radical-mediated coupling of glycosyl halides with allylic sulfides or sulfones usually affords α-*C*-glycosides with high stereoselectivity.[17] Glycosyl bromides are good candidates for such reactions because of the labile anomeric C-Br bond. However, several other radical precursors have also been used in literature such as sulfide, sulfone, trichloroimidate and selenide. In our case the bromide precursor was easily prepared and the C-Br bond could be readily cleaved in a homolytic manner to form a stable anomeric radical on glycosides (Scheme 2).[18]

Scheme 2: Preparation of α-C-allylated galactose[18]



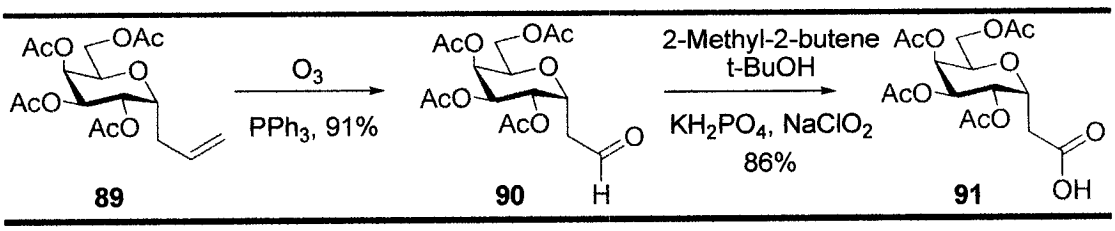
As illustrated in Scheme 2 the synthesis of α-*C*-allylated galactose started from D-galactose pentaacetate. Bromination with 30% hydrogen bromide in acetic acid solution afforded bromogalactose in quantitative yield. The product formed was the pure α-anomer, due to the strong anomeric effect that favors the thermodynamically more stable α-anomeric radical.

Conversion of bromide **29** to the desired allyl intermediate **89** was then accomplished using the procedure of Keck.[19] In this procedure Keck used a combination of allyl sulfides and hexabutyldistannane to replace allyl stannanes and successfully extended the application of the reaction.[17] Magnusson was the first to employ this strategy in carbohydrate synthesis.[18] Following Magnusson's procedure, α-*C*-allylated galactose tetraacetate **89** was prepared by a radical-mediated allylation of the galactosyl bromide with allyl phenyl sulfone. After initiation, the stannyl radical could either abstract bromine from the substrate or be trapped by the olefin in allyl phenyl sulfone. Both reaction pathways would lead to the desired α–allylation product, which was obtained in 87% yield. This is presumably due to the fact that the axially positioned anomeric radical can be stabilized by the interaction of the lone pair of the intramolecular oxygen with the SOMO of the radical. Therefore, the allylation of the anomeric radical gave only the α-anomer product.

5.2.2 Assembly of the building blocks for C-linked AFGP analogues

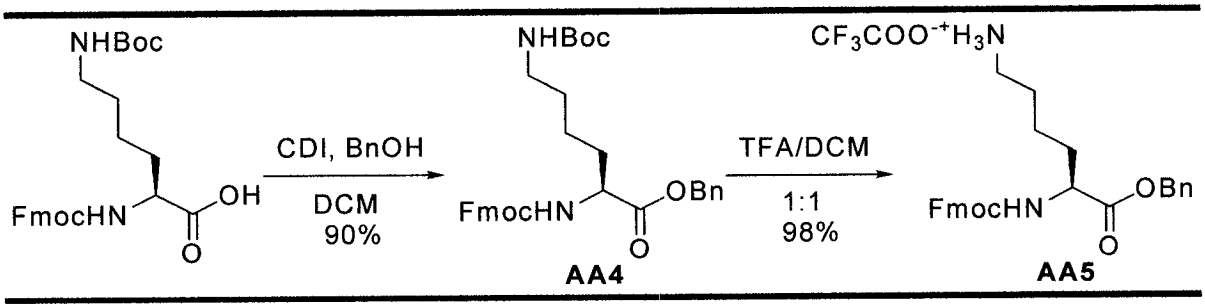
The allyl galactose **89** was then oxidized using ozone and triphenylphosphine as a reducing agent to afford aldehyde **90** (Scheme 3). This aldehyde was further oxidized in 86% yield to afford the carboxylic acid precursor **91** necessary for the formation of the amide bond in the desired AFGP analogue.[20]

Scheme 3: Oxidation of C-allylated galactosyl derivative[20]

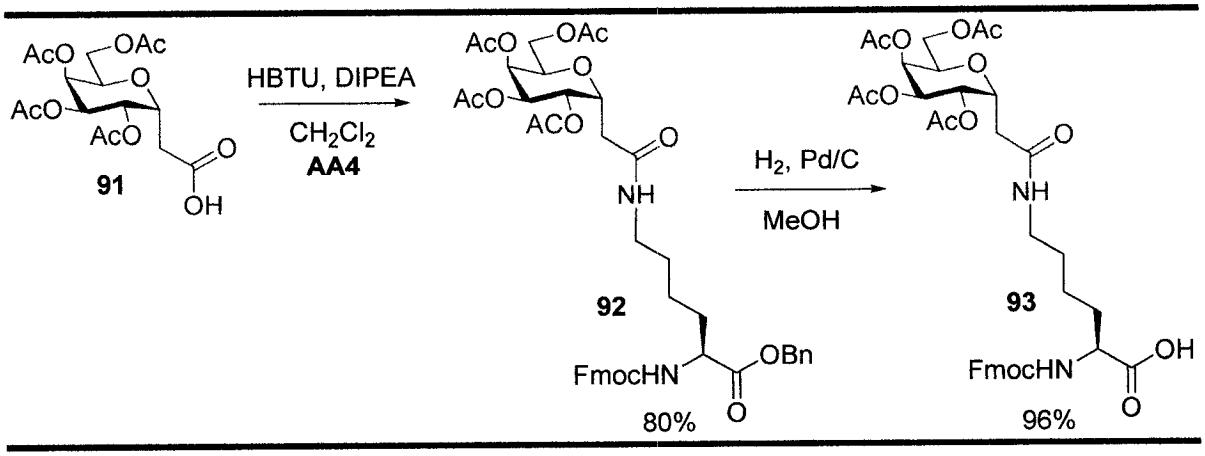


Once we were in possession of the carboxylic acid synthon **91**, we coupled it with the lysine precursor **AA5**. This amino acid was prepared in high yield using commercially available Fmoc-Lys(Boc)-OH via benzylation followed by subsequent treatment with 50% TFA in CH_2Cl_2. The final product was crystallized in ether (Scheme 4). The coupling of **91** and **AA5** using HBTU and DIPEA then afforded the desired compound **92** in 80% yield. Compound **92** was then debenzylated using H_2 and Pd/C to afford the desired building block **93** for solid phase synthesis (Scheme 5). The building block FmocLys[α-*C*-Gal(OAc)$_4$]OH **93** was therefore obtained in 6 steps from galactose pentaacetate in 52% overall yield.

Scheme 4: Preparation of the peptide backbone



Scheme 5: Preparation of the lysine galactose building block



5.3. Synthesis of *C*-linked GalNAc containing building blocks

5.3.1 Synthesis of the monosaccharide building block

In this section we desired to investigate the relative importance of an *N*-acetyl in the C-2 position of *C*-linked AFGP analogues (Figure 4). To do this, an acetamide function in the C-2 position was incorporated into our most promising analogue $[OGG(Gal)]_4$. In the *O*-linked system, the *N*-acetyl analogue results in an increase of 12% in RI activity compared to the galactose analogue. We desired to verify if this trend could be preserved in the *C*-linked analogues.

In addition, we desired to observe the influence of alanine in the peptide chain of the *C*-linked GalNAc analogues and determine whether or not this alanine would have the same effect than in the *C*-linked galactose serie. Toward this goal, we will compare $[LAA(Gal)_4]$ to $[LGG(Gal)]_4$ for the galactose serie and $[OGG(GalNAc)]_4$ to $[OAA(GalNAc)]_4$ for the GalNAc serie. Toward this goal, we will compare the activity of $[OGG(GalNAc)]_4$ and $[OAA(GalNAc)]_4$.

Finally, we wanted to determine whether the influence of the disaccharide *β*-D-galactosyl-(1,3)-D-*N*-acetylgalactosamine present in the native system would be the same than for the *O*-linked analogues. In our previous chapter we demonstrated by comparing $[TAA(GalNAc)]_4$ **81** to AFGP8, that all the TH and DIS activity as well as 2/3 of the RI were lost when the terminal galactose was removed. It is possible that similar phenomena would be observed for our *C*-linked analogues.

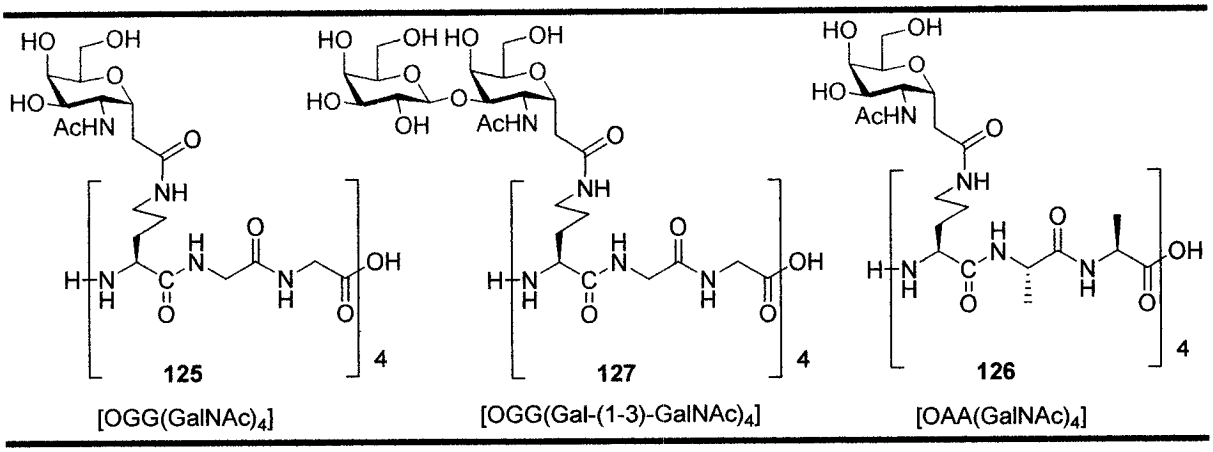
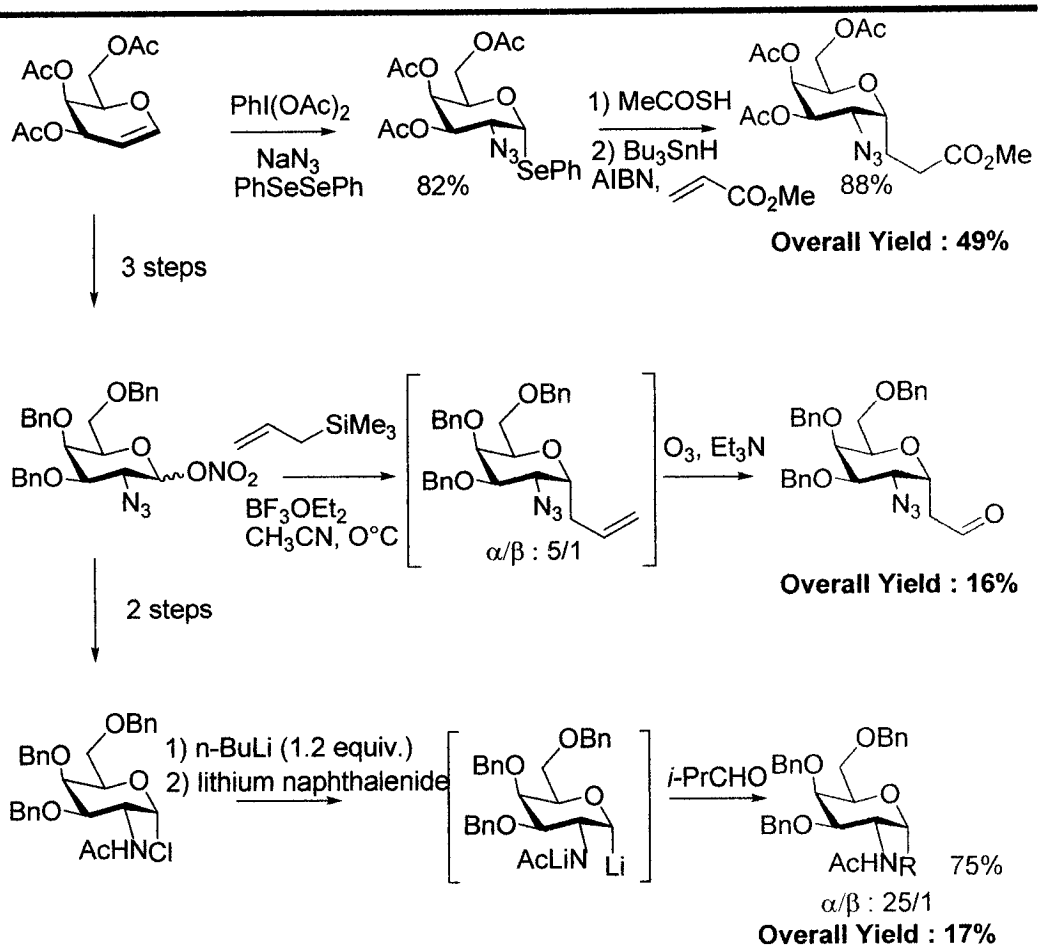


Figure 4: C-linked GalNAc targets

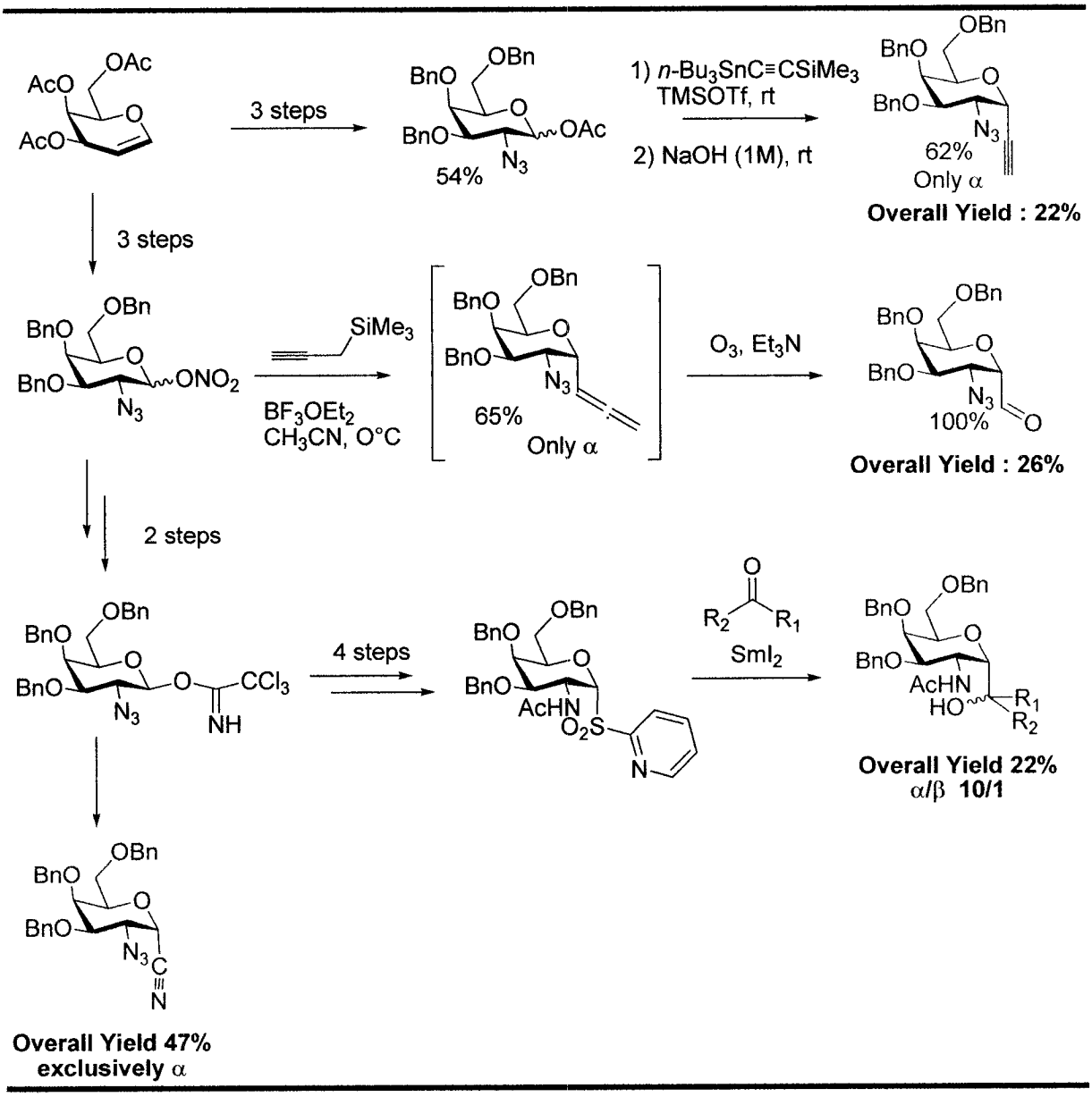
5.3.1.1 Preparation of C-GalNAc alkenes

Although many methods have been developed to prepare *C*-linked carbohydrate derivatives, the synthesis of *C*-linked glycosyl amine derivatives is still lengthy and inefficient.[21] The preparation of *C*-linked 2-deoxy-2-acetamido-α-D-galactopyranose derivatives is especially difficult due to the incompatibility of C-2 nitrogen protecting groups with most *C*-glycosylation strategies.[22] Currently reported syntheses of *C*-linked 2-deoxy-2-acetamido-α-D-galactopyranose derivatives utilize glucosyl derivatives (*via* oxime intermediates)[23] or galactosyl pyranose[24] derivatives (*via* galactal intermediates) as starting materials. The latter approach often introduces a C-2 acetamide precursor *via* azido nitration,[24a] azido chlorination[24b] or azido selenation,[24c-f] and these intermediates are then subjected to stereoselective carbon–carbon bond forming reactions (Scheme 6 and Scheme 7). For instance, acetylenic,[24g] allylic,[24h] allenic,[24h] cyano[24i] and other derivatives[24j-l] have been prepared following azidonitration or chlorination of the appropriate glycal.

Scheme 6: Previous synthesis to obtain C-linked -2-acetamido-α-D-galactose[24]



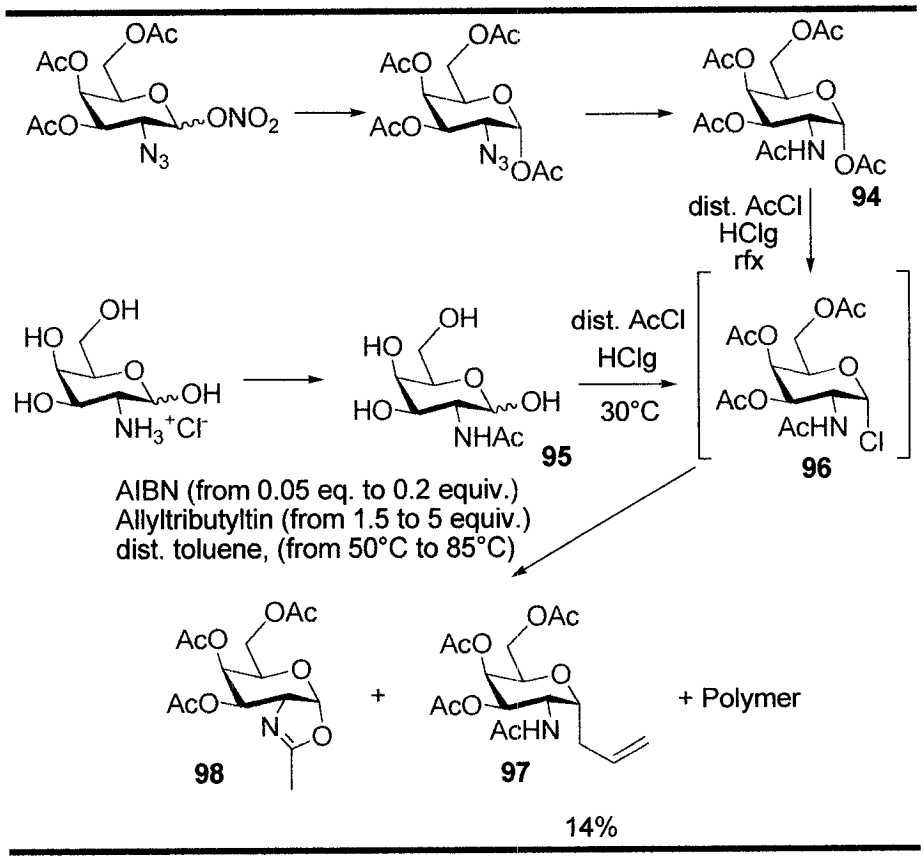
Scheme 7: Previous synthesis to obtain C-linked -2-acetamido-α-D-galactose (continued)[24]



Similarly, a number of *C*-linked GalNAc derivatives[24m,n] including *C*-linked disaccharide GalNAc derivatives[24o] have been prepared *via* azido selenation of galactal. *C*-linked *N*-acetyl galactosamine derivatives have also been prepared *via* direct Keck allylation[22,25,26a] of *N*-acetyl galactosamine.[26b] Based upon literature precedent, we chose to perform a direct allylation on *N*-acetyl galactosamine derivative **96** (Scheme 8).[26b]

Chloropyranose **96** was generated from peracetylated *N*-acetyl galactosamine **94,** which in turn was prepared *via* azido nitration of the corresponding glycal.[24a]

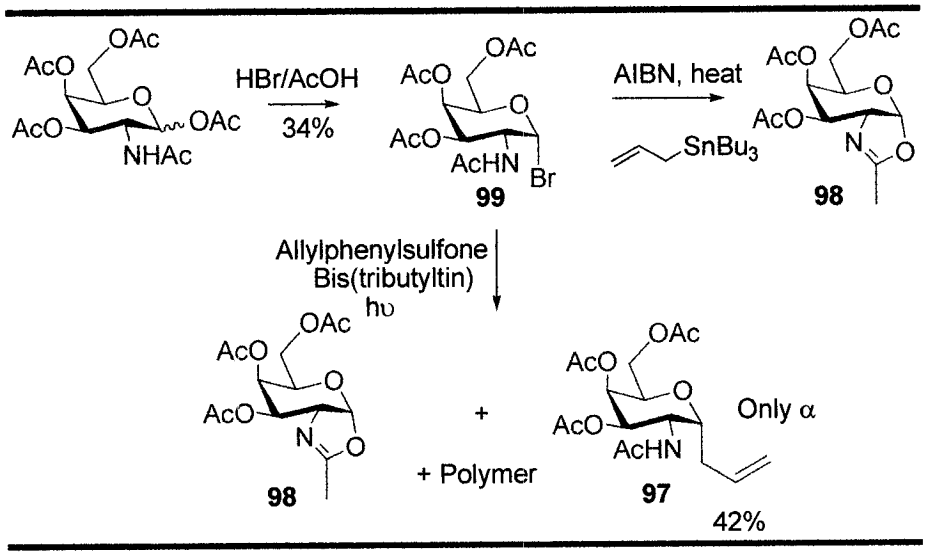
Scheme 8: Preparation and allylation of intermediate 96



Unfortunately, in our hands, conversion of **94** to **97** *via* generation of the chloro intermediate **96** failed to proceed with yields greater than 14%, even when previously reported optimal conditions (3 equivalents allyl tributylstannane and 0.15 mole equivalents AIBN) were employed.[26b] Changes in temperature, different Lewis acids ($AlCl_3$ and $TiCl_4$) and the saturation of the distilled acetyl chloride with HCl also failed to improve the yield of **97**.[27] Generation of **96** starting from D-galactosamine produced a similar result via the intermediate **95**. These results can be explained by the rapid degradation of the intermediate **96** and the inability to purify this intermediate by crystallization or silica-gel flash column chromatography. Therefore, we explored the possibility of performing either a Keck allylation[26a] or photochemically mediated radical allylation on bromo derivative **99** (Scheme 9). Unlike chloro intermediate **96**, we were able to purify bromo derivative **99** using silica

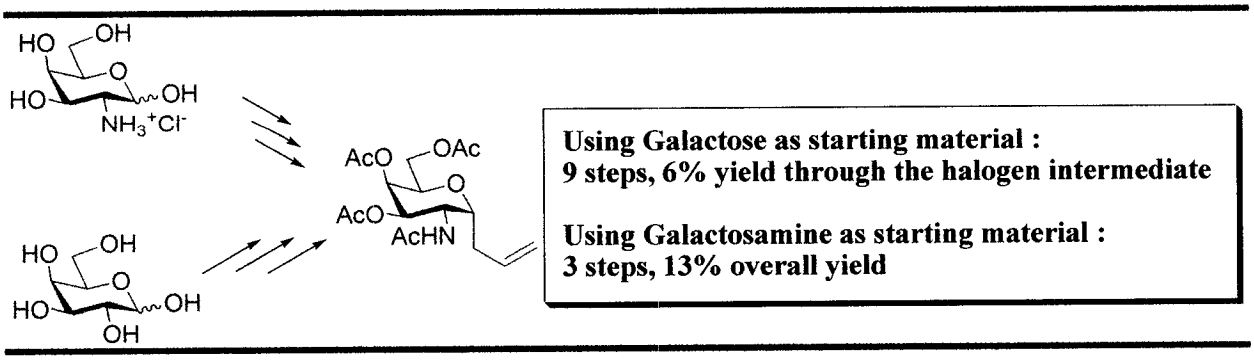
gel flash chromatography. However, when **99** was subjected to a Keck allylation, the only product obtained was the bicyclic oxazoline **98**. A photochemically mediated allylation[28] slightly improved the yield of **97** to 42%.

Scheme 9: Preparation and allylation of intermediate 99



While this sequence seems attractive given that only the α anomer is generated, preparation of the requisite starting material (2-acetamido-l,3,4,6-tetra-*O*-acetyl-2-deoxy-α,β-D-galactopyranose) is not trivial. For instance, preparation from D-galactosamine hydrochloride requires only one step,[27a] but galactosamine is prohibitively expensive. Furthermore, the preparation from galactose pentaacetate requires 5 to 7 steps[24a], with tedious purifications by column chromatography (Scheme 10). Consequently this approach was abandoned.

Scheme 10: Overall yields of our different strategies



It would be interesting to later complete this study by probing whether the position and nature of the terminal saccharide influences specific antifreeze activities (TH, DIS and RI) (Figure 5). Since it would be useful to obtain disaccharide analogues substituted at various positions, we decided to synthesize a fully orthogonally protected *C*-linked GalNAc precursor which could be used as a generic synthon for our present analogues as well as for our future analogues.

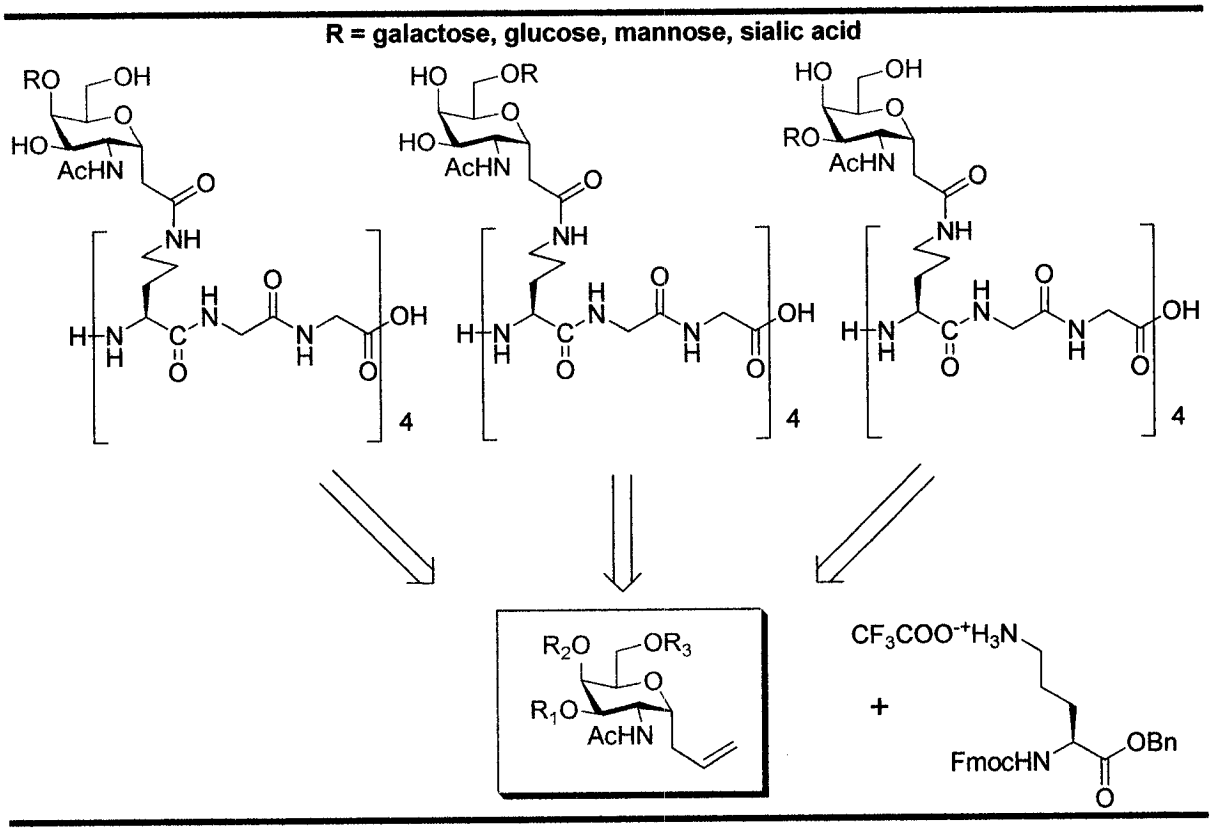


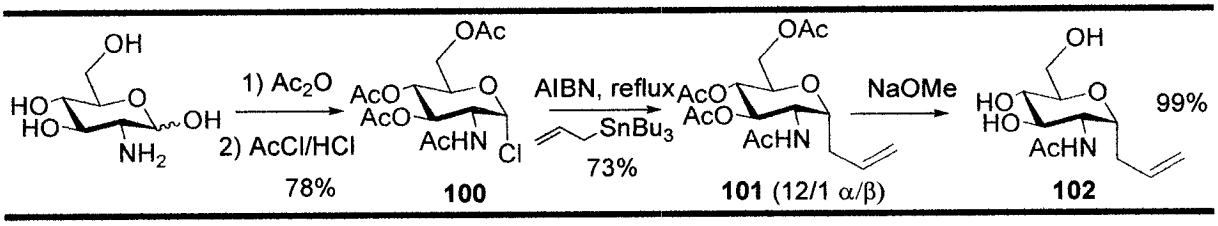
Figure 5: Generic precursor for all our C-linked galactosamine analogues

Given that Keck allylation of D-glucosamine derivatives[25] has been reported, we decided to combine different literature strategies and explore a novel synthetic route to access *C*-linked 2-deoxy-2-acetamido-α-D-galacto-pyranose derivatives **104** and **107** (Scheme 12) starting from commercially available and inexpensive D-glucosamine. The proposed strategy installs an allyl group at the anomeric position and inverts the stereochemistry at C-4 late in the synthesis. The inversion of stereochemistry at C-4 in various pyranose derivatives[29a,b] has been well studied on *O*-linked oligosaccharides.[29c-e] In

fact, a similar inversion has been reported in the synthesis of *C*-linked 2-acetamido-2-deoxy-α-D-galactose where the C-2 acetamido group was introduced using an oxime early in the synthesis.[23a] This approach was also attractive in that the C-3 hydroxyl group of **104** and **107** would be unprotected and available for a glycosidic coupling after the migration/epimerization sequence.

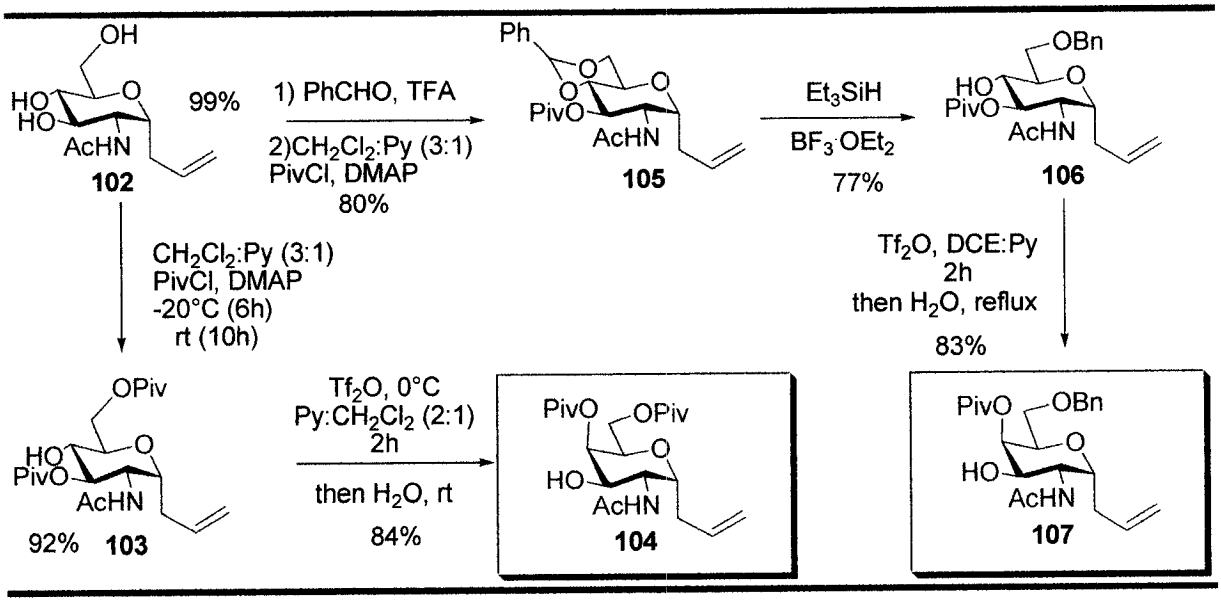
The first step in the synthesis was the preparation of 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride **100** (Scheme 11 and 12) from D-glucosamine which proceeded in 78% yield.[27b,c] Compound **100** was smoothly converted to *C*-linked derivative **101** *via* Keck allylation[26a] in 73% yield as a 12:1 α:β mixture of diastereomers. Optimized conditions for this allylation utilized 4.8 equivalents of allyl tributylstannane and 0.3 equivalents of AIBN in refluxing THF. While the α/β ratios were not as high as those previously reported,[27c] these conditions were advantageous in that no polymer by-products were observed. Furthermore, the amount of allyl stannane in the reaction mixture can be reduced by half compared to the Bertozzi procedure.[22] Deacetylation of **101** using NaOMe then furnished intermediate **102** in 99% yield. Compound **102** was then converted into *C*-linked 2-deoxy-2-acetamido-α-D-galacto-pyranose derivatives **104** and **107**.

Scheme 11: Preparation of the intermediate 102



With respect to the latter, the C-6 and C-3 hydroxyls were selectively protected using 2 equivalents of pivaloyl chloride to produce **104**. After some modifications of the original procedure (allowing the reaction to go overnight at RT following 7h at 0°C), we obtained the desired compound selectively with two pivaloyl protecting groups on C-6 and C-3 in 92% yield as an inseparable α:β 12:1 mixture.

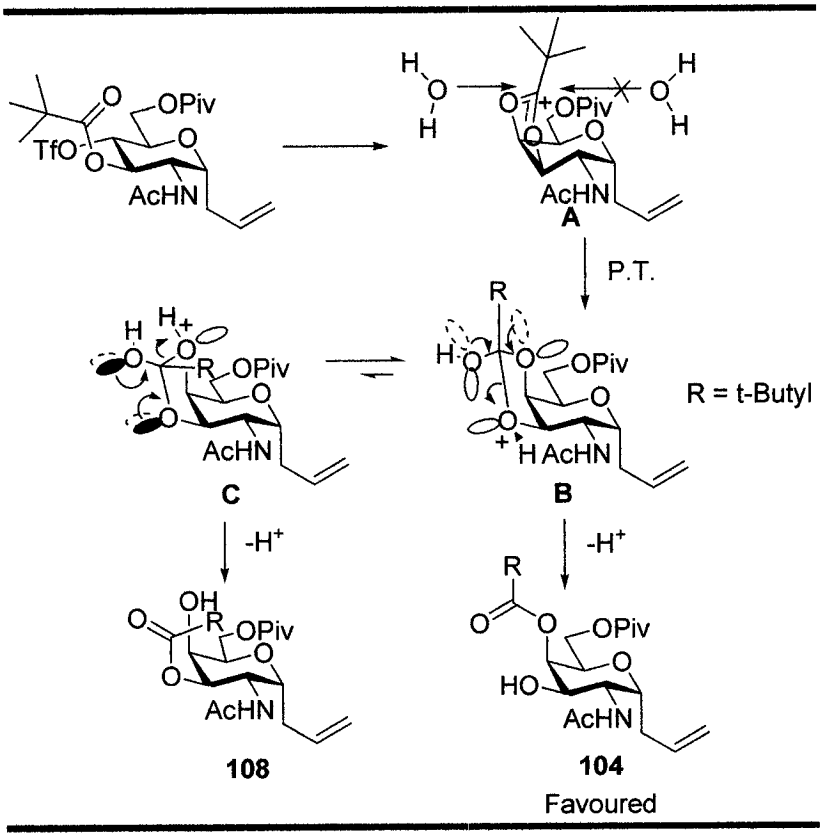
Scheme 12: Synthesis of the C-linked 2-deoxy-2-acetamido-α-D-galacto-pyranose derivatives 103 and 107



Several precautions must be taken for this reaction, in that all the reactants and the solvents have to be freshly distilled and extremely dry. Inversion of the C-4 hydroxyl group was accomplished by treatment with triflic anhydride followed by stirring in water overnight. This led to *C*-linked 2-acetamido-α-D-galactose derivative **104** in 6 steps and an overall yield of 44%.

The inversion of the C-4 hydroxyl group has been studied using various adjacent esters[14] and is rationalized as outlined in Scheme 13. Intramolecular displacement of the triflate *via* the carbonyl oxygen of the pivaloyl group at C-3 produces dioxolenium ion intermediate **A.** Hydration occurs from the least hindered face to give hemi-orthoester intermediate **B** after proton transfer. Intermediate **B** collapses with the assistance of two primary stereoelectronic effects[30] resulting in the formation of axial ester **104** as the α-anomer following purification by column chromatography. While the corresponding equatorial ester **108** can be formed *via* intermediate **C**, this product is not observed due to severe steric interactions between the *t*-butyl group and the pyranose ring. In addition, collapse of intermediate **B** to form **104** furnishes an ester with the more stable "*Z*" conformation while collapse of intermediate **C** produces an ester **108** with the less stable "*E*" ester conformation.

Scheme 13: Stereoselectivity in the ring opening of hemi-orthoester intermediates B and C[30]



Preparation of orthogonally protected *C*-linked 2-deoxy-2-acetamido-α-D-galacto-pyranose derivative **107** is possible by conversion of **102** to benzylidene acetal **105** in 80% overall yield, *via* the benzylidene acetal intermediate (84% yield) and subsequent pivaloylation (95% yield). The diastereomeric mixture (α/β = 12:1) was used directly in the next step. Selective cleavage of benzylidene acetal[31] **105** using triethylsilane and boron trifluoride diethyl etherate ($BF_3.OEt_2$) afforded the α-anomer of **107** in 77% yield following purification by column chromatography (Table 1). The $BF_3.OEt_2$ method was preferred over the TFA or TfOH method because in the TFA and TfOH procedures several degradation products were observed making the purification challenging.

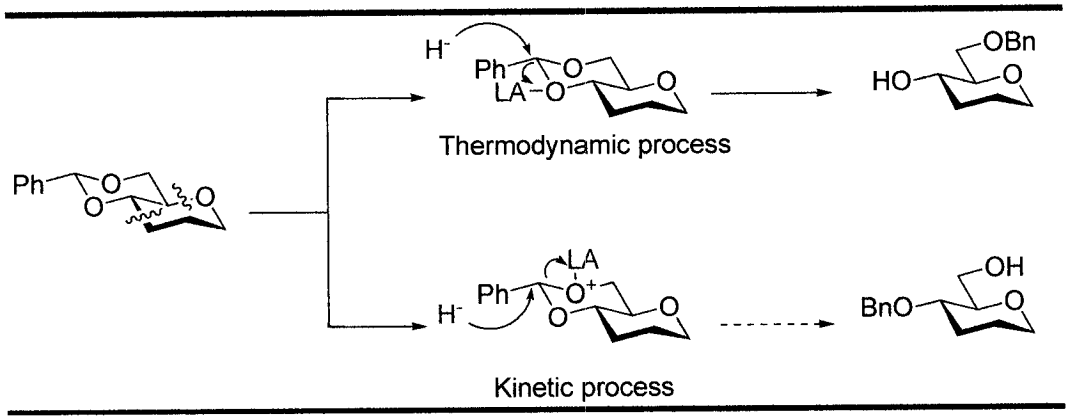
Table 1: Optimization of the benzylidene ring opening

Et_3SiH (equiv.)	Acid	Equiv.	T(°C)	Time (hours)	Yields
3	TfOH	1	-78	24	**(SM)**
3	TfOH	2	-78	24	**(SM)**
3	TfOH	3.4	-78	1	**74%**
3	TfOH	3.4	0→RT	1.5	**29%**
5	TFA	5	0(1h) + 0→RT	1h+7h	**70%+13%(SM)**
5	TFA	5	0→RT	12	**80%**
5	**$BF_3.OEt_2$**	**2**	**0→RT**	**2**	**77%**
12	$BF_3.OEt_2$	2	0→RT	8	**61%**

All the reaction were completed with 5 equivalents of triethylsilane

The selective benzylidene ring opening has been rationalized[32] by Samuelsson *et al.* as described in Scheme 14. Reaction conditions have been developed to favour the formation of either regioisomer.

Scheme 14: Conditions of the selective benzylidene ring opening[32]

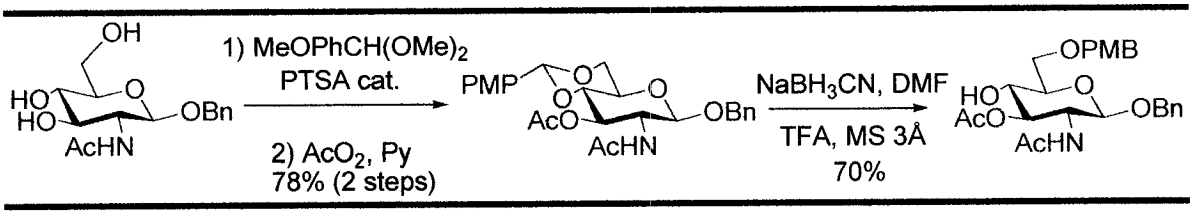


For example, reduction of 4,6-benzylidene-protected glucopyranosides with $NaBH_3CN$ under protic conditions produces 6-*O*-benzyl (OBn) ethers in high yield, whereas reductive cleavage with $i Bu_2AlH_3$ or $NaBH_3CN$ and TMSCl yields 4-*O*-benzyl ethers.

Regioselectivity is thought to be determined in the first case by differences in Lewis basicity of the acetal oxygens (O-4 > O-6) and in the second case by accessibility to the bulky Lewis acid (O-6> O-4). The Curtin-Hammett principle could also be used to predict the outcome of this ring opening reaction.

Interestingly, in the *O*-linked system the *para*-methoxybenzylidene acetal of glucose was synthesized and selectively opened to afford the *para*-methoxybenzyl ether in the C-6 position (Scheme 15).[33] This strategy is likely applicable to our system and would afford an alternative orthogonal protecting group.

Scheme 15: Synthesis of the C-6 para-methoxybenzyl ether of glucose[33]



Inversion of the C-4 hydroxyl group and migration of the pivaloyl protecting group was accomplished by treatment with triflic anhydride in refluxing dichloroethane/water to afford the orthogonally protected *C*-linked 2-deoxy-2-acetamido-α-D-galacto-pyranose derivative **106** in 83% yield (Table 2). This route required only 8 steps and produced **107** in 29% overall yield.

Table 2: Optimization of the C-3 inversion of compound 102

Entry	**Triflate Formation[a]**			**GalNAc Formation**		**Yields (%)**
	Solvents[b]	**T**	**Time (h)**	**Conditions**	**Time (h)**	
1	$EtCl_2$/Py	0°C to rt	1.5	H_2O/reflux DCE	0.5	**45**
2	CH_2Cl_2/Py	0°C	1	H_2O/reflux CH_2Cl_2	2	**57**
3	CH_2Cl_2/Py	0°C	1	H_2O/reflux CH_2Cl_2	2	**77**
4	**$EtCl_2$/Py**	**0°C**	**1**	**H_2O/reflux DCE**	**1.5**	**83**

a) All reactions were run using 1.25 equivalents of Tf_2O added dropwise at -20°C. b) All solvents are freshly distilled and the ratio of solvent:pyridine is 10:1

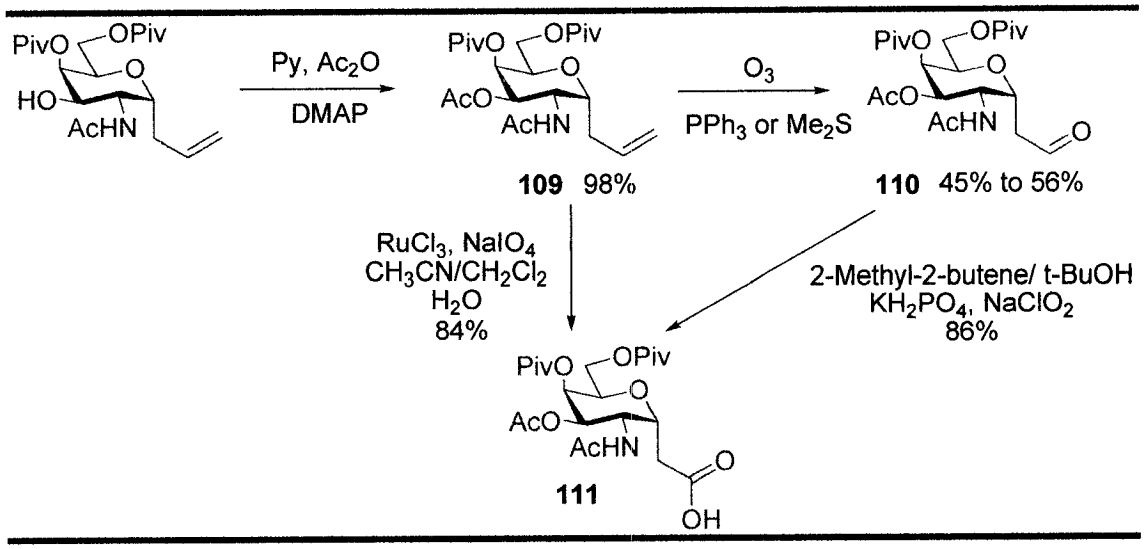
In summary, the synthesis of intermediates **104** and **107** can be accomplished in 8 and 6 steps respectively with overall yields of 29% and 44%. To date, the most efficient synthesis of **104** starting from D-glucose was carried out in fourteen steps and 11% overall yield[8a] and no syntheses using D-glucosamine have been reported. The synthetic sequence is amenable to large scale and consequently could be used to generate large quantities of **104** as well as differentially protected galactosamine derivative **107**. In both cases, a diastereomeric mixture (α/β anomers) is carried through until late into the synthesis where facile purification by column chromatography furnishes exclusively α products. Other attractive features of this approach include the fact that D-glucosamine is inexpensive, and all steps in the synthetic sequences are high yielding.

5.3.1.2 Oxidation and coupling to obtain C-linked GalNAc building block

We then decided to use the small amount of compound **104** that we synthesized in order to obtain the monosaccharide GalNAc analogues. Therefore we acetylated the hydroxyl group at the C-3 position to obtain **109**. This compound was then converted into the aldehyde **110** by ozonolysis and further oxidized to give the desired carboxylic acid **111** (Scheme 16).

As yields from the ozonolysis step were rather low and several degradation products were observed, we opted for a direct conversion of **109** to the desired carboxylic acid **111**. This was accomplished by treatment with catalytic ruthenium trichloride in the presence of sodium periodate in acetonitrile and water, which produced **111** in 84% yield.[34] Once we were in possession of the carboxylic acid synthon, we coupled it with the ornithine amino acid **AA7**. This amino acid was prepared using commercially available Fmoc-Orn(Boc)-OH, which was first benzylated and then treated with 50% TFA in CH_2Cl_2.

Scheme 16: Synthesis of the carboxylic acid synthon prior to coupling



Both reactions were high yielding and the final product was crystallized from ether. Coupling using HBTU and DIPEA afforded the desired compound **112** in 85% yield. Compound **112** was then debenzylated to afford the desired building block **113** that was ready for solid phase synthesis (Scheme 17).

Scheme 17: Final steps to obtain GalNAc building block

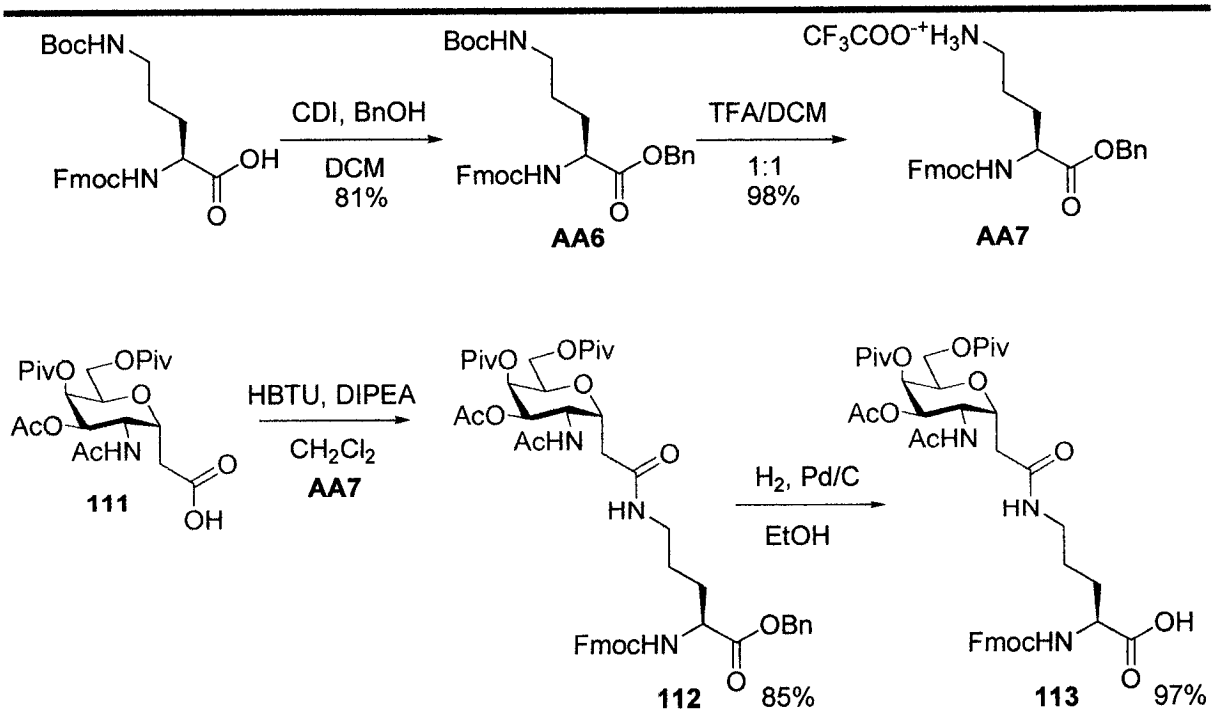


Table 3 reviews the yield and number of steps required in order to obtain a suitable GalNAc building block in order to synthesize our AFGP8 analogues.

Table 3: Summary of the synthesis of compound 113

***C*-Glycosylated Ornithine building block**	***Overall yield*** **(10 steps) from galactosamine hydrochloride**
FmocOrn(α-*C*-GalNAc)OH	**113 : 29%**

5.3.2 An approach to the synthesis of the disaccharide (1, 3)-β Gal-GalNAc building block

5.3.2.1 Synthesis of the (1, 3)-β Gal-GalNAc building block

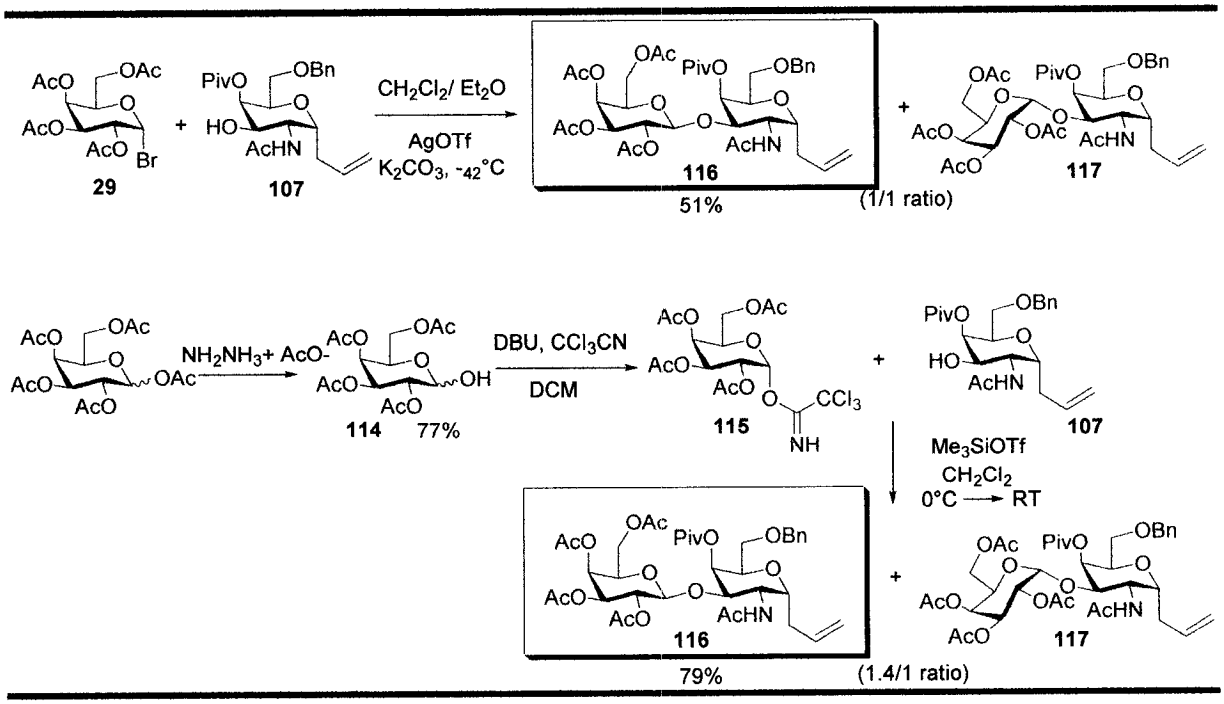
The difference in antifreeze activities between the native AFGP8 and its *O*-linked GalNAc analogue **81** are significant. The native AFGP8 exhibits an important TH activity and 3 times the RI activity than its monosaccharide analogue **81**. Moreover none of our *C*-linked analogues exhibit any TH activity. It was therefore essential to investigate the reasons for this low activity.

With the monosaccharide building block **113** in hand, we focused on the synthesis of the disaccharide building block **116** (Scheme 18). Our first attempt utilized the brominated galactose as a glycosyl donor. We expected that the anchimeric effect would favor the β diastereoisomer. However, in this case we obtained a 1:1 α:β mixture.

At this point it was impossible for us to identify the individual α/β anomers spectroscopically. The presence of several overlaps in the ^{1}H NMR spectrum prevented us from determining either the conformation or the configuration of these two molecules, furthermore we observe that the coupling constant of the anomeric protons of both anomer were similar making their differentiation impossible (H_1' = 6.02 ppm; 4.2Hz and H_1' = 6.05 ppm; 4.8 Hz). In the regular chair conformation (for galactose derivative) α-linkage results

in a coupling constant ranging from 3 to 6 Hz and β-linkage ranging from 7 to 10 Hz. The observed coupling constants confirm the fact that the comformation of these compounds were not chair-like. In theory a one dimensional TOCSI experiment should have allowed us to determine the conformation and therefore the configuration of the different compounds. However, the signals observed by this technique were too broad and it was impossible to obtain a reliable coupling constant value. We also tried to crystallize both compounds using slow diffusion techniques with hexane and dichloromethane but the crystals obtained were too fragile to survive the X-ray crystallography experiment. The stereochemistry of these two compounds were determined at a later step (refer to ozonolysis of **116** and **117** in section 5.3.2.1) and consequently the α/β ratio could be resolved.

Scheme 18: First attempt in order to obtain 116 with high diastereomeric excess



Following our first attempt, the reaction was then carried out at RT instead of 0°C. The yield was improved to 76%, but the α/β ratio, determined by NMR, increased to 1.5:1 in favor of the undesired α diastereoisomer. We thought at this point that perhaps the oxycarbenium intermediate was "stable" and that the product resulting from the anchimeric effect was not as favored as we had originally thought, even though in the previous coupling

for our β *O*-linked building blocks the ratio were much higher. Therefore, the trichloroimidate method using direct S_N2 displacement was employed. The coupling was attempted and the ratio and yield were slightly improved, but were still lower than expected. We attempted to perform the same reaction at RT but here again the ratio favored the α anomer.

Direct displacement of the trichloroimidate was expected to afford the β compound. Consequently several reactions with less polar solvents such as hexane were used in order to disfavor the formation of the oxycarbenium intermediate. Unfortunately this only slightly improved the α/β ratio in favor of the β diastereoisomer but also significantly decreased the yield of the reaction.

We also observed that these reactions were unusually slow. There are two reasons that could explain why these reactions were so slow compared to the similar β coupling in Chapter 4. First, the tetraacetate donors **29** and **115** were "disarmed" and therefore not very reactive and secondly, in the first coupling reactions in Chapter 4, acetates were in the C-4 position and were imposing less steric hindrance than the pivaloyl of the glycosyl acceptor **107**. Schmidt demonstrated[35] that longer times reaction tend to favor the formation of the α-anomer. Therefore, the fact that the glycosyl acceptor **107** is sterically hindered might slow down the reaction and increase the percentage of α- anomer observed.

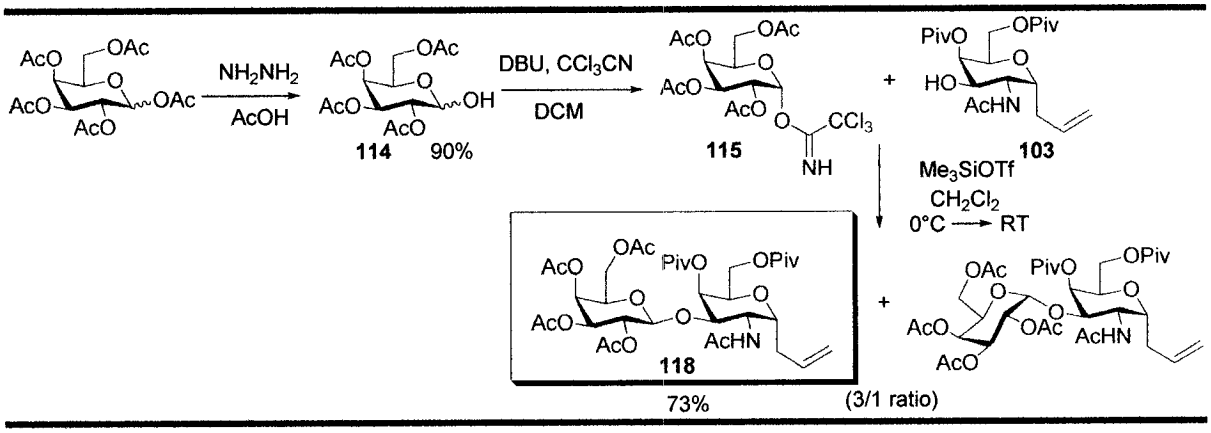
Expecting that the next step could allow us to determine the configuration of the two analogues, the ozonolysis of both disaccharides were performed. Triphenylphosphine was used as a reductor, but unfortunately it was impossible to fully separate the triphenylphosphine oxide from our compounds. However, these compounds allowed us to determine which compound was the α anomer (H_1' = 6.04 ppm; 4.2 Hz) and which compound was the β (H_1' = 6.01 ppm; 8.1 Hz). Since there was ^{1}H NMR and mass spectral evidence that we had the desired compound, we proceeded to the next step and oxidized the α analogue using $NaClO_2$, KH_2PO_4, 2-metyl-2-butene and *t*BuOH. Unfortunatly, several problems appeared. First, the triphenylphosphine oxide seemed to interfere with the reaction, and second, the oxidation conditions seemed too harsh for our compound and several hydrolyzed by-products were formed. We believe that either a $RuCl_3$-catalysed oxidation or an ozonolysis using dimethylsulphide followed by TEMPO or TPAP addition could solve the problem. However, at this stage the supply of starting material was low and

therefore we decided to synthesize more of **103** instead of **107** to obtain the disaccharide building block.

The synthesis of compound **103** was shorter than **107** and was also amenable to the synthesis of the AFGP native disaccharide unit. We then decided to use our optimized conditions for the synthesis of **118** in order to perform the coupling of **115** and **103**.

We observed that this coupling was much faster than the coupling of **115** with **107** and we also determined that using **115** as a glycosyl donor increased the yield as well as the α/β ratio in favor of the β diastereoisomer. The yield was 73% and the α/β ratio was 1/3.

Scheme 19: Coupling using 103 as glycosyl acceptor



The fact that a pivaloyl ester functionality on the C-6 hydroxyl changed the α/β ratio compared to a benzyl ether functionality might be because the conformations of the two acceptors are different. Therefore the hydroxyl group of the glycosyl donor **103** might be more accessible than for the glycosyl donor **107**. We expected that the yield and the α/β ratio could subsequently be improved by using a milder Lewis acid such as $BF_3.OEt_2$. The use of such Lewis acid should disfavor the formation of the oxocarbenium intermediate and hopefully improve our reaction conditions compared to the use of TMSOTf.

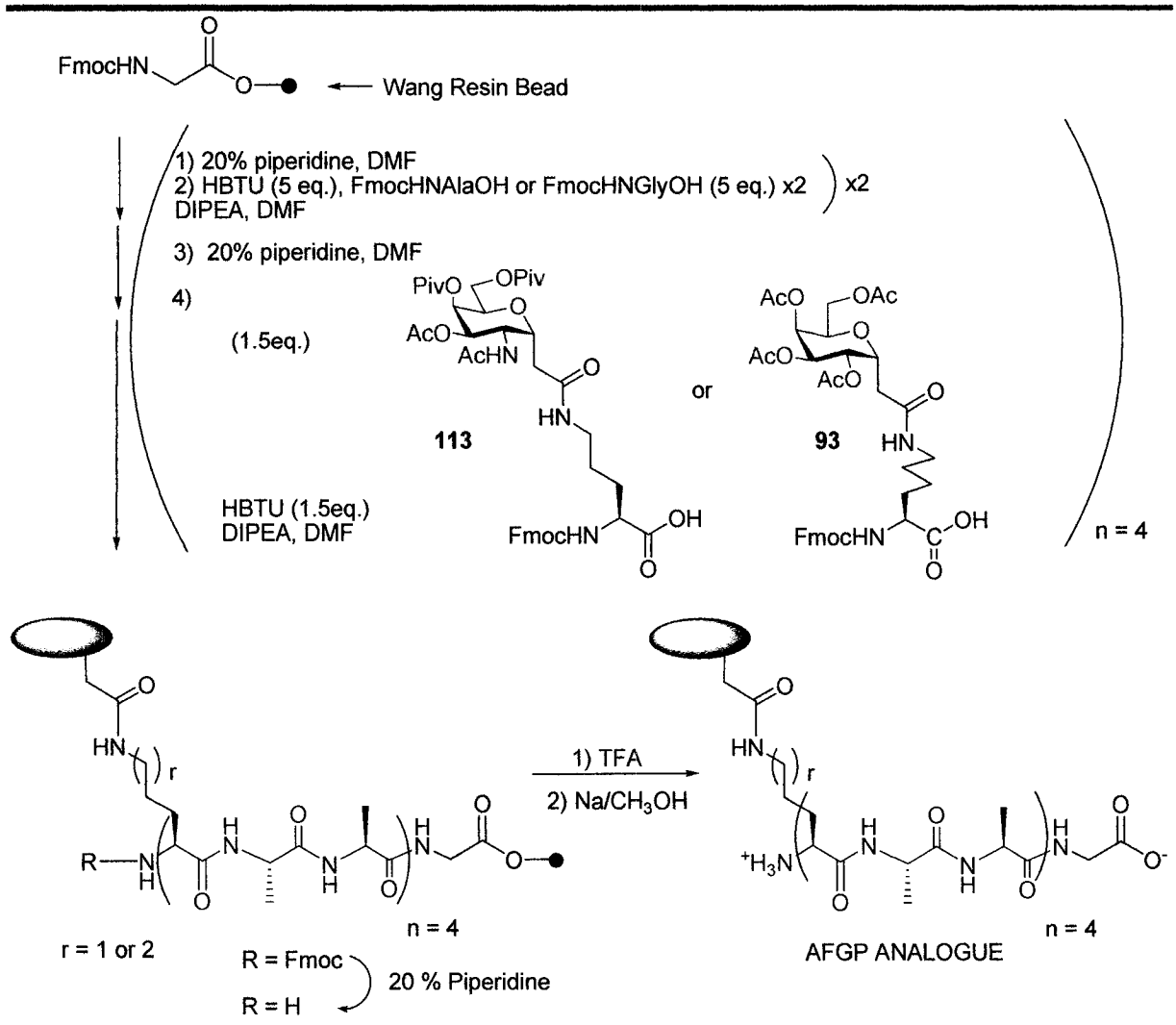
Once compound **118** was obtained, it was treated with $RuCl_3$ and $NaIO_4$ to afford almost no desired carboxylic acid. The hydrolysis of the glycosidic linkage was the main reaction. At this stage we decided to try an ozonolysis using dimethyl sulfide as the reducing agent. Unfortunately, the yield obtained was very low (25%) and several degradation products were observed. This low yield was probably due to the dimethyl sulfide being

oxidized to DMSO (verified by NMR). We were unfortunately out of starting material at this point and the amount of carboxylic acid was not sufficient to successfully carry out the coupling and obtain more than the necessary 200 mg for solid phase synthesis.

5.3.3 Solid phase synthesis of [LGG(Gal)]$_4$, [LAA(Gal)]$_4$, [OGG(GalNAc)]$_4$ and [OAA(GalNAc)]$_4$

C-Linked AFGP analogues were prepared by automated sequential assembly of building blocks **93** and **113** with *N*-Fmoc alanine or *N*-Fmoc glycine (Scheme 20). 20% piperidine in DMF solution (2h) was used to deprotect the Fmoc group at the α-amino position and HBTU was necessary for the couplings in order to generate the *in situ* activating reagent for amide bond coupling in the presence of diisopropylethylamine (DIPEA). The coupling reactions were repeated twelve times using either *N*-Fmoc alanine, *N*-Fmoc-glycine or building blocks **93** or **113** to afford the resin-bound polypeptides. The coupling with the commercially available amino acids was optimized using 5 equivalents of the amino acid and HBTU, and a double coupling technique (12h each). The building block coupling reactions were carried out with 1.5 equivalents of building block and HBTU, with a single coupling method of 48h. The solvent for coupling was also changed to NMP, which has been demonstrated to give better yields in coupling reactions. All the compounds were also solubilized in a HOBT solution in order to minimize racemization. After the last building blocks were assembled, the *N*-Fmoc groups were removed and the polypeptides were cleaved from the resin with 50% TFA in dichloromethane to furnish acetate protected analogues **119-122**. The Polypeptides were then precipitated from the reaction mixtures with diethyl ether, re-dissolved in water, and lyophilized in order to eliminate traces of TFA. Before further modifications, these analogues were purified by reverse phase HPLC. To afford the final *C*-linked AFGP analogues **123-126**, polypeptides **119-122** were treated with a basic solution of sodium methoxide (pH = 10) with one drop of water in order to avoid the methylation of the *C*-terminal acid, followed by neutralization with Amberlite IR-120H resin. Purification of the AFGP analogues was performed with reverse phase HPLC and the identities of the analogues were verified with MALDI mass spectrometry.

Scheme 20: Solid phase synthesis of C-linked analogues



119: $[Gal(OAc)_4](LGG)_4G$
120: $[Gal(OAc)_4](LAA)_4G$
121: $[GalNAc(OAc)(OPiv)_2](OGG)_4G$
122: $[GalNAc(OAc)(OPiv)_2](OAA)_4G$

MeOH / Na
1 drop H_2O

123: $[Gal(OH)_4](LGG)_4G$
124: $[Gal(OH)_4](LAA)_4G$
125: $[GalNAc(OH)_3](OGG)_4G$
126: $[GalNAc(OH)_3](OAA)_4G$

During these syntheses, we optimized the automated solid phase synthesis using different solvents, concentrations, sequences of addition, reaction times and mixing speeds. Optimizing these parameters allowed us to obtain excellent yields using the automated solid phase synthesizer. Unfortunately during the synthesis of $OGG(GalNAc)_4$, a computer system error occurred between the addition of HBTU and our amino acid. Five equivalents of HBTU were therefore mixed with the resin for 12h and we suspect that it capped off the resin resulting in the failure of the solid phase synthesis and the loss of our building block.

To avoid this problem, in the future we modified the sequence to add the amino acid before the addition of HBTU. Therefore if any error occurs, the amino acid would already be present in the well with the resin and no capping off should occur.

Table 4 summarizes the yield of the solid phase synthesis and the overall yields of all our *C*-linked AFGP8 analogues.

Table 4: Overall yields for the synthesis of our different C-linked AFGP8 analogues

***C*-Glycosylated Derivatives**	*Yields* **Solid Phase synthesis**	*Overall yields* **(From pyranose pentaacetate)**
Galactose$(LGG)_4G$: **123**	**85%**	**44%**
Galactose-$(LAA)_4G$: **124**	**90%**	**47%**
GalNAc-$(OGG)_4G$: **125**	**Failed***	**Failed***
GalNAc-$(OAA)_4G$: **126**	**91%**	**27%**

* Due to computer error during solid phase synthesis

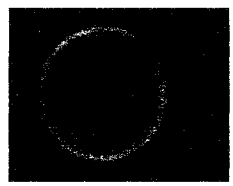
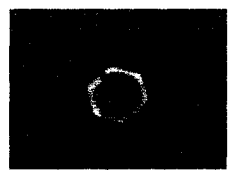
5.4 Physical and biological properties of our *C*-linked AFGP8 analogues

5.4.1 Thermal hysteresis and dynamic ice shaping activities

None of the *C*-linked analogues exhibited any difference between the melting point and the freezing point, which indicated that they do not possess any thermal hysteresis activity at 10 mg/mL. At this concentration AFGP8 already exhibits a significant TH gap of 0.12°C. At this concentration only $[LGG(Gal)]_4$ exhibited DIS activity within the TH gap, however, ice grew as circular disks at temperatures below the TH gap. Other *C*-linked analogues did not possess any dynamic ice shaping activity and therefore they did not have the ability to adsorb on the ice surface (Table 4). These interesting results confirm that our

C-linked analogues possess the adequate structural features for minimization of TH and DIS activity combined with the ability to exhibit potent RI activity.

Table 5: Dynamic ice shaping and thermal hysteresis of C-linked analogues

Analogues	Ice shaping	TH
Lys-Gly-Gly (Gal) 3-mer[36]		No TH
Lys-Gly-Gly (Gal) 4-mer : 123		**No TH**
Lys-Gly-Gly (Gal) 6-mer[36]		0.6°C
Lys-Gly-Gly (Gal) 9-mer[36]		0.6°C
Orn-Gly-Gly (Gal) 4-mer[36]		No TH
Lys-Pro-Gly (Gal) 4-mer[36]	No ice shaping	No TH
Lys-Ala-Ala (Gal) 4-mer : 124	**No ice shaping**	**No TH**
Orn-Ala-Ala (GalNAc) 4-mer : 126	**No ice shaping**	**No TH**

At this point in the discussion, our goals and the progress made will be reviewed.

a) *Determining the optimal length for minimal TH and DIS activities*

By comparing the TH and DIS activities of the [LGG(Gal)] family of analogues (3-mer, 4-mer, 6-mer and 9-mer), we can observe that the 4-mer analogue does not exhibit any TH activity compared to the 6-mer and 9-mer[36] that display an activity of approximately 50% of AFGP8 TH activity. This confirms what was observed in the *O*-linked analogues, in that the 4-mer seems to be of optimal length to obtain minimal TH activity while still displaying RI activity. The result that none of our 4-mer analogues exhibit any TH activity suggests that 4-mer analogues with monosaccharides are detrimental for TH activity.

b) *Investigating the influence of the length of the peptide side chain for our C-linked 4-mer analogues on TH and DIS activities*

By comparing the TH and DIS activities of the glycosylated ornithine $[OGG(Gal)]_4$ and lysine analogues $[LGG(Gal)]_4$, it was observed that modifying the length of the linker between the peptide backbone and the carbohydrate moiety by 1 carbon did not have any influence on the TH or DIS activity in these 4-mer analogues. Indeed, both analogues do not exhibit any TH activity and both of them display some interaction with the ice lattice by DIS activity. One plausible explanation for the fact that they display DIS activity and no TH activity is that we are at the threshold of the TH activity, and therefore it is possible that for longer analogues the length of the peptide side chain would have an influence on TH and DIS activities.

c) *Determining if the substitution with the proline in the peptide backbone $[LPG(Gal)]_4$ has any influence on TH or DIS activities*

Proline substituted AFGP8 was demonstrated to have the same antifreeze activities as its alanine counterpart.[37] By comparing $[LPG(Gal)]_4$ with $[LGG(Gal)]_4$, we desired to determine the influence of this proline insertion. We observed that proline substitution is detrimental for DIS activity for our *C*-linked analogues, as $[LPG(Gal)]_4$ did not exhibit any DIS activity while $[LGG(Gal)]_4$ does.

d) *Examining the role of the N-acetyl function in the C-2 position on TH and DIS activities.*

When compared to $[OGG(Gal)]_4$, the $[OGG(GalNAc)]_4$ analogue should have given us information about the influence of this *N*-acetyl function on our *C*-linked analogues.

Unfortunately the solid phase synthesis gave unsatisfactory characterization and we were unable to investigate the influence of the *N*-acetyl function.

e) *Examining the influence of other substitutions on the peptide backbone of C-linked AFGP analogues on TH and DIS activities.*

The peptide backbone of the native system is mainly composed of alanine, whereas our *C*-linked analogues contain glycine. Investigating the influence of alanine substitution on our *C*-linked analogues might explain the role of the alanine and its influence on TH and DIS activities in our *C*-linked analogues. Therefore, we tested the $[LAA(Gal)]_4$ analogue and compared it to its glycine counterpart $[LGG(Gal)]_4$. Our results indicated that alanine is detrimental for DIS activity in our *C*-linked analogues because a total loss of DIS was observed in $[LAA(gal)]_4$ compared to its glycine counterpart.

The effect of alanine substitution in the peptide backbone of our *C*-linked analogues should also be compared to its effect on our *O*-linked analogues to verify if any analogy could be made. Unfortunately $[SGG(Gal)]_4$ solid phase synthesis also failed and this analogy would have to be made in future work.

5.4.2 Recrystallization inhibition activity

Like for our *O*-linked analogues, six pictures were taken and the ten larger crystals of each picture were analysed. The SEM and *t*-value were also calculated for each sample, and the SEM were represented as error bar on Figure 6. The *t*-value of every of our 5.54 10^{-6} M samples were resulting in a significance above 99.5% (α level = 0.005).

The RI assay results of our *C*-linked galactosyl AFGP analogues **123, 124** and **126** are shown in Figure 7 and Table 6. Previous studies presented in Chapter 1 showed that the length of our polymers had an influence on TH and RI (Figure 6). The $[LGG(Gal)]_4$ analogues allowed us to observe that the 4-mer seems to be the minimal length in order to observe RI activity. TH and DIS activities do not appear to be significant for this length; therefore all of our subsequent analogues will be 4-mers.

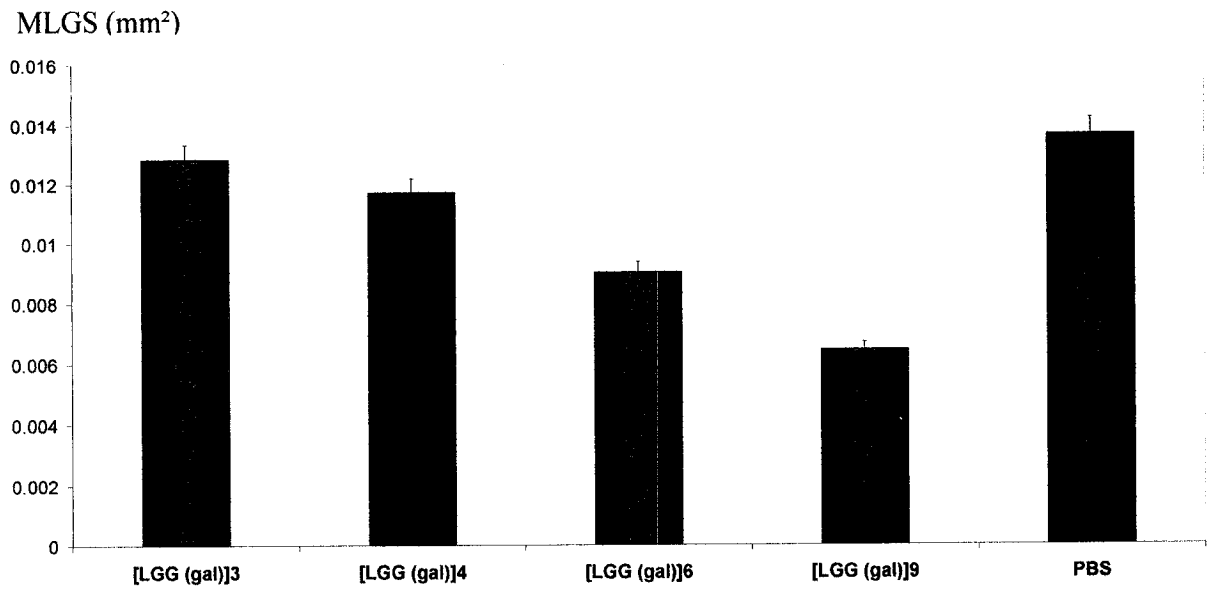


Figure 6: RI activities of [LGG(Gal)]n glycopolymers[36]

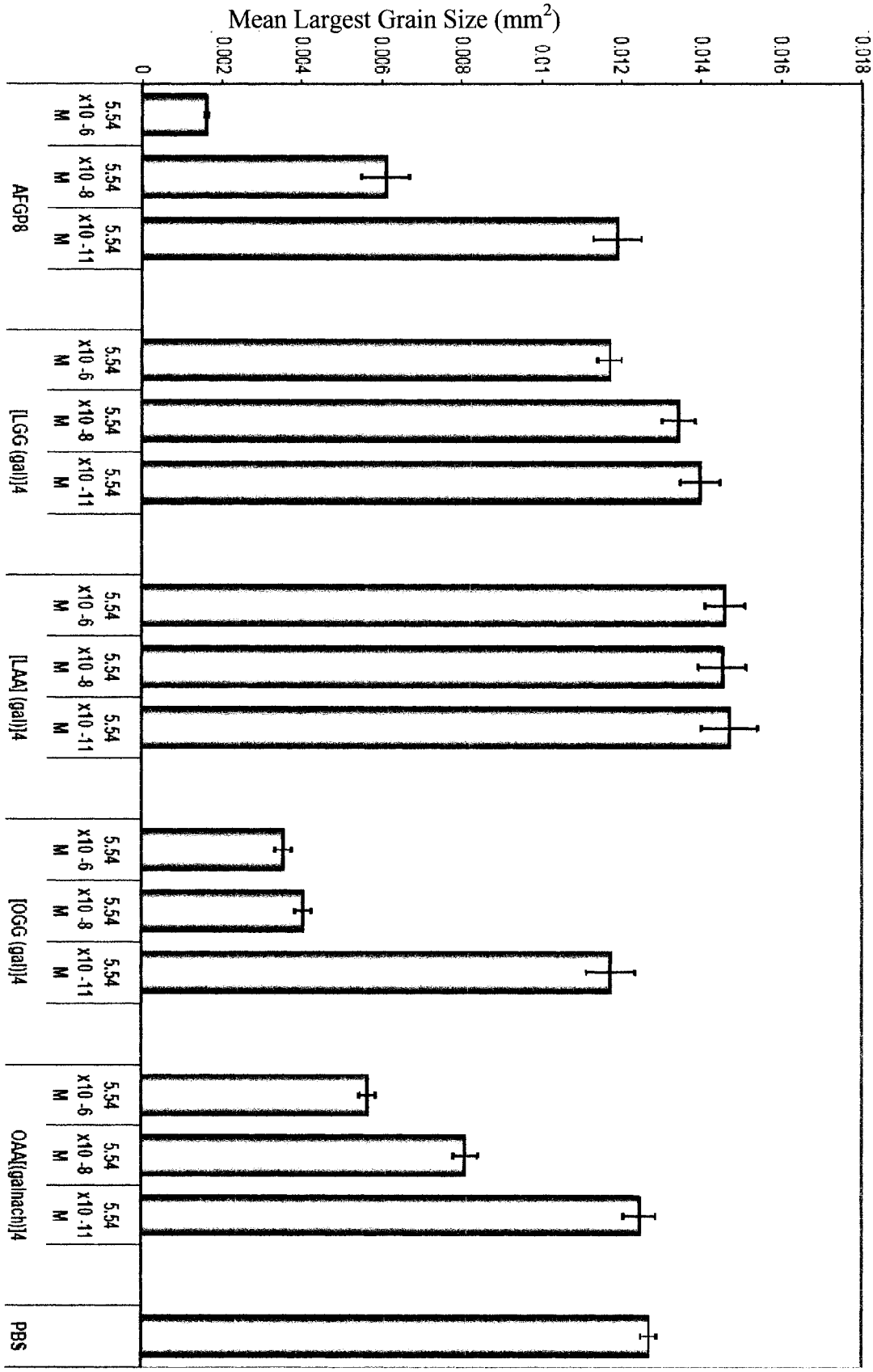


Figure 7: RI data for C-linked analogues

Table 6: Mean largest grain size of AFGP8, our C-linked analogues and PBS (surface area unit: mm^2).

Conc. (M)	AFGP8	$[LGG(Gal)]_4$	$[LAA(Gal)]_4$	PBS
5.54e-06	**0.0016±3.9e^{-5}**	**0.0117±0.0003**	**0.0146±0.0005**	**0.0127±0.0002**
5.54e-08	0.0061±0.0006	0.0134±0.0004	0.0145±0.0006	
5.54e-11	0.0119±0.0006	0.0150±0.0005	0.0147±0.0007	

Conc. (M)	AFGP8	$[OGG(Gal)]_4$	$[OAA(GalNAc)]_4$	$[LPG(Gal)]_4$
5.54e-06	**0.0016±3.9e^{-5}**	**0.0035±0.0002**	**0.0057±0.0002**	**0.0095±0.0003**
5.54e-08	0.0061±0.0006	0.0040±0.0002	0.0081±0.0003	0.0101±0.0003
5.54e-11	0.0119±0.0006	0.0117±0.0006	0.01247±0.0004	0.0121±0.0004

In order to improve RI activity versus TH or DIS activities, we will review and discuss our results for RI activity of our *C*-linked analogues.

a) Determining the optimal length of the peptide backbone for minimal TH and DIS activities and significant RI activity

As mentioned before, $[LGG(Gal)]_4$ exhibits a small but real non-specific RI activity and no TH. It displays approximately 8% of the activity of our negative control (PBS). Therefore the length of the peptide backbone seems to be one of the key structural features for TH activity. 4-mer analogues appear to possess the optimal peptide backbone length required for minimal TH activity and significant RI activity while using monosaccharide *C*-linked analogues.

b) Investigating the influence of the length of the peptide side chain for our C-linked 4-mer analogues on TH and DIS activities.

By comparing the RI activity of glycosylated ornithine $[OGG(Gal)]_4$ and lysine analogue $[LGG(Gal)]_4$, we observe that increasing the length of the linker between the peptide backbone and the carbohydrate moiety by one carbon seems to significantly decrease the RI activity of our 4-mer analogues. Indeed the ornithine analogue is approximately four times more active than its lysine counterpart.

Experiments conducted on $[LGG(Gal)]_4$ and $[OGG(Gal)]_4$ provided evidence that a shorter side chain,(*i.e.* ornithine), exhibited higher RI activity than longer side chains, (*i.e.* lysine). The peptide side chain of $[OGG(Gal)]_4$ brings the carbohydrate moiety closer to the

peptide backbone compared to $[LGG(Gal)]_4$. This proximity might favour some interactions between the carbohydrate moiety and the peptide backbone and therefore introduce some rigidity in the orientation of the carbohydrate moiety and the overall conformation of the $[OGG(Gal)]_4$. It is possible that having more rigidity is favourable for RI activity. One can hypothesize that the rigidity of the ornithine favours some specific orientation between all the carbohydrates of a single polymer. Consequently, this relative orientation might generate a carbohydrate cluster that could be favourable for ice interaction, whereas a higher degree of free rotation, i.e. lysine, would allow too many relative orientations which would compromise the ice interaction of this carbohydrate cluster.

A similar hypothesis was also proposed for the *C*-serine analogues previously prepared in our laboratory.[38] It was observed that some rigidity in the side chain between the carbohydrate and the peptide moieties was necessary in order to observe any RI activity. The rigidity of the side chain can directly affect the orientation of the carbohydrate moieties, and therefore the RI activity.

c) Determining if the substitution of proline in the peptide backbone $[LPG(Gal)]_4$ has any influence on RI activity.

Proline substituted AFGP8 has been demonstrated to have the same antifreeze activities as its alanine counterpart. By comparing $[LPG(Gal)]_4$ with $[LGG(Gal)]_4$, we desired to determine the influence of this proline insertion (Figure 8). We observed that proline substitution seemed to be slightly beneficial for RI activity in our *C*-linked analogues. Indeed $[LPG(Gal)]_4$ is approximately 20% more active than its glycine counterpart. Proline insertion seems to be very promising approach for the synthesis of future analogues, because it is detrimental for DIS activity and beneficial for RI activity.

Following the previous hypothesis, the rigidity added by the addition of proline to the peptide backbone seems to be favourable for the RI activity. It is possible that this additional rigidity provides the same ability to increase carbohydrate clustering as ornithine did compared to lysine in $[LGG(gal)]_4$ and $[OGG(gal)]_4$ analogues.

d) *Examine the influence of other substitutions on the peptide sequence on TH and DIS activities.*

As for TH activity, we tested the RI activity of $[LAA(Gal)]_4$ analogue and compared it to its glycine counterpart $[LGG(Gal)]_4$. It appears that alanine is detrimental for RI activity in our *C*-linked analogues because a total loss of RI is observed in $[LAA(gal)]_4$.

The effect of alanine substitution in the peptide backbone of our *C*-linked analogues should also be compared to its effect on our *O*-linked analogues to verify if any analogy could be made.

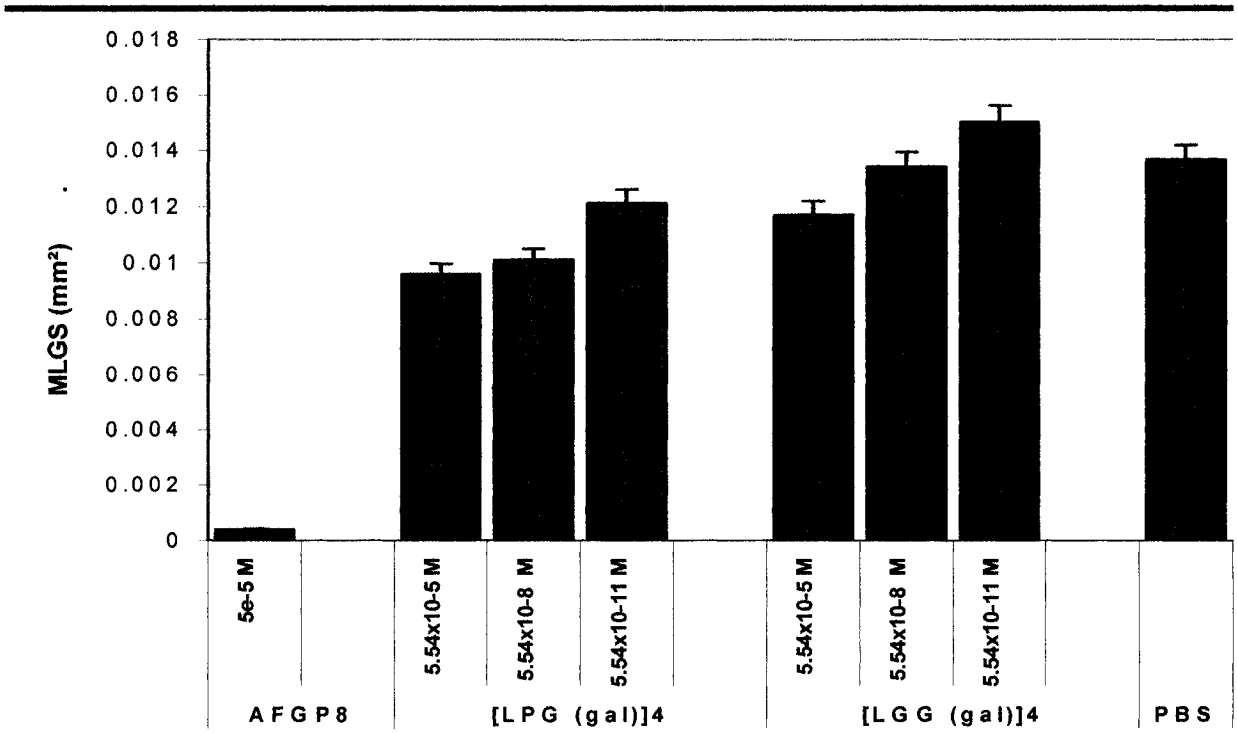


Figure 8: Comparison of the RI activity of $[LPG(Gal)]_4$[36] and $[LGG(Gal)]_4$

5.4.3 Conformational study of our analogues using Circular Dichroism (CD)

The CD spectra of our *C*-linked AFGP analogues were measured in water at 25 °C (Figure 9).

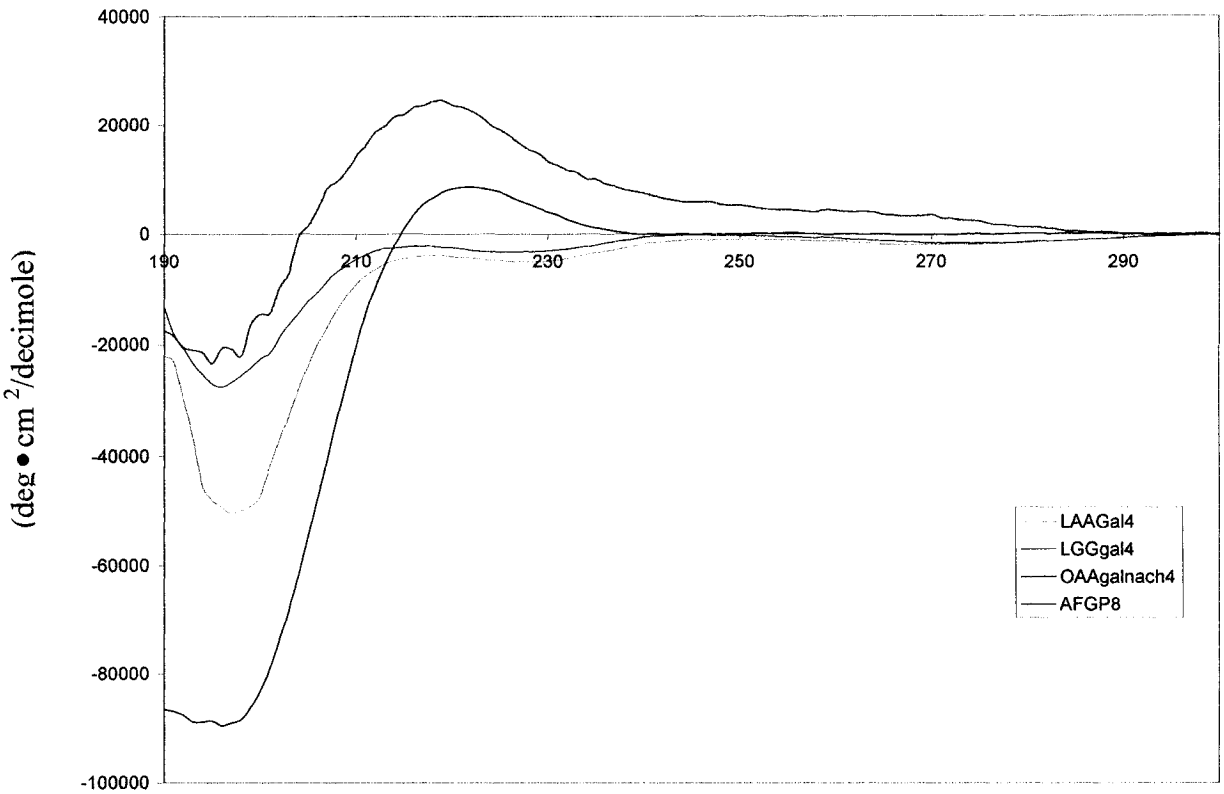


Figure 9: Circular dichroism of our C-linked analogues

All CD spectra were deconvoluted using the same software (SELCON3) and parameters (IBASIS3 and IBASIS4). In order to confirm the conformational trends that we could extract from CD spectra, two different parameters were used for this study: IBASIS3 was reported in red and IBASIS4 in blue.

Our *C*-linked analogues gave only partial convergence with IBASIS5 and therefore no satisfactory results were obtained using these parameters. Consequently IBASIS3 and IBASIS4 were used and gave satisfactory convergence with our analogues. IBASIS3 is composed of a set of reference proteins containing 37 proteins with regular and distorted α-helix, regular and distorted β-strand, β-turn and unordered conformations with optimal wavelength 185-240 nm; whereas IBASIS4 is composed of the same set of 37 proteins plus 6 proteins isolated and characterized by X-ray. The optimal wavelength of this reference set is between 190 and 240 nm. These two databases should and did afford very similar results.[39]

It is not possible to directly correlate the conformation obtained by CD deconvolution of our *C*-linked analogues with AFGP8 or our *O*-linked analogues. Indeed our *C*-linked analogues possess on the side chain an amide bond that will absorb and therefore perturb the CD spectra obtained. Several studies using glycosidic amide ligation on a peptide usually neglect the influence of this amide bond, however, in this case these bonds represent 1/5 of the amide bonds present on the molecule. Therefore, it is difficult to predict their relative influence on the CD spectra and on the deconvolution results.

However, by performing all the deconvolution calculations with the same set of reference proteins, useful information can still be gained on the conformational differences between the *C*-linked AFGP analogues. For want of anything better some conformational trends could also be cautiously extracted and maybe compared to AFGP8. The deconvolution analysis results of the *C*-linked AFGP analogues and AFGP8 are listed in Table 7.

Table 7: CD deconvolution of C-linked analogues using IBASIS3 and IBASIS4

Compound	Regular α-helix	Distorted α-helix	Regular β-strand	Distorted β-strand	β-turn	Random coil	Sum
[OAA (GalNAc)]$_4$	0.012	0.089	0.179	0.089	0.169	0.273	0.811
[OAA (GalNAc)]$_4$	0.008	0.085	0.179	0.093	0.170	0.290	0.834
[LAA(Gal)]$_4$	0.041	0.122	0.013	0.071	0.235	0.531	1.013
[LAA(Gal)]$_4$	0.001	0.087	0.024	0.062	0.219	0.605	0.998
[LGG(Gal)]$_4$	0.019	0.056	0.085	0.095	0.241	0.509	1.005
[LGG(Gal)]$_4$	-0.012	0.038	0.075	0.090	0.252	0.551	0.994

After analyzing the deconvolution data, it appears that our *C*-linked analogues exhibit a random coil conformation as their major conformation. However, several other conformations coexist at the same time. In the case of [LAA(Gal)]$_4$ even though all activity

seemed to be lost compared to $[LGG(Gal)]_4$, the alanine does not appear to have much influence on the conformation. An interesting pattern is observed with our most active compound: $[OAA(GalNAc)]_4$ compared to $[LAA(Gal)]_4$ and $[LGG(Gal)]_4$ displays a smaller population of random coil and exhibit a large population of β-strand which is a more ordered and rigid conformation. This preliminary CD spectroscopic study indicates that the conformation of our *C*-linked AFGP analogues might be an important factor, though not the only factor, for antifreeze-specific activities.

The trend observed for our $[OAA(GalNAc)]_4$ analogue is quite different for what was previously observed for the *C*-linked analogues prepared by Murphy.[36] Her analogues tended to display a very important population of random coil while displaying only small amounts of β-sheet and α-helix. This might be due to the fact that different parameters were used for the deconvolution or that a certain degree of rigidity is favorable but not necessary for RI activity and that some specific structural features could be sufficient for the RI activity to occur.

[1] Postema, M. H. D. *C-glycoside Synthesis*; CRC press: Boca Raton, **1995**.

[2] Levy, D. E.; Tang, C. *The Chemistry of C-glycosides*; Pergamon: Oxford, **1995**.

[3] Sinay, P. *Pure and Appl. Chem.* **1997**, 69, 459.

[4] Beau, J. M.; Gollagher, T. *Topics Curr. Chem.* **1997**, 187, 1.

[5] Nicotra, F. *Topics Curr. Chem.* 1997, 187, 55.

[6] Du, Y. Lindhardt R. J. *Tetrahedron* **1998**, 54, 9913.

[7] Dorgan, B. J., Jackson, R. F. W. *Synlett* **1996**, 859.

[8] Urban, D.; Skrydstrup, T.; Beau, J. *Chem. Commun.* **1998**, 955.

[9] (a) Dondoni, A., Marra, A., Massi, A. *Chem. Commun.* **1998**, 1741-1742.

(b) Dondoni, A., Marra, A., Massi, A. *J. Org. Chem.* **1999**, 64, 933.

[10] (a) Ben, R. N., Orellana, A., Arya, P. *J. Org. Chem.* **1998**, 63, 4817.

(b) Dondoni, A., Mariotti, G., Marra, A. *Tetrahedron Lett.* **2000**, 41, 3483.

(c) Dondoni, A., Mariotti, G., Marra, A. *J. Org. Chem.* **2002**, 67, 4475.

(d) Vincent, S. P.; Schleyer, A.; Wong, C.-H. *J. Org. Chem.* **2000**, 65, 4440.

[11] Roe, B. A., Boojamra, C. G., Griggs, J. L., Bertozzi, C. R. *J. Org. Chem.* **1996**, 61, 6442.

[12] (a) Lewis, M. D., Cha, J. K., Kishi, Y. *J. Am. Chem. Soc.* **1982**, 104, 4976.

(b) Kozikowski, A. P., Sorgi, K. L., Wang, B. C., Xu, Z.-B. *Tetrahedron Lett.* **1983**, 24, 1563.

(c) Hosomi, A., Sakata, Y., Sakurai, H. *Tetrahedron Lett.* **1984**, 25, 2383.

(d) Giannis, A., Sandhoff, K. *Tetrahedron Lett.* **1985**, 26, 1479.

(e) Horton, D. Miyake, T. *Carbohydr. Res.* **1988**, 184, 221.

(f) Panek, J. S., Sparks, M. A. *J. Org. Chem.* **1989**, 54, 2034.

(g) Bertozzi, C. R., Cook, D. G., Kobert, W. R., Gonzalez-Scarano, F., Bednarski, M. D. *J. Am. Chem. Soc.* **1992**, 114, 10639.

[13] (a) Keck, G. E., Enholm, E. J. *Tetrahedron* **1985**, 41, 4079.

(b) Giese, B., Linker, T., Muhn, R. *Tetrahedron* **1989**, *45*, 935.

[14] Keck, G.E.; Boden, E.P. *Tetrahedron Lett.* **1984**, 25, 265.

[15] Bertozzi, C. R., Bednarski, M. D. *Tetrahedron Lett.* **1992**, 33, 3109.

[16] Boons, G. J., Hale, K. J. *Organic Synthesis with Carbohydrates*, Blackwell Science: Malden, MA. **2000**.

[17] Keck, G. E., Yates, J. B. *J. Am. Chem. Soc.* **1982**, 104, 5829.

[18] Pontén, F., Magnusson, G. *J. Org. Chem.* **1996**, 61, 7463.

[19] Keck, G. E., Byers, J. H. *J. Org. Chem.* **1985**, 50, 5444.

[20] Eniade, A.; Murphy, A.V.; Landreau, G.; Ben, R.N. *Bioconjugate Chemistry* **2001**, 12, 817.

[21] For representative approaches see:

(a) Nicotra, F.; Russo, G.; Ronchetti, F.; Toma, T. *Carbohydr. Res.* **1983,** 124, C5.

(b) Giannis, A.; Sandhoff, K. *Carbohydr. Res.* **1987,** 171, 201.

(c) Carcano, M.; Nicotra, F.; Panza, L.; Russo, G. *J. Chem. Soc., Chem. Commun.* **1989,** 297.

(d) Grondin, R.; Leblanc, Y.; Hoogsteen, K. *Tetrahedron Lett.* **1991,** 32, 5021.

(e) Leteux, C.; Veyrieres, A. *J. Chem. Soc., Perkin Trans. 1,* **1994,** 2647.

(f) Kim, K.; Hollingsworth, R. I. *Tetrahedron Lett.* **1994,** 35, 1031.

(g) Hoffman, M.; Kessler, H. *Tetrahedron Lett.* **1994,** 35, 6067.

[22] Roe, B. A.; Boojamra, C. G.; Griggs J. L.; Bertozzi, C. *J. Org. Chem.* **1996**, 61, 6442.

[23] (a) Cipolla L.; La Ferla, B.; Lay, L.; Peri, F.; Nicotra, F. *Tetrahedron: Asymmetry* **2000**, 11, 295.

(b) Cipolla L.; Lay, L.; Nicotra, F. *J. Org. Chem.* **1997**, 62, 6678.

[24] For representative approaches see:

(a) Lemieux, R. U.; Ratcliffe R. M. *Can. J. Chem.* **1979**, 57,1244

(b) Bovin, N. V.; Zurabyan, S. E., Khorlin, A. Y. Carbohydr. Res. **1981**, 98, 25.

(c) Czernecki, S.; Randriamandimby, D. *Tetrahedron Lett.* **1993**, 34, 7915.

(d) Czernecki, S.; Ayadi, E.; Randriamandimby, D. *J. Org. Chem.* **1994**, 59, 8256.

(e) Santoyo-Gonzalez, F.; Calvo-Flores, F. G.; Garcýa-Mendoza, P.; Hernandez-Mateo, F.; Isac-Garcýa, J.; Robles-Dýaz, R. *J. Org. Chem.* **1993**, 58, 6122.

(f) Giuliano, R. M.; Davis, R. S.; Boyko, W. J. *J. Carbohyd. Chem.* **1994**, 13, 1135.

(g) Dondoni, A.; Mariotti, G.; Marra, A. *J. Org. Chem.* **2002**, 67, 4475.

(h) Bertozzi, C. R.; Bednarski, M. D. *Tet. Lett.* **1992**, 33, 3109.

(i) Hoffmann, M. G.; Schmidt R. R. *Liebigs Ann. Chem.* **1985**, 2403.

(j) Burkhart, F. and Kessler, H. *Tet. Lett.* **1998**, 39, 255.

(k) Urban, D.; Skrysdstrup T.; Beau, J.-M. *J. Org. Chem.* **1998**, 63, 2507.

(l) Shäfer, A.; Thiem, J. *J. Org. Chem.* **2000**, 65, 24.

(m) Rubinstenn, G,; Esnault, J.; Mallet, J.-M.; Sinaÿ, P. *Tetrahedron: Asymmetry* **1997**, 8, 1327.

(n) Sanmartin, R.; Tavassoli, K. E.; Walsh K. E.; Walter, D. S.; Gallagher, T. *Org. Lett.* **2000**, 2, 4051.

(o) Grant, L.; Liu Y.; Walsh, K. E.; Walter, D. S.; Gallagher T. *Org. Lett.* **2002**, 4, 4623.

[25] McGarvey, G. J.; Schmidtmann, F. W.; Benedum, T. E.; Kizer, D. E. *Tet. Lett.* **2003**, 44, 3775.

[26] (a) Keck, G. E.; Yates, J. B. *J. Am. Chem. Soc.* **1982**, 104, 5829.

(b) Cui, J.; Horton, D. *Carbohydr. Res.* **1998**, 309, 319.

[27] (a) Tarasiejska, Z.; Jeanloz, R. W. *J. Am. Chem. Soc.*; **1958**; 80; 6325.

(b) Heidlas, J. E.; Lees, W. J.; Pale, P.; Whitesides, G. M. *J. Org. Chem.* **1992**, 57, 146.

(c) Horton D. *Methods Carbohydr. Chem.* **1972**, 6, 282.

(d) Baker, B. R.; Joseph, J.; Schaub, R. E.; Williams, J. H. *J. Org. Chem.* **1954**, 19, 1786.

[28] Ponten, F.; Magnusson, G. *J. Org. Chem.* **1996**, 61, 7463.

[29] (a) Binkley, R. W.; Sivik, M. R. *J. Org. Chem.* **1986**, 51, 2621.

(b) Binkley, R. W. *J. Org. Chem.* **1991**, 56, 3892.

(c) Lay, L.; Nicotra, F.; Panza, L.; Russo, G. *Helv. Chimica. Acta.* **1994**, 77, 509.

(d) Lubineau, A.; Bienayme, H. *Carbohydr. Res.* **1991**, 212, 267.

(e) Nashed, M. A.; El-Sokkary, R. I.; Rateb, L. *Carbohydr. Res.* **1984**, 131, 47.

[30] Deslongchamps, P. Stereoelectronic Effects in Organic Chemistry. Pergamon Press Ltd., Oxford, **1983**, p83.

[31] (a) DeNinno, M. P.; Etienne, J. B.; Daplantier, K. C. *Tet. Lett.* **1995**, 36, 669.

(b) Debenham, S. D.; Toone, E. J. *Tetrahedron Asym.* **2000**, 11, 385.

(c) Barroca, N.; Jacquinet, J.-C. *Carbohydr. Res.* **2002**, 337, 673.

[32] Johansson R., Samuelsson B., *Chem. Commun.* **1984**. 201.

[33] Lubineau A.; Bonnaffé D. *Eur. J. Org. Chem.* **1999**, 10, 2523.

[34] Buskas, T.; Soderberg, E.; Konradsson, P.; Fraser-Reid, B. *J. Org. Chem.* **2000**, 65, 958.

[35] Schmidt R.R. *Modern Methods in Carbohydrate Synthesis*, Khan, S.H.; O'Neill, R.A. eds, Harwood Academic, Netherlands, **1996**, p20.

[36] Murphy, A.V. Ph.D. thesis SUNY Binghamton, Binghamton **2004**.

[37] Geoghegan, K.F.; Osuga, D.T.; Ahmed, A.I.; Yeh, Y. Feeney, R.E. *J. Biol. Chem.* **1980**, 255, 663.

[38] For more detail see Chapter 1 regarding our *C*-linked serine analogues

[39] The list of the proteins used by this reference set is described at http://lamar.colostate.edu/~sreeram/CDPro.

FUTURE WORK

The SAR studies have identified several different structural features necessary to obtain RI or TH activity in *O*-linked or *C*-linked AFGP8 analogues. However, several other structural modifications need to be investigated in order to optimize our *C*-linked analogues for cryopreservation. Toward this end, future work should aim to achieve the following:

Understanding of the role of the peptide backbone alanines vs glycines on TH and RI activity. Our *C*-linked analogues substituted the alanine present in native AFGP8 with glycine and exhibited interesting antifreeze activity. It would be interesting to see the effect of glycine substitution in the peptide backbone of our *O*-linked analogues by comparing $[SAA(Gal]_4$ and $[SGG(Gal)]_4$ for RI and TH activity. We should also compare the effects obtained on our *O*-linked analogues to our *C*-linked analogues, ($[LGG(Gal)]_4$ and $[LAA(Gal)]_4$), in order to determine if any correlation could be made. The effect of glycine substitution could also be investigated on our most active *O*-linked analogues. Therefore $[TGG(GalNAc)]_4$ would need to be synthesized and compared to $[TAA(GalNAc)]_4$.

Examination of the influence of the carbohydrate exo-cyclic oxygen on RI and TH activities. Knowing the significant RI activity of *C*-$[SGG(Gal)]_4$, it would be interesting to compare this *C*-linked analogue with an equivalent *O*-linked analogue. By synthesizing $[SGG(Gal)]_4$ we could compare the two directly. Comparison of the two analogues would allow us to determine whether or not the unusual RI activity observed with the *C*-$[SGG(Gal)]_4$ is due to the relative hydrophobicity of the CH_2 compared to the exo-cyclic oxygen of the *O*-linked analogues or if the peptide backbone modification is the main reason for this unexpected increase in RI activity.

Examination of the role of the peptide side chain amide bond on RI activity. The fact that the *C*-$[(S+2)GG(Gal)]_4$ analogue does not exhibit any RI activity (Chapter 1, Graph 7), while $[OGG(Gal)]_4$ is one of our most active analogues, suggests that the amide bond present in our $[OGG(gal)]_4$ analogue is involved with the occurrence of RI activity. In order to understand what feature of the amide bond favors RI activity several analogues would need

to be synthesized and tested (Figure 1). Compound B and C will give information about the importance of the amine and the carbonyl, compound D will indicate if the position of this amine and carbonyl are relevant and finally compound E will allow the investigation of the geometry of the amide bound. Finally all these analogues allow us to fully understand the role of the carbonyl, the amine, and the geometry of the amide bond on RI activity.

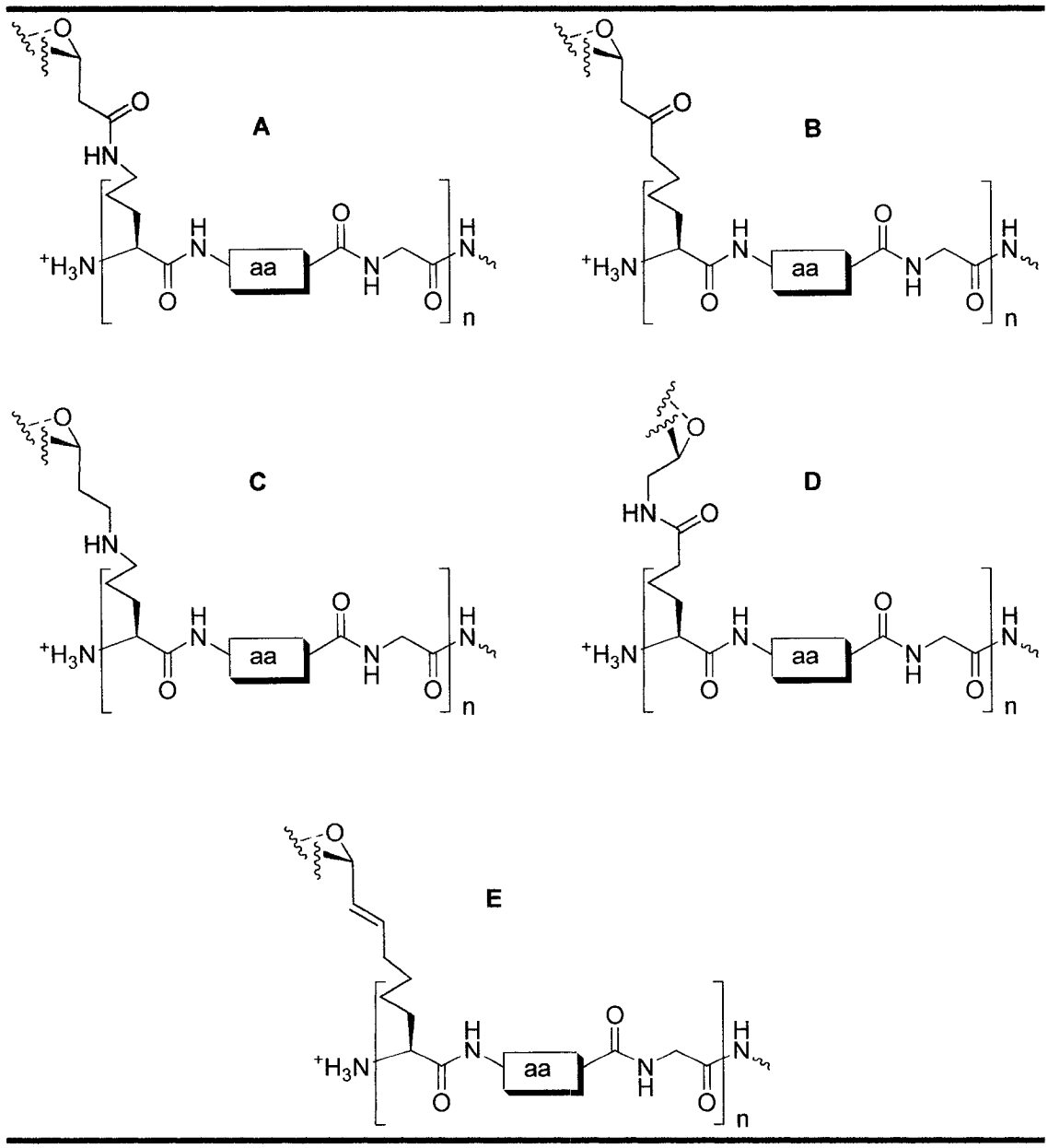
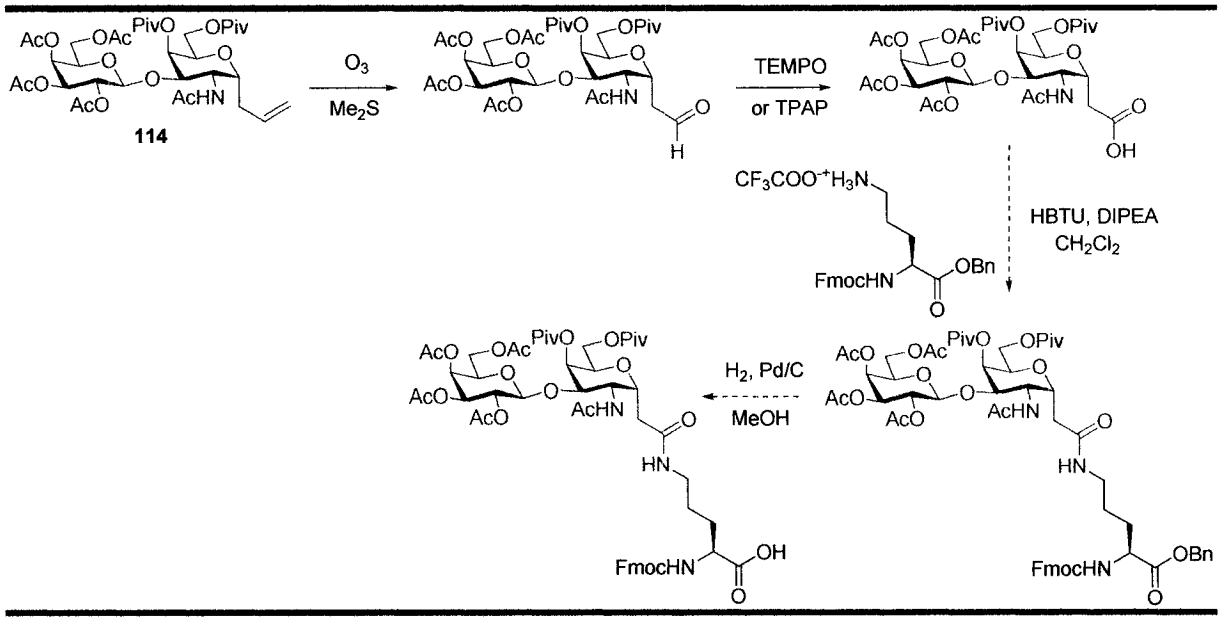


Figure 1: Different AFGP8 analogues with side chain modifications designed to be compared to our 1st generation compounds

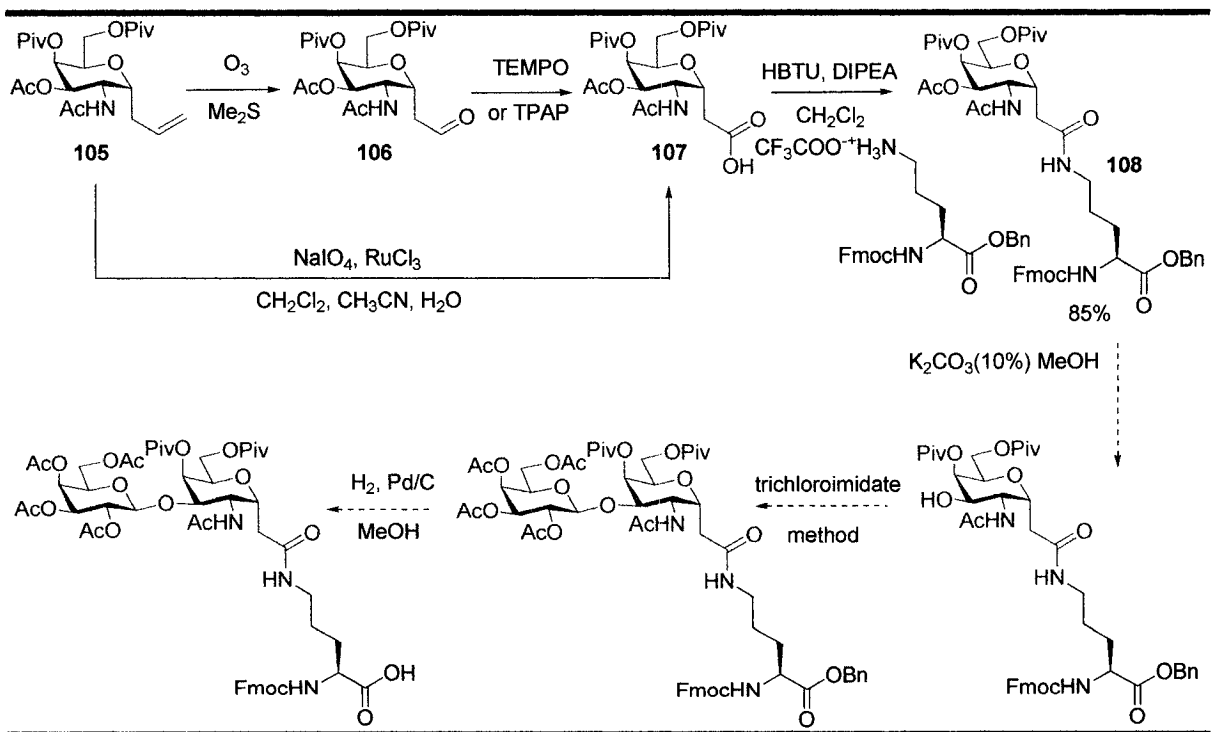
Synthesis and analysis of [OGG(GalNAc)]$_4$ and [OPG(GalNAc)]$_4$. The presence of alanine residues seemed to decrease RI activity when [LGG(Gal)]$_4$ was compared to [LAA(Gal)]$_4$. Therefore, because [OAA(GalNAc)]$_4$ is exhibiting considerable RI activity: it would be necessary to prepare the analogue [OGG(GalNAc)]$_4$ to verify if the alanine trend observed with [LAA(Gal)]$_4$ is followed. If this is the case, [OGG(GalNAc)]$_4$ could particularly active for RI activity. Moreover, the presence of a proline in [LPG(gal)]$_4$ seems to increase RI activity when compared to [LGG(gal)]$_4$. Insertion of a proline residue in the peptide backbone of an analogue of [OGG(GalNAc)]$_4$ ([OPG(GalNAc)]$_4$), should also be attempted.

Synthesis and analysis of [OGG(β-(1, 3)-Gal-GalNAc)]$_4$ analogues. The synthesis of *C*-linked analogues containing the native AFGP disaccharide moiety as well as the synthesis of disaccharide analogues with different carbohydrate moieties will be completed in the Ben laboratory. Two pathways to obtain the necessary building blocks for the [OGG(β-(1, 3)-Gal-GalNAc)]$_4$ analogues are possible. One approach is to oxidize the disaccharide **114** using dimethyl sulfide and ozone, followed by a TEMPO or TPAP oxidation to obtain the carboxylic acid functional group necessary for coupling with the amino acid. An alternate approach is to attempt the oxidation and coupling of the carbohydrate moiety with the amino acid using a monosacharide residue that would later be transformed to a disaccharide (Scheme 1 and Scheme 2).

Scheme 1: First possible pathway to obtain a disaccharide analogue building block



Scheme 2: Second possible pathway to obtain a disaccharide analogue building block



General Experimental Methods:

In this experimental section several compounds that have been previously reported in literature are presented and described again. The reasons for this are that either the methodology in order to obtain these compounds has been improved or that different complimentary characterization data have been added. Several syntheses present in this thesis will have to be reproduced for complementary studies in our laboratory and even though NMR data for some of our building block are described in literature no spectra are available, therefore for conveniency these spectra will be added.

Infrared (IR) spectra were recorded using a FTIR-8400 S, Fourier transform infra red spectrophotometer (Simadzu) and Arid Zone™ MB Series ABB Bomen interfaced with a Dell computer. ^{1}H and ^{13}C NMR spectra were recorded using a Bruker Avance 500 or 300 MHz in the appropriate deuterated solvent. Chemical shift values are reported in ppm downfield from TMS as an internal standard. Multiplicities are reported as s, singlet; bs, broad singlet; d, doublet; dd, doublet of doublets; t, triplet; m, multiplet. Low-resolution mass spectra were obtained using a Kratos Concept 2H Analytical (CI, EI) or a Micromass Quattro LC (ESI) mass spectrometer. High-resolution mass spectra were obtained using a Kratos Concept Analytical high-resolution mass spectrometer using sodium iodide as an internal standard. Chromatographic separations were performed using 230-400 Mesh silica gel from Natland Inc.. All solvents were dried and distilled prior to utilization; THF was distilled over sodium-benzophenone and methylene chloride was distilled over calcium hydride. N,N dimethylformamide (DMF) was dried with activated molecular sieves and stored under argon atmosphere. All solution phase reactions were performed under an argon atmosphere unless otherwise stated. Solution phase reactions were monitored by thin layer chromatography (TLC), and visualized with UV lamp (254 nm) and stained with i) ceric ammonium molybdate (CAM), and ii) dinitrophenyl hydrazone (DNPH). **Proper drying of silica gel is important because of the instability of Pfp esters on wet silica. The silica was dried at 125°C for 48h.**

General protocol for the manual solid phase synthesis of O-linked and C-linked analogues

80 mg of Fmoc-Glycine Wang resin (0.047 mmol) was swollen in 5 mL of DMF for 20 minutes. After this time, the solvent was drained and 20% piperidine solution in DMF was added. The solution was allowed to stir for 2 hours, drained and the resin was washed with three aliquots of DMF (20 mL). The resin was then tested for free amine using Kaiser and TNBS tests. Alanine or glycine (2 equivalents) were premixed with HBTU (1.8 equivalents) and DIPEA (0.01 mL) in DMF. The reaction was stirred for 30 minutes and then transferred to the SPPS flask and stirred for 8 hours. The flask was drained, the resin was rinsed three times with DMF (45 mL total), and Kaiser and TNBS test were performed to verify the coupling had reached completion (negative test outcome). The resin was treated with 20% piperidine solution in DMF for 2 hours, the flask was drained and the resin was washed thoroughly with DMF (50 mL). Kaiser and TNBS test were performed to ensure the presence of free amine. The previous coupling was repeated a second time with the desired amino acid and at the third coupling our building block was added. The building block (1.5 equivalents) was premixed with 1.4 equivalents of HBTU in DMF, followed by the addition of 0.01 mL of DIPEA, in 3 mL DMF for 30 minutes. For the pentafluorophenyl ester analogues this preactivation is not necessary and the compound can be directly transferred to the SPPS flask. The SPPS is let to stir for 16 hours, rinsed and tested using Kaiser and TNBS tests. At this time, we have a tripeptide core unit linked to the resin. The procedure was repeated till the desired length for the peptide was obtained. To cleave the glycopeptide from the resin, the resin was rinsed with dichloromethane (20 mL), and stirred for 2 hours in 1:1 (v:v) trifluoroacetic acid and dichloromethane. The solution was filtered and concentrated and the product crystallized from diethyl ether to produce a white powder. The glycopolymer was dissolved in a sodium methoxide (pH 10) solution (4 mL sodium methoxide: 1 mL water) and stirred for 5 hours if the protecting groups are acetate. The solution was acidified with IR-120 ion exchange resin, filtered, concentrated to remove methanol and lyophilized to give product as a white powder. Purification using reverse phase chromatography (C-18) was then performed to obtain our glycopeptide. If the protecting groups are benzyl ether a hydrogenolysis is performed using, H_2, Pd in EtOAc or EtOH depending of the solubility of the polymer and the reaction is stirred for 48 hours.

Kaiser test: Two drops from each of the following solutions were added to a sample of a few dry resin beads in a vial:

1) 5 g of ninhydrin in 100 mL of ethanol

2) 80 g of phenol in 20 mL of ethanol

3) 2 mL of 0.001 M aqueous solution of potassium cyanide in 98 mL of pyridine

The vial was put in an oven at 120 °C for 5 minutes. A positive test was indicated by the appearance of a dark blue solution.

TNBS test: One drop of each of the following solutions was added to a sample of a few dry resin beads in a vial.

1) 10 % DIPEA in DMF

2) 1 % TNBS in DMF

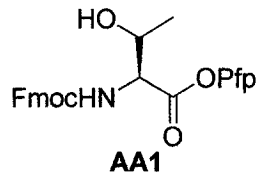
The vial was shaken and left standing at room temperature for 5 minutes. The solution was diluted with 1mL of DMF. A positive test was indicated by the dark orange color of the beads.

General protocol for the automated solid phase synthesis of O-linked and C-linked analogues

A similar protocol was used for the automated solid phase synthesis. Several modifications were carried to the liquid phase protocol in order to improve the yields. NMP was used as solvent for the coupling steps instead of DMF and our coupling procedure involved a double coupling. For each amino acid added to the resin, we performed a first coupling with 5 equivalents of the commercially available amino acid and let the reaction stirred for 12 hours. Then the resin is washed and instead of proceeding to the deprotection of the Fmoc protecting group, we performed a second coupling with 5 equivalents of the same amino acid. This coupling technique improved the yield of each coupling. For the coupling of our building block a single coupling with 1.5 equivalents was performed and the reaction was stirred for 48 hours. The chemicals used for the coupling, deprotection and cleavage from the resin as well as the procedures were the same than for manual SPPS.

The following amino acid derivatives have been previously described in the literature or are commercially available but are important building blocks used in our laboratory and are therefore described for conveniency.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-L-threonine-pentafluorophenyl ester (AA1)[1]

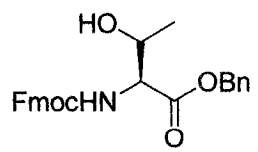


AA1

This compound was synthesized according to a slightly modified procedure of Istvan *et al.*.[1] To a solution of FmocThrOH (5 mmol) and pentafluorophenol (1.64 g, 5 mmol) in dry DMF (30 mL), DCI (5 mmol) was added at 0°C. The reaction was stirred at 0°C for 1 hour and 2 hours at 25°C. Dicyclohexylurea was filtered off and the solvent was evaporated. The residue was triturated with hexane, filtered and recrystallized in EtOAc/hexane to afford a white powder (1.98 g, 3.9 mmol, 78%).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.73 (d, J = 7.5 Hz, 2H), 7.60 (d, J = 5.4 Hz, 2H), 7.36 (t, J = 7.2, 2H), 7.27 (t, J = 7.3, 2H), 6.16 (d, J = 9.3 Hz, 1H), 4.70 (dd, J_1 = 2.1 Hz, J_2 = 9.3 Hz, 1H), 4.53-4.45 (m, NH+2H), 4.22 (t, J = 6.9Hz, 1H), 2.37 (bs, OH), 1.33 (d, J = 6.3 Hz, 3H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 167.8, 157.4, 144.0, 143.9, 143.8, 139.9, 139.8 (m), 136.7 (m), 128.2, 127.9, 127.5, 125.4, 125.1, 120.4, 68.2, 67.9, 59.7, 47.4, 20.2. LRMS (ES, NH_4^+) m/z: calcd for $C_{25}H_{18}O_5NF_5$ ($M+K^+$), 546.1; found, 546.2.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-L-threonine-benzyl ester (AA2)[2]



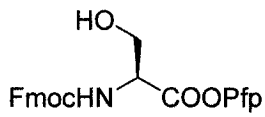
AA2

This compound was synthesized according to a slightly modified procedure of Istvan *et al.*.[1] To a solution of FmocThr(*t*Bu)OH (2.5 mmol) and benzyl alcohol (2.5 mmol) in dry

DMF (20 mL), DCI (2.5 mmol) was added at 0°C. The reaction was stirred at 0°C for 1 hour and 2 hours at 25°C. Dicyclohexylurea was filtered off and the solvent was evaporated. A white powder corresponding to FmocThr(*t*Bu)OBn was obtained (2.28 mmol, 91%). The compound was then solubilized in 15 mL of TFA/CH_2Cl_2 solution1:1 for 90 minutes at room temperature and was evaporated under reduced pressure at 30°C. Finally after recrystallization in EtOAc/hexane compound **AA2** (2.16 mmol) was obtained in 95% yield. The spectroscopy data are in agreement with the literature.[2]

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.76 (d, J = 7.4 Hz, 2H), 7.60 (d, J = 7.1 Hz, 2H), 7.34 (m, 9H), 5.63 (d, J = 9.1 Hz, 1H), 5.21 (t, J = 5.6 Hz, 2H), 4.45-4.37 (m, NH+3H), 4.23 (m, 1H), 1.93 (br, OH), 1.24 (d, J = 5.5 Hz, 3H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 171.2, 156.9, 144.0, 143.8, 141.5, 141.4, 135.3, 129.1, 129.0, 128.9, 128.8, 128.8, 128.7, 128.7, 128.4, 128.4, 127.9, 127.3, 125.3, 120.2, 68.2, 67.6, 67.4, 59.4, 47.3, 20.1. LRMS (ES, NH_4^+) m/z: calcd for $C_{26}H_{25}O_5N$ ($M+Na^+$), 454.2; found, 454.3.

N^{α}-[(Fluoren-9-ylmethoxy)carbonyl]-L-serine pentafluorophenyl ester (AA3)[1]



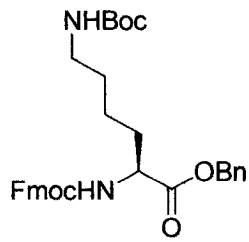
AA3

This compound was synthesized according to a slightly modified procedure of Istvan *et al.*.[1] To a solution of FmocSerOH (2 mmol) and pentafluorophenol (654 mg, 2 mmol) in dry DMF (10 mL), DCI (2 mmol) was added at 0°C. The reaction was stirred at 0°C for 1 hour and 2 hours at 25°C. Dicyclohexylurea was filtered off and the solvent was evaporated. The residue was triturated with hexane, filtered and recrystallized in EtOAc/hexane to afford a white powder (690 mg, 1.64 mmol, 82%).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.73 (d, J = 7.4 Hz, 2H), 7.61 (d, J = 5.4 Hz, 2H), 7.34 (t, J = 7.0, 2H), 7.26 (t, J = 7.1, 2H), 5.93 (d, J = 7.2 Hz, 1H), 4.51–4.36 (m, 2H), 4.22–4.18 (m, 3H), 4.00 (dd, J_1 = 11.3 Hz, J_2 = 2.7 Hz, 1H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 167.8, 157.4, 144.0, 143.9, 143.8, 139.9, 139.8 (m), 136.7 (m), 128.2, 127.9, 127.5, 125.4, 125.1, 120.4,

68.2, 67.9, 59.7, 47.4, 20.2. LRMS (ES, NH_4^+) m/z: calcd for $C_{24}H_{16}O_5NF_5$ ($M+Na^+$), 516.1; found, 516.1.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O-tert*-butyloxycarbonyl-L-lysine-benzyl ester (AA4)

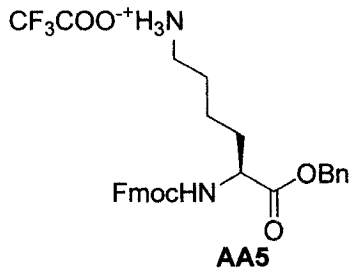


AA4

This compound was synthesized according to a slightly modified procedure of Istvan *et al.*.[1] To a solution of Fmoc-Lys(Boc)-OH (2.5 g, 5.33 mmol) in 45 mL CH_2Cl_2 was added CDI (0.95 g, 5.86 mmol). The mixture was stirred at room temperature for 40 minutes. Benzyl alcohol (0.6 mL, 5.33 mmol) was then added and allowed to stir at 25°C overnight. The product was extracted with CH_2Cl_2 and washed successively with a saturated ammonium chloride solution (100 mL), water (100 mL) and brine (100 mL). The organic phase was dried and concentrated. Flash column chromatography (50:1, CH_2Cl_2/MeOH), afforded the title compound **AA4** (2.67 g, 4.80 mmol, 90 %).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.76 (d, J = 7.5 Hz, 2H), 7.6 (d, J = 7.5 Hz, 2H), 7.42-7.28 (m, 9H), 5.41 (d, J = 7.8 Hz, 1H), 5.23-5.13 (m, 2H), 4.45-4.33 (m, 4H), 4.23-4.19 (m, 1H), 3.06 (bs, 2H), 1.92-1.81 (m, 1H), 1.76-1.64 (m, 1H), 1.43 (s, 10H), 1.38-1.22 (m, 3H); ^{13}C NMR (300 MHz, $CDCl_3$) δ: 172.2, 156.0, 143.8, 143.6, 141.2, 135.1, 128.6, 128.3, 127.6, 127.0, 125.0, 119.9, 79.1, 66.9, 53.7, 47.1, 40.0, 32.1, 29.5, 28.3, 22.2; LRMS (ES, H^+) m/z: calcd for $C_{33}H_{38}N_2O_6$ ($M+H^+$) 559.3; found, 559.1.

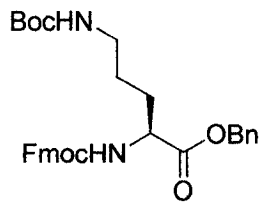
N^{α}-(Fluoren-9-ylmethoxycarbonyl)-L-lysine-benzyl ester trifluoroacetic acid salt (AA5)



AA5

Fmoc-Lys(Boc)-OBn (1.84 g, 3.30 mmol) was dissolved in a solution of CH_2Cl_2 and TFA (1:1) and was stirred for 1 hour at 25°C. Solvents were removed under reduce pressure to yield compound **AA5** (1.84 g, 3.20 mmol, 98%). This compound was directly converted to afford compound **89**. LRMS (ESI) m/z: calcd for $C_{27}H_{28}N_2O_4$ (M-TFA+H^+) 445.5; found, 445.4.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O-tert*-butyloxycarbonyl-L-ornithine-benzyl ester (AA6)



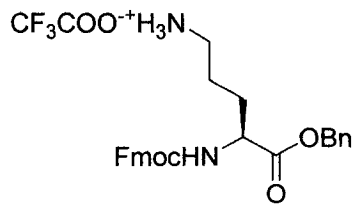
AA6

This compound was synthesized according to a slightly modified procedure of Istvan *et al.*.[1] The starting α-*N*-Fmoc-δ-*N*-Boc ornithine (1 g, 2.20 mmol) was dissolved in freshly distilled dichloromethane (30 mL), cooled to 0°C and stirred under argon for 30 min. 1,1'-Carbonyl diimidazole (392 mg, 2.42 mmol) was added and the solution was stirred for 1 hr while slowly warming to room temperature. Benzyl alcohol (0.250 mL, 2.42 mmol) was added and the reaction mixture was stirred for 24 hr. Although TLC showed that starting material was still present, the reaction was stopped and the solution was concentrated *in vacuo*. The desired compound was first purified by silica gel chromatography eluting with a

gradient of hexane/EtOAc (9:1→1:2) and crystallized from Et_2O to give a white solid (977 mg, 1.78 mmol, 81%).

^1H NMR (300MHz, $CDCl_3$) δ: 7.77 (d, J = 7.5 Hz, 2H), 7.60 (d, J = 7.4 Hz, 2H), 7.44–7.29 (m, 9H), 5.47–5.42 (m, 1H), 5.21 (d, J = 12.2 Hz, 1H), 5.16 (d, J = 12.2 Hz, 1H), 4.50–4.39 (m, 4H), 4.21 (t, J = 6.9 Hz, 1H), 3.10 (bs, 2H), 1.93–1.85 (m, 1H), 1.73–1.63 (m, 1H), 1.55–1.40 (m, 2H), 1.44 (s, 9H). LRMS (ESI) m/z: calcd for $C_{31}H_{34}N_2O_6$ (M^++H) 531.2; found 531.4.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-L-lysine-benzyl ester trifluoroacetic acid salt (AA7)



AA7

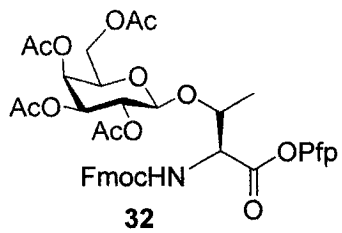
α-*N*-Fmoc-δ-*N*-Boc ornithine benzyl ester (410 mg, 0.753 mmol) was dissolved in an anhydrous 1:1 mixture of trifluoroacetic acid:CH_2Cl_2 (10 mL). The resulting solution was stirred at room temperature for 4 hr, after which it was concentrated under reduced pressure to yield a beige solid. The compound was used in the next step without further purification. LRMS (ESI) m/z: calcd for $C_{26}H_{26}N_2O_4$ (M-TFA+H^+) 431.2; found, 431.1

General procedure for coupling with glycoside bromide as glycoside donor

To a solution of glycoside bromide (2 g, 4.87 mmol, 5 equivalents) and protected amino acid (1 equivalent) in dry dichloromethane (30 mL) was added K_2CO_3 (4 equivalents) and AgOTf (6 equivalents), under an argon atmosphere at -42°C. The solution was degassed and sonicated for 10 minutes with 4Å molecular sieves at -42°C and let to stir for 16 hour at 0°C. The mixture was filtered over Celite, washed successively with $NaHCO_3$ (15 mL) and H_2O (15 mL), dried ($MgSO_4$) and concentrated. The residue was purified by silica gel column chromatography.

Some of the following glycoamino acids have been described in previous publications, however the experimental procedures were modified and the yields were improved.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-L-threonine-pentafluorophenyl ester (32)[3]

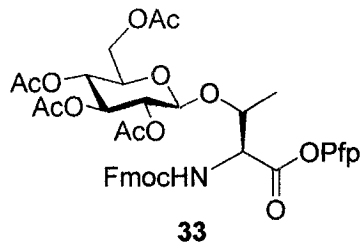


32

The α/β (1/9) mixture obtained was chromatographed on silica gel with (2:1) hexane/EtOAc as eluent and gave the desired compound (73%). Rf 0.25 (hexane-EtOAc 1:1). The spectroscopy data are in agreement with the literature.[3]

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.78 (d, J = 7.4 Hz, 2H), 7.62 (m, 2H), 7.35-7.21 (m, 4H), 6.23 (d, J = 9.3 Hz, 1H), 5.69-5.53 (m, 1H), 5.30 (s, 2H), 5.13 (t, J = 8.8 Hz, 1H), 4.90-5.08 (m, 2H), 4.49-4.25 (m, 3H), 4.23-4.15 (m, 2H), 4.12-3.95 (m, 1H), 3.84 (t, J = 5.7 Hz, 1H), 2.05 (s, 3H), 2.04 (s, 6H), 2.00 (s, 3H), 1.37 (d, J = 6.1 Hz, 3H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 171.0, 170.6, 170.1, 169.7, 156.8, 144.1, 143.9, 141.6, 128.1, 127.5, 125.5, 120.4, 99.6, 74.2, 72.7, 72.2, 71.6, 68.5, 67.8, 62.0, 58.8, 47.4, 21.0, 21,0, 20,9, 20.8, 17.1. LRMS (ES, K^+) m/z: calcd for $C_{39}H_{36}NO_{14}F_5$ ($M+K^+$), 876.8; found, 876.0.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-threonine-pentafluorophenyl ester (33)[3]

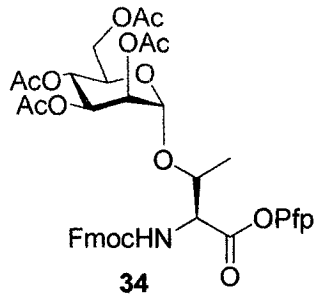


33

The α/β (1/7) mixture obtained was chromatographed on silica gel with (2:1) hexane/EtOAc as eluent and gave the desired compound (71%). Rf 0.23 (hexane-EtOAc 1:1). The spectroscopy data are in agreement with the literature.[3]

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.78 (d, J = 7.5 Hz, 2H), 7.68-7-60 (m, 2H), 7.40 (t, J = 7.4 Hz, 2H), 7.31 (t, J = 7.3 Hz, 2H), 5.75 (d, J = 8.9 Hz, 1H), 5.24-5.20 (m, 2H), 5.09 (t, J = 9.5 Hz, 1H), 5.00 (t, J = 8.8 Hz, 1H), 4.73 (dd, J_1 = 2.0 Hz, J_2 = 8.9 Hz, 1H), 4.53-4.47 (m, 3H), 4.30-4.00 (m, 3H), 3.69 (d, J = 8.7 Hz, 1H), 2.05 (s, 3H), 2.04 (s, 6H), 2.00 (s, 3H), 1.31 (d, J = 6.2 Hz, 3H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 170.7, 170.6, 170.5, 170.2, 169.8, 169.3, 156.9, 156.7, 144.3, 143.9, 141.6, 141.5, 128.1, 128.0, 127.0, 125.4, 120.4, 98.9, 71.3, 69.3, 69.2, 67.9, 66.8, 61.1, 60.7, 47.4, 21.0, 21,0, 20,9, 20.8, 14.5. LRMS (ES, K^+) m/z: calcd for $C_{39}H_{36}O_{14}F_5$ ($M+K^+$), 876.8; found, 876.0.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyl)-L-threonine-pentafluorophenyl ester (34)[3]

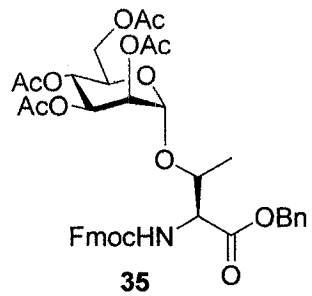


34

The α/β (60/1) mixture obtained was chromatographed on silica gel with (1:1) hexane/EtOAc as eluent and gave the desired compound (68%). Rf 0.21 (hexane-EtOAc 1:1). The spectroscopy data are in agreement with the literature.[3]

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.75 (d, J = 7.5 Hz, 2H), 7.63 (d, J = 7.5 Hz, 2H), 7.41-7.27 (m, 4H), 5.68 (d, J = 9.7 Hz, 1H), 5.32-5.18 (m, 2H), 5.13 (bs, 1H), 4.97 (s, 1H), 4.81 (dd, J_1 = 10.1 Hz, J_2 = 1.8 Hz, 1H), 4.60-4.40 (m, 3H), 4.26 (dd, J_1 = 12.3 Hz, J_2 = 6.1 Hz, 2H), 4.16-4.02 (m, 2H), 2.12 (s, 3H), 2.07 (s, 3H), 2.05 (s, 3H), 1.97 (s, 3H), 1.41 (d, J = 6.3 Hz,.1H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 170.9, 170.2, 170.0, 170.0, 166.9, 156.8, 144.0, 143.9, 143.0, 141.6, 139.9, 136.6, 128.1, 127.5, 125.5, 120.4, 99.5, 77.6, 69.6, 69.0, 68.1, 66.6, 62.9, 58.9, 47.4, 21.4, 21.0, 18.4. LRMS (ES, K^+) m/z: calcd for $C_{39}H_{36}NO_{14}F_5$ ($M+K^+$), 876.7; found, 876.2.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyl)-L-threonine-benzyl ester (35)[4a]

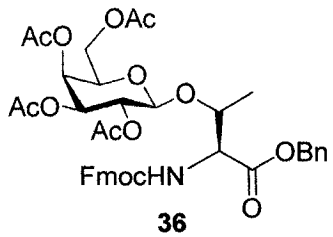


35

The α/β (60/1) mixture obtained was chromatographed on silica gel with (1:2) hexane/EtOAc as eluent and gave the desired compound (75%). Rf 0.20 (hexane-EtOAc 2:5). The spectroscopy data are in agreement with the literature.[4a]

^{1}H NMR (300 MHz,$CDCl_3$) δ: 7.75 (d, J = 7.5 Hz, 2H), 7.63 (d, J = 7.5 Hz, 2H), 7.35-7.29 (m, 9H), 5.68 (d, J = 9.7 Hz, 1H), 5.35-5.14 (m, 5H), 4.95 (s, 1H), 4.81 (dd, J_1 = 10.1 Hz, J_2 = 1.8 Hz, 1H), 4.60-4.40 (m, 3H), 4.28-4.23 (m, 2H), 4.15-4.04 (m, 2H), 2.14 (s, 3H), 2.08 (s, 3H), 2.07 (s, 3H), 1.99 (s, 3H), 1.47 (d, J = 6.1 Hz, 1H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 170.6, 170.4, 170.4, 170.0, 167.2, 156.8, 144.0, 143.9, 143.0, 141.6, 139.9, 136.6, 128.9, 128.5, 128.4, 128.1, 127.5, 125.5, 120.4, 99.5, 77.6, 69.6, 69.0, 68.1, 66.6, 62.9, 58.9, 47.4, 21.4, 21.0, 18.4. LRMS (ES, NH_4^+) m/z: calcd for $C_{39}H_{41}NO_{14}$ ($M+H^+$), 748.3; found, 748.4. HRMS m/z calcd. for $C_{39}H_{40}NO_{14}$ (M^+) 747.2527; found, 747.2520.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl)-L-threonine-benzyl ester (36)[4a]

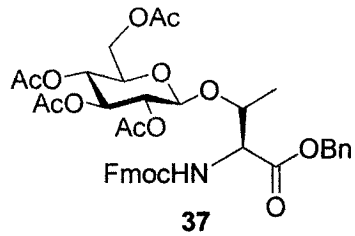


36

The α/β (1/9) mixture obtained was chromatographed on silica gel with (1:2) hexane/EtOAc as eluent and gave the desired compound (75%). Rf 0.21 (hexane-EtOAc 2:5). The spectroscopy data are in agreement with the literature.[4a]

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.73 (d, J = 7.5 Hz, 2H), 7.61 (t, J = 6.4 Hz, 2H), 7.37-7.29 (m, 9H), 5.63 (d, J = 9.4 Hz, 1H), 5.31 (d, J = 2.8 Hz, 1H), 5.19 (d, J = 12.3 Hz, 1H), 5.15 (t, J = 12.4 Hz, 1H), 4.91 (dd, J = 9.5 Hz, J = 8.3 Hz, 1H), 5.11 (dd, J_1 =10.4 Hz, J_2 = 7.9 Hz, 1H), 4.92 (dd, J_1 =10.7 Hz, J_2 = 3.3 Hz, 1H), 4.45-4.31 (m, 5H), 4.22 (t, J = 7.3, 1H), 4.10-3.99 (m, 2H), 3.67 (t, J = 6.8 Hz, 1H), 2.11 (s, 3H), 2.01 (s, 3H),. 2.00 (s, 3H), 1.97 (s, 3H), 1.20 (d, J = 5.6 Hz, 1H). ^{13}C NMR (93 MHz, $CDCl_3$) δ: 170.6, 170.5, 170.4, 170.1, 169.6, 157.0, 144.2, 144.0, 141.6, 141.5, 135.6, 128.9, 128.7, 128.5, 127.9, 127.4, 125.5, 125.5, 120.2, 99.8, 75.6, 70.9, 70.8, 69.1, 67.6, 67.6, 67.0, 61.1, 60.7, 58.8, 47.4, 21.3, 21.0, 20.9, 20.8, 17.7. LRMS (ES, NH_4^+) m/z: calcd for $C_{39}H_{41}NO_{14}$ ($M+H^+$), 748.3; found, 748.4. HRMS m/z calcd. for $C_{39}H_{40}NO_{14}$ (M^+) 747.2527; found, 747.2528.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-threonine-,benzyl ester (37)[4a]



37

The α/β (1/7) mixture obtained was chromatographed on silica gel with (1:2) hexane/EtOAc as eluent and gave the desired compound (82%). Rf 0.20 (hexane-EtOAc 2:5). The spectroscopy data are in agreement with the literature.[4a]

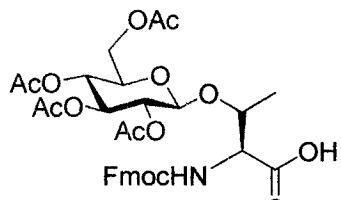
^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.76 (d, J = 7.6 Hz, 2H), 7.62 (dd, J_1 = 7.5 Hz, J_2 = 3.2 Hz, 2H), 7.41-7.28 (m, 9H), 5.63 (d, J = 9.3 Hz, 1H), 5.22 (d, J = 12.5 Hz, 1H), 5.15 (d, J = 12.5 Hz, 1H), 5.03 (t, J = 9.3 Hz, 1H), 4.91 (dd, J_1 = 9.5 Hz, J_2 = 8.3 Hz, 1H), 5.14 (t, J = 9.3 Hz, 1H), 4.47-4.33 (m, 6H), 4.24 (t, J = 7.3, 1H), 4.19 (dd, J_1 = 12.5 Hz, J_2 = 4.3 Hz, 1H), 4.02 (dd, J_1 = 13.8 Hz, J_2 = 2.0 Hz, 1H), 3.47-3.43 (m, 1H), 2.04 (s, 3H), 2.02 (s, 3H), 2.02 (s, 3H), 2.00 (s, 3H), 1.22 (d, J = 6.1 Hz, 1H). ^{13}C NMR (93 MHz, $CDCl_3$) δ: 170.5, 170.1, 169.7, 169.2, 169.1, 156.6, 143.8, 143.5, 141.1, 141.1, 135.2, 128.5, 128.4, 128.1, 127.6, 126.9, 125.0, 119.8, 98.7, 74.9, 72.3, 71.5, 71.1, 67.9, 67.2, 61.4, 58.3, 53.3, 47.0, 20.5, 20.5, 20.5, 20.4, 17.0. LRMS (ES, NH_4^+) m/z: calcd for $C_{39}H_{41}NO_{14}$ ($M+H^+$), 748.3; found, 748.3. HRMS m/z calcd. for $C_{39}H_{40}NO_{14}$ (M^+), 747.2527; found, 747.2528.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-galacopyranosyl)-L-threonine (38)[4b]

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-β-D-glucopyranosyl)-L-threonine (39)[4b]

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-acetyl-α-D-mannopyranosyl)-L-threonine (40)[4b]

Compounds **35**, **36** and **37** were dissolved in an EtOH or EtOAc (depending of the solubility) solution containing 10% $^{w}/_{w}$ Pd/C and the reactions, stirred for 15 hours at 25°C. After this time the mixtures were filtered and chromatographed on silica gel (CH_2Cl_2/MeOH (90/10 to 95/5)). Further purification of the products was done by C-18 reverse phase HPLC[5] (isocrat, 45% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compounds **38**, **39**, **40** were obtained with a retention time of 7.70 min, 7.25 min and 6.95 min and 93%, 96% and 93% yield respectively. Each compound was fully characterized. Due to the hydrogen bonding generated by the acid function all the signal were broad, however the integration and the chemical displacements corresponded to previously reported data.[4b] ^{13}C Spectra and HRMS were also in agreement with these data..



39

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.73 (d, J = 7.5 Hz, 2H), 7.60 (t, J = 7.5 Hz, 1H), 7.37 (t, J = 7.4 Hz, 2H), 7.30 (t, J = 7.0 Hz, 2H), 5.63 (d, J = 8.9 Hz, 1H), 5.22 (m, 2H), 5.09 (t, J = 9.5 Hz, 1H), 5.00 (t, J = 8.8 Hz, 1H), 4.73 (dd, J_1 = 2.0 Hz, J_2 = 8.9 Hz, 1H), 4.50 (m, 3H), 4.30-4.00 (m, 3H), 3.69 (d, J = 8.7 Hz, 1H), 2.05 (s, 3H), 2.04 (s, 6H), 2.00 (s, 3H), 1.31 (d, J = 6.2 Hz, 3H). ^{13}C NMR (93 MHz, $CDCl_3$) δ: 171.0, 170.6, 170.1, 169.7, 156.8, 144.1, 143.9, 141.6, 128.1, 127.5, 125.5, 120.4, 99.6, 74.2, 72.7, 72.2, 71.6, 68.5, 67.8, 62.0, 58.8, 47.4,

21.0, 21.0, 20.9, 20.8, 17.1. LRMS (ES, H^+) m/z: calcd for $C_{33}H_{37}NO_{14}$ (M^++Na), 694.2; found, 694.1.

38 : HRMS m/z calcd. for $C_{33}H_{37}NO_{14}$ (M^+), 671.2214; found, 671.2214.

Rf 0.25 (CH_2Cl_2-MeOH 9:1).

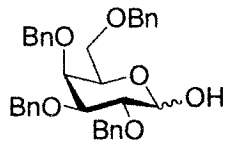
39 : HRMS m/z calcd. for $C_{33}H_{37}NO_{14}$ (M^+), 671.2214; found, 671.2210.

Rf 0.23 (CH_2Cl_2-MeOH 9:1).

40 : HRMS m/z calcd. for $C_{33}H_{37}NO_{14}$ (M^+), 671.2214; found, 671.2222.

Rf 0.22 (CH_2Cl_2-MeOH 9:1).

2,3,4,6-Tetra-*O*-benzyl-D-galactopyranose (47)[6]

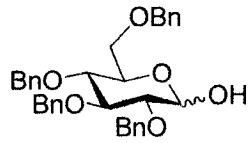


47

This compound was synthesized according to the procedure of Motawia *et al.*.[6] The α/β (5/2) mixture obtained was chromatographed on silica gel with (9:1) hexane/EtOAc as the initial eluent to remove all impurities and (7:3) hexane–EtOAc to elute **47** to give the desired compound (98%).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.84-7.80 (m, 2H), 7.55-7.51 (m, 3H), 5.54-5.50 (m, 17H), 5.37 (m, 3H), 5.20 (d, J = 11.5 Hz, 1H), 5.03 (d, J = 10.3 Hz, 1H), 4.96 (d, J = 10.4 Hz, 1H), 4.96-4.79 (m, 4H), 4.68 (d, J = 11.8 Hz, 1H), 4.61 (d, J = 11.8 Hz, 1H), 4.25-4.17 (m, 2H), 3.89-3.81 (m, 4H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 138.7, 138.3, 138.2, 137.8, 134.1, 131.3, 128.7, 128.3, 128.2, 128.2, 128.1, 127.8, 127.7, 127.7, 127.6, 127.6, 127.5, 127.4, 126.9, 87.5, 84.0, 77.1, 77.1, 75.5, 74.4, 73.4, 73.4, 72.5, 68.7. LRMS (ES, NH_4^+) m/z: calcd for $C_{34}H_{36}O_6$ (M+NH_4^+), 558.3; found, 558.3.

2,3,4,6-Tetra-*O*-benzyl-D-glucopyranose (48)[6]



48

This compound was synthesized according to the procedure of Motawia *et al.*.[6] The α/β (5/2) mixture obtained was chromatographed on silica gel with (9:1) hexane/EtOAc as the initial eluent to remove all impurities and (7:3) hexane–EtOAc to elute **47** and gave the desired compound (74%).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.76 (m, 2H), 7.45 (m, 23H), 5.02 (m, 4H), 4.89 (d, J = 10.4 Hz, 1H), 4.85 (d, J = 10.1 Hz, 1H), 4.72 (m, 3H), 3.87 (m, 4H), 3.69 (m, 2H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 138.4, 138.3, 138.1, 138.1, 133.8, 131.9, 128.9, 128.5, 128.5, 128.4, 128.2, 128.2, 128.0, 127.9, 127.9, 127.8, 127.7, 127.7, 127.6, 127.5, 87.4, 86.8, 80.8, 79.1, 77.8, 75.8, 75.4, 75.0, 73.4, 69.0. LRMS (ES, NH_4^+) m/z: calcd for $C_{34}H_{36}O_6$ ($M+NH_4^+$), 558.3; found, 558.3.

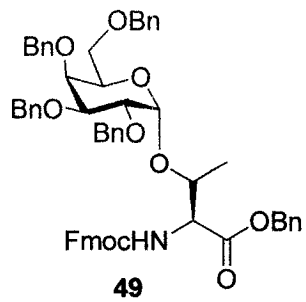
General procedure for coupling with a trichloroimidate as glycoside donor

The protected glycopyranoside with anomeric free hydroxyl (1 equivalent) was dissolved in dry dichloromethane under an argon atmosphere at 25°C. The solution was degassed and sonicated for 15 minutes with 3Å molecular sieves. The mixture was treated with Cl_3CCN (40 equivalents) and K_2CO_3 (25 equivalents) and stirred at 25°C for 10 hours or until the reaction was completed. The mixture was filtered over Celite and silica gel and was concentrated to give an amorphous solid. The residue was dried overnight over phosphorus pentoxide under reduce pressure and was used for the coupling without any further purification.

To a solution of trichloroimidate (1.5 equivalents), TBDMSOTf (0.8 equivalent) and 4Å molecular sieves in dry CH_2Cl_2/ Et_2O (1:1) under an argon atmosphere and at 0°C was added the protected amino acid (1 equivalent) dropwise. The solution was stirred for 12 hours at 25°C and $NaHCO_3$ was added. The mixture was stirred for 15 minutes, filtered

over Celite and concentrated. The residue was eluted from a column of silica gel. For the synthesis of compound **60,** $BF_3.Et_2O$ (0.8 equivalent) was used instead of TBDMSOTf.

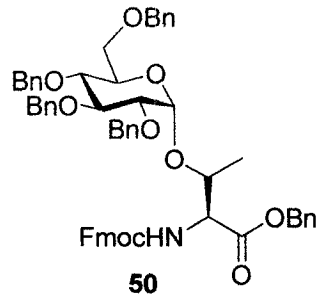
N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-benzyl-α-D-galactopyranosyl)-L-threonine-benzyl ester (49)[7]



49

The α/β (6/1) mixture obtained was chromatographed on silica gel with (97:3) $CHCl_3/Et_2O$ as eluent and gave the desired compound (62%). Rf 0.25 ($CHCl_3$-Et_2O 97:3). ^1H NMR (500 MHz, $CDCl_3$) δ: 7.76 (dd, J_1 = 7.3 Hz, J_2 = 4.2 Hz, 2H), 7.65 (t, J = 7.5 Hz, 2H), 7.43-7.23 (m, 29H), 6.26 (d, J = 8.0 Hz, 1H), 5.13 (s, 1H), 4.98 (d, J = 11.4 Hz, 1H), 4.93 (d, J = 3.5 Hz, 1H), 4.85 (d, J = 12.6 Hz, 1H), 4.81 (d, J = 11.8 Hz, 1H), 4.76-4.70 (m, 2H), 4.64-4.59 (m, 1H), 4.54-4.38 (m, 6H), 4.26 (t, J = 7.2 Hz, 1H), 4.06-4.02 (m, 2H), 4.00-3.96 (m, 1H), 3.64-3.55 (m, 2H), 1.36 (d, J = 6.3 Hz, 1H). ^{13}C NMR (93 MHz, $CDCl_3$) δ: 170.5, 156.7, 143.9, 143.7, 141.2, 141.1, 138.5, 138.1, 137.8, 135.2, 128.4, 128.4, 128.3, 128.2, 128.2, 128.1, 128.1, 128.0, 127.8, 127.7, 127.6, 127.5, 127.5, 127.4, 127.4, 127.3, 127.2, 127.0, 126.9, 125.1, 125.0, 119.8, 98.4, 78.9, 75.7, 74.9, 74.7, 73.4, 72.7, 69.9, 68.7, 67.1, 67.0, 65.7, 59.1, 47.1, 19.1, 15.2. LRMS (ES, NH_4^+) m/z: calcd for $C_{59}H_{57}NO_{10}$ ($M+H^+$), 940.4; found, 940.4. HRMS m/z calcd. for $C_{59}H_{56}NO_{10}$ (M^+), 939.3982; found, 939.3982.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-benzyl-α-D-glucopyranosyl)-L-threonine-benzyl ester (50)[7]

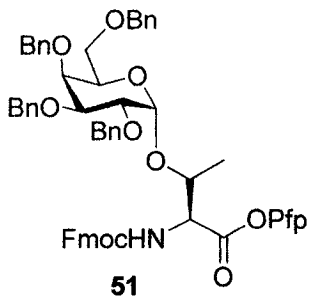


50

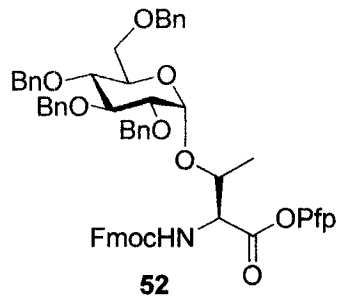
The α/β (10/1) mixture obtained was chromatographed on silica gel with (97:3) $CHCl_3/Et_2O$ as eluent and gave the desired compound (42%). Rf 0.27 ($CHCl_3$-Et_2O 97:3). ^1H NMR (500 MHz, $CDCl_3$) δ: 7.76 (dd, J_1 = 7.5 Hz, J_2 = 3.6 Hz, 2H), 7.65 (dd, J_1 = 6.9 Hz, J_2 = 5.2 Hz, 2H), 7.43-7.26 (m, 29H), 7.21 (dd, J_1 = 7.9 Hz, J_2 = 2.0 Hz, 2H), 6.12 (d, J = 8.2 Hz, 1H), 5.14 (s, 1H), 5.02 (d, J = 11.1 Hz, 1H), 4.95-4.87 (m, 3H), 4.73-4.60 (m, 3H), 4.53 (dd, J_1 = 11.3. Hz, J_2 = 6.6 Hz, 2H), 4.49-4.43 (m, 3H), 4.27 (t, J = 7.0 Hz, 1H), 4.03 (t, J = 9.3 Hz, 1H), 3.96-3.90 (m, 1H), 3.82 (dd, J_1 = 10.6 Hz, J_2 = 3.4 Hz, 1H), 3.74-3.68 (m, 2H), 3.55 (dd, J_1 = 9.6 Hz, J_2 = 3.5 Hz, 1H), 1.38 (d, J = 6.1 Hz, 1H). ^{13}C NMR (93 MHz, $CDCl_3$) δ: 170.6, 156.9, 144.1, 143.9, 141.4, 141.4, 138.8, 138.2, 138.0, 138.0, 135.4, 128.7, 128.5, 128.5, 128.3, 128.1, 128.0, 128.0, 127.9, 127.8, 127.2, 125.3, 125.2, 120.1, 98.3, 81.9, 79.6, 77.7, 75.9, 75.7, 75.4, 73.7, 73.3, 71.1, 68.5, 67.4, 67.3, 59.3, 47.3, 19.3, 19.3. LRMS (ES, NH_4^+) m/z: calcd for $C_{59}H_{57}NO_{10}$ ($M+H^+$), 940.4; found, 940.4. HRMS m/z calcd. for $C_{59}H_{56}NO_{10}$ (M^+) 939.3982, found 939.3982.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-benzyl-α-D-galactopyranosyl)-L-threonine-pentafluorophenyl ester (51)

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-benzyl-α-D-glucopyranosyl)-L-threonine-pentafluorophenyl ester (52)



51



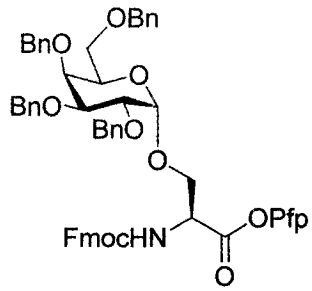
52

These compounds were obtained at Binghamton University. Having no more access to the NMR software for such files and this methology being redesigned toward the benzyl esters, the spectra are not be described herein. However, the mass spectrum corresponded to the desired mass. The ^{1}H and ^{13}C NMR spectra for compound **52** were salvaged from a paper copy and are presented in the Appendix.

51 : The α/β (6/1) mixture obtained was chromatographed on silica gel with (98:2) $CHCl_3/Et_2O$ as eluent and gave the desired compound (82%). Rf 0.28 ($CHCl_3$-Et_2O 98:2). LRMS (ES, NH_4^+) m/z: calcd for $C_{59}H_{52}F_5NO_{10}$ ($M+Na^+$), 1052.4; found, 1052.8.

52 : The α/β (10/1) mixture obtained was chromatographed on silica gel with (98:2) $CHCl_3/Et_2O$ as eluent and gave the desired compound (49%). Rf 0.28 ($CHCl_3$-Et_2O 98:2). LRMS (ES, NH_4^+) m/z: calcd for $C_{59}H_{52}F_5NO_{10}$ ($M+H^+$), 1030.4; found, 1030.4.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2,3,4,6-tetra-*O*-benzyl-α-D-galactopyranosyl)-L-threonine-pentafluorophenyl ester (53)

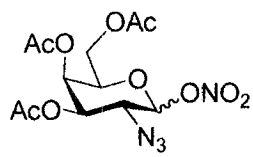


53

The α/β (5/1) mixture obtained was chromatographed on silica gel with (98:2) $CHCl_3/Et_2O$ as eluent and gave the desired compound (49%). Rf 0.26 ($CHCl_3$-Et_2O 98:2).

^1H NMR (500 MHz, $CDCl_3$) δ: 7.70 (dd, J_1 = 7.3 Hz, J_2 = 3.9 Hz, 2H), 7.53 (dd, J_1 = 12.1 Hz, J_2 = 7.5 Hz, 2H), 7.40-7.15 (m, 29H), 6.69 (d, J = 9.2 Hz, 1H), 4.92 (d, J = 11.6 Hz, 1H), 4.86-4.77 (m, 3H), 4.76 (d, J = 3.8 Hz, 1H), 4.71 (d, J = 11.7 Hz, 1H), 4.62 (d, J = 12.0 Hz, 1H), 4.53 (d, J = 11.6 Hz, 1H), 4.51 (dd, J_1 = 11.9 Hz, J_2 = 4.2 Hz, 1H), 4.41 (d, J = 12.3 Hz, 1H), 4.35 (dd, J_1 = 10.4 Hz, J_2 = 7.2 Hz, 1H), 4.31-4.25 (m, 2H), 4.11 (t, J = 7.2 Hz, 1H), 4.05 (dd, J_1 = 9.6 Hz, J_2 = 4.0 Hz, 1H), 3.91 (t, J = 6.1 Hz, 1H), 3.88-3.81 (m, 3H), 3.53 (dd, J_1 = 9.3 Hz, J_2 = 7.0 Hz, 1H), 3.33 (dd, J_1 = 9.3 Hz, J_2 = 5.2 Hz, 1H). HRMS m/z calcd. for $C_{58}H_{50}NO_{10}F_5$ (M^+), 1015.3355; found, 1015.3354.

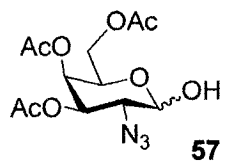
3,4,6-Tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactopyranosyl nitrate (55b)[8]



This compound was synthesized according to a slightly modified procedure of Lemieux *et al.*.[8] To a stirred solution of **54** (2.3 g, 10.9 mmol) in distilled acetonitrile (46 mL) under argon at -15°C was added sodium azide (0.82 g, 12.6 mmol) and cerium ammonium nitrate (13.8 g, 30.6 mmol). The reaction was stirred for 1 hr at -15°C and 1 hr at 25°C. The reaction mixture was then diluted in cold Et_2O (40 mL) and H_2O (20 mL) and

washed with water (3*40 mL). The organic layer was then dried over magnesium sulfate, filtered and concentrated under reduced pressure. The residue obtained was chromatographed on silica gel chromatography with (3:2) hexane/EtOAc as eluent to give the desired compound **55b** (2.70g, 7.2mmol, 85%). The characterization data are in accordance with the published ones.

3,4,6-Tri-*O*-acetyl-2-azido-2-deoxy-α-D-galactopyranose (57) (α/β = 1.3/1)[9]



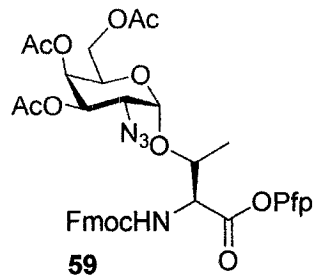
57

This compound was synthesized according to the procedure of Llen *et al.*[9]

α analogue : IR (film) 2121 (N_3) cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ: 5.42-5.38 (m, 2H), 4.82 (dd, J_1 = 3.4 Hz, J_2 = 9.9 Hz, 1H), 4.48 (t, J = 6.6, 1H), 4.15-4.09 (m, 2H), 3.76-3.70 (m, J = 3.4, 2H), 2.16 (s, 3H), 2.06 (s, 6H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 170.9, 170.4, 170.3, 92.3, 70.8, 67.8, 66.4, 58.1, 20.7, 20.7, 20.7.

β analogue : IR (film) 2121 (N_3) cm^{-1}; ^1H NMR (300 MHz, $CDCl_3$) δ: 5.42-5.38 (m, 2H), 4.72 (d, J_1 = 8.0 Hz, 1H), 4.14-4.10 (m, 2H), 3.94 (t, J = 6.6 Hz, 1H), 3.70-3.65 (m, J = 8.0, 2H), 2.17 (s, 3H), 2.07 (s, 6H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 170.9, 170.4, 170.2, 96.4, 71.3, 68.4, 66.6, 62.0, 60.7, 61.9, 61.7, 58.1, 20.7, 20.7, 20.6. LRMS (CI, Iso) m/z: calcd for $C_{12}H_{17}N_3O_8$ $(M)^+$, 331; found, 331, $C_{12}H_{17}N_3O_8$ (M-OH), 314; found, 314.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(3,4,6-tri-*O*-acetyl-2-azido-2-desoxy-α-D-galactopyranosyl)-L-threonine-pentafluorophenyl ester (59)[3]

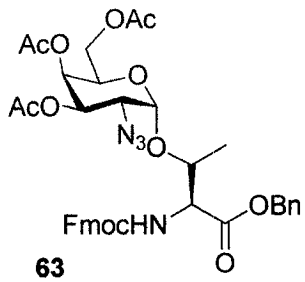


59

The α/β (3/1) mixture obtained was chromatographed on silica gel with (3:1) hexane/EtOAc as eluent and gave the desired compound (82%). Rf 0.25 (hexane-EtOAc 3:1). The spectroscopy data are in agreement with the literature.[3]

IR (film) 2120 (N_3) cm^{-1}; ^1H NMR (360 MHz, $CDCl_3$) δ: 7.75 (d, J = 7.5 Hz, 2H), 7.61 (d, J = 7.9 Hz, 2H), 7.41-7.24 (m, 4H), 5.87 (d, J = 9.4 Hz, 1H), 5.46 (d, J = 2.2 Hz, 2H), 5.29 (dd, J_1 = 11.2 Hz, J_2 = 3.3 Hz, 1H), 5.16 (d, J = 3.6 Hz, 1H), 4.76 (dd, J_1 = 9.4 Hz, J_2 = 1.8 Hz, 1H), 4.60-4.55 (m, 1H), 4.50-4.39 (m, 2H), 4.31-4.25 (m, 2H), 4.11-4.09 (d, J = 10.4 Hz, 2H), 3.75 (dd, J_1 = 10.8 Hz, J_2 = 4.0 Hz, 1H), 2.15 (s, 3H), 2.07 (s, 3H), 2.04 (s, 3H), 1.44 (d, J = 6.5 Hz, 3H). LRMS (ES, H^+) m/z: calcd for $C_{37}H_{33}N_4O_{12}F_5$ ($M+Na^+$), 843.0; found, 843.2.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(3,4,6-tri-*O*-acetyl-2-azido-2-desoxy-α-D-galactopyranosyl)-L-threonine-benzyl ester (63)[10]

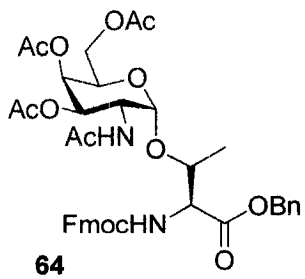


63

The α/β (7/1) mixture obtained was chromatographed on silica gel with (4:1) hexane/EtOAc as eluent and gave the desired compound (62%). Rf 0.22 (hexane-EtOAc 4:1). The spectroscopy data are in agreement with the literature.[10]

IR (film) 2121 (N_3) cm^{-1}, ^1H NMR (300 MHz, $CDCl_3$) δ: 7.76 (d, J = 7.5 Hz, 2H), 7.62 (d, J = 7.2 Hz, 2H), 7.40-7.26 (m, 9H), 5.71 (d, J = 9.6 Hz, 1H), 5.46-5.39 (m, 2H), 5.38-5.19 (m, 3H), 4.89 (d, J = 3.6 Hz, 1H), 4.49-4.04 (m, 8H), 2.15 (s, 3H), 2.06 (s, 3H), 2.04 (s, 3H), 1.34 (d, J = 6.3 Hz, 3H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 170.7, 170.4, 170.4, 170.3, 170.2, 157.1, 144.1, 144.0, 141.6, 135.1, 129.0, 129.0, 128.9, 128.0, 127.4, 125.5, 120.2, 99.6, 92.7, 77.1, 68.6, 68.2, 68.1, 67.7, 67.3, 66.9, 62.0, 58.9, 68.0, 57.9, 47.3, 21.0, 20.9, 18.8, 14.4. LRMS (ES, H^+) m/z: calcd for $C_{38}H_{40}N_4O_{12}$ ($M+Na^+$), 767.4; found, 767.1.

N^{α}-(Fluoren-9-ylmethoxycarbonyl)-*O*-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl)-L-threonine-benzyl ester (64)[10]

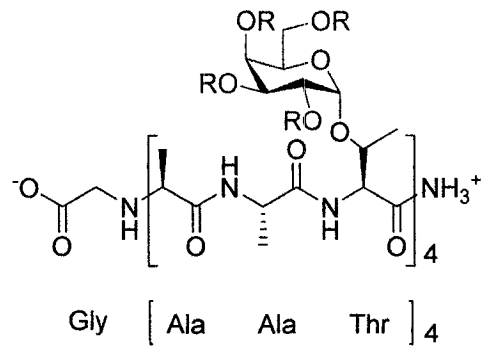


64

The residue was chromatographed on silica gel with (3:1) hexane/EtOAc as eluent and gave the desired compound (70%). Rf 0.23 (hexane-EtOAc 3:1). The spectroscopy data are in agreement with the literature.[10]

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.75 (d, J = 7.3 Hz, 2H), 7.61 (d, J = 7.3 Hz, 2H), 7.39-7.27 (m, 9H), 5.85 (d, J = 9.6 Hz, 1H), 5.72 (d, J = 9.5 Hz, 1H), 5.34 (d, J = 2.1 Hz, 1H), 5.16 (d, J = 11.8 Hz, 1H), 5.07-5.02 (m, 2H), 4.76 (d, J = 3.5 Hz, 1H), 4.52-4.37 (m, 4H), 4.24-4.22 (m, 2H), 4.16 (t, J = 6.0 Hz, 1H), 4.11-3.90 (m, 2H), 2.14 (s, 3H), 2.00 (s, 3H), 1.97 (s, 3H), 1.95 (s, 3H), 1.27 (d, J = 6.3 Hz, 3H). ^{13}C NMR (93 MHz, $CDCl_3$) δ: 171.3, 171.1, 170.7, 170.7, 170.7, 156.8, 144.1, 144.0, 141.7, 134.7, 129.3, 129.3, 128.9, 128.8, 128.2, 127.5, 125.5, 125.4, 120.4, 100.3, 77.6, 68.8, 68.2, 67.9, 67.7, 67.6, 62.5, 58.9, 47.8, 47.5, 23.6, 21,1, 21.1, 21.0, 18.5. LRMS (ES, H^+) m/z: calcd for $C_{40}H_{44}N_2O_{13}$ ($M+Na^+$), 783.0; found, 783.4.

α –Gal (4OH)(TAA)$_4$G (70)



66 : R = Bn α-Gal(4OBn)(TAA)$_4$G
70 : R = H α-Gal(4OH)(TAA)$_4$G

After cleavage from the resin, compound **66** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **66** was obtained with a retention time of 12.53 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{178}H_{209}N_{13}O_{38}$ $(M-H)^+$, 3135.4; found, 3135.4.

The compound was then readily debenzylated to afford compound **70** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **70** was

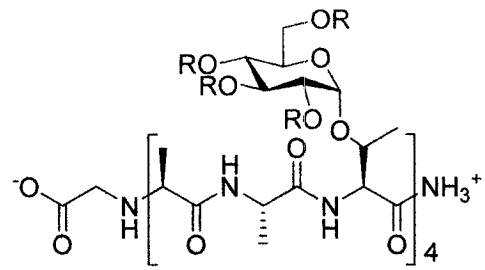
obtained with a retention time of 9.45 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{66}H_{113}N_{13}O_{38}$ $(M-H)^+$, 1695.7; found, 1694.2.

66 : ^1H NMR (300 MHz, $CDCl_3$) δ: 7.74-7.53 (7H), 7.31-7.22 (7H), 5.37 (14H), 1.83-1.70 (bs), 0.90-0.65 (bs),^{13}C NMR (56 MHz, $CDCl_3$) 172.8, 171.8, 171.9, .123.4, 122.4, 118.5, 99.9, 99.9, 96.1, 89.5, 32,3, 30.1, 29.8, 29.7, 29.6, 29.5, 29.4, 23.1, 23.0, 14.5

70 : IR (neat) cm^{-1}: 3330, 1637.

^1H NMR (500 MHz, $CDCl_3$) δ: 5.95, 4.70 (10H), 4.51-4.11 (1H), 4.27-3.90 (1H), 3.72-3.38 (2H), 3.32-3.25 (4H), 2.35-2.22 (1H), 2.07-1.91 (1H), 1.83-1.50 (2H), 1.41-1.23 (bs 20H) 0.90-0.65 (bs, 5H),^{13}C NMR (93 MHz, $CDCl_3$) 172.8, 171.8, 171.8, 100.1, 99.9, 89.5, 32,3, 30.1, 29.8, 29.7, 29.6, 29.5, 29.4, 23.1, 23.1.

α –Glc (4OH)(TAA)$_4$G (71)



67 : R = Bn α-Glc(4OBn)(TAA)$_4$G
71 : R = H α-Glc(4OH)(TAA)$_4$G

After cleavage from the resin, compound **67** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **67** was obtained with a retention time of 11.94 min. (MALDI, MeOH, positive ion) m/z: calcd for $C_{178}H_{209}N_{13}O_{38}$ $(M+Na)^+$, 3159.4; found, 3159.0.

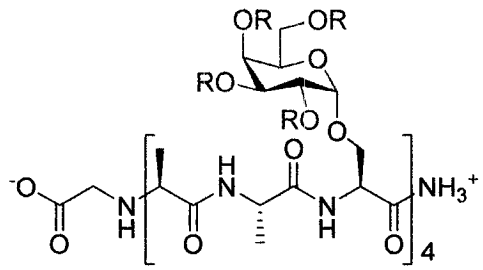
The compound was then readily debenzylated to afford compound **71** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **71** was

obtained with a retention time of 10.05 min. (MALDI, MeOH, positive ion) m/z: calcd for $C_{66}H_{113}N_{13}O_{38}$ $(M+Na)^+$, 1718.7; found, 1718.2.

71 : IR (neat) cm^{-1}: 3335, 1641.

^1H NMR (300 MHz, D_2O) δ: 4.75, 4.68 (20H), 4.38-4.20 (0.7H), 4.06-3.78 (0.8H), 3.72-3.58 (1H), 3.35-2.75 (1H), 2.10-1.61 (0.7H), 1.58-1.12 (3.6H), ^{13}C NMR (93 MHz, $CDCl_3$) 170.2, 170.1, 169.9, 101.3, 101.2, 99.4, 90.5. 90.4, 70.5-66.0 (20C), 60.2, 58.3, 53.3, 46.9, 20.6, 20.6, 20.6, 20.4, 20.4, 17.3

α –Gal (4OH)(SAA)$_4$G (72)



Gly [Ala Ala Ser]$_4$

68 : R = Bn α-Gal(4OBn)(SAA)$_4$G
72 : R = H α-Gal(4OH)(SAA)$_4$G

After cleavage from the resin, compound **68** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **68** was obtained with a retention time of 13.02 min. (MALDI, MeOH, positive ion) m/z: calcd for $C_{174}H_{201}N_{13}O_{38}$ $(M+Na)^+$, 3103.2; found, 3104.0.

The compound was then readily debenzylated to afford compound **72** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **72** was obtained with a retention time of 9.27 min. (MALDI, MeOH, positive ion) m/z: calcd for $C_{62}H_{105}N_{13}O_{38}$ $(M+Na)^+$, 1662.5; found, 1663.2.

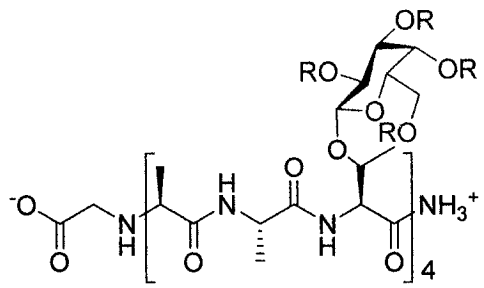
68 : ^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.40-7.25 (15H), 5.00 (1H), 4.62 (1H), 4.60-4.29 (2H), 4.25-3.85 (3H), 3.42-3.38 (6H), 2.91-2.67 (10H), 2.51-2.42 (6H), 1.83-2.05 (6H), 1.41-1.23 (10H) 0.90-0.65 (20H),^{13}C NMR (93 MHz, $CDCl_3$) 175.8, 175.8, 175.6, 175.3, 175.2, 128.4, 128.4, 128.4, 128.3, 128.3, 128.2, 127.8, 126.2, 99.9, 99.8, 96.1, 49.3-46.9 (25C) , 38.6, 31.7, 31.5, 31.4, 29.2-28.2 (15C), 22.2, 17.1, 16.8, 12.8.

72 : IR (neat) cm^{-1}: 3331, 1639.

^{1}H NMR (500 MHz, $CDCl_3$) δ: 5.01 (1H), 4.60 (13H), 4.50-4.21 (1H), 4.25-3.85 (1H), 3.42-3.38 (3H), 2.91-2.67 (1H), 2.51-2.42 (1H), 1.83-2.05 (2H), 1.41-1.23 (5H) 0.90-0.65 (3H).

^{13}C NMR (93 MHz, $CDCl_3$) δ: 172.8, 172.8, 171. 6, 171.3, 171.2, 99.8, 99.7, 96.1, 49.6, 31.7, 31.4, 30.4, 29.7, 29.7, 29.6, 29.5,29.4, 29.2, 29.1, 29.0, 22.5, 20.4, 17.4, 13.9, 11.3.

β –Gal (4OH)(TAA)$_4$G (78)



Gly [Ala Ala Thr]$_4$

74 : R = Ac β-Gal(4OAc)(TAA)$_4$G
78 : R = H β-Gal(4OH)(TAA)$_4$G

After cleavage from the resin, compound **74** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **74** was obtained with a retention time of 12.63 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{98}H_{145}N_{13}O_{54}$ $(M-H)^+$, 2369.2; found, 2367.9.

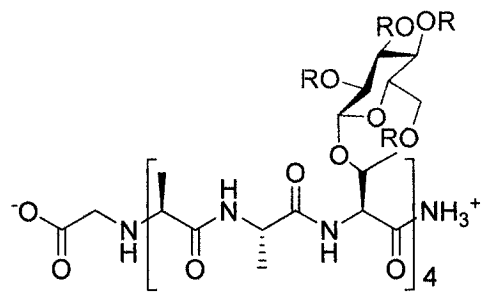
The compound was then readily debenzylated to afford compound **78** which was also purified by C-18 reverse phase then HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **78** was

obtained with a retention time of 9.97 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{66}H_{113}N_{13}O_{38}$ $(M-H)^+$, 1695.7; found, 1694.9.

78 : IR (neat) cm^{-1}: 3328, 1645.

^1H NMR (300 MHz, D_2O) δ: 5.23 (5H), 5.2 (9H), 4.8 (45H), 4.51 (4.5H), 3.05-2.97 (12H), 2.47-2.27 (45.3H), 1.25-1.12 (17.5H), 1.58-1.12 (3.6H), ^{13}C NMR (56 MHz, $CDCl_3$) δ: 178.2, 177.9, 176.6, 175.9, 174.4-165.5 (6C), 122.3, 120.0, 117.6, 115.3, 75.4, 74.9, 72.0, 71.5, 70.0, 63.3, 54.7, 50.9, 49.7, 44.5-37.8 (18C), 60.2, 58.3, 53.3, 46.9, 20.6, 20.6, 20.6, 20.4, 20.4, 17.3.

β –Glc (4OH)(TAA)$_4$G (79)



75 : R = Ac β-Glc(4OAc)(TAA)$_4$G
79 : R = H β-Glc(4OH)(TAA)$_4$G

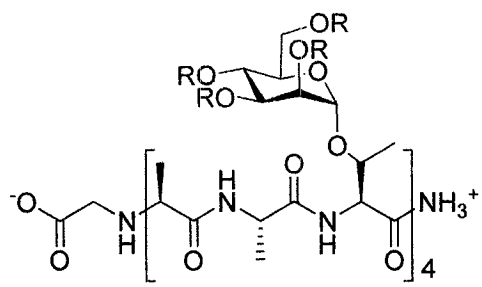
After cleavage from the resin, compound **75** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **75** was obtained with a retention time of 11.75 min. (MALDI, MeOH, positive ion) m/z: calcd for $C_{98}H_{145}N_{13}O_{54}$ $(M+K)^+$, 2406.9; found, 2406.1.

The compound was then readily debenzylated to afford compound **79** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **79** was obtained with a retention time of 9.76 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{66}H_{113}N_{13}O_{38}$ $(M-H)^+$, 1695.7; found, 1694.9.

79 : IR (neat) cm^{-1}: 3331, 1636.

^1H NMR (300 MHz, MeOD) δ: 5.04, 4.70 (20H), 4.53-4.12 (2H), 4.10-3.85 (2H), 3.72-3.41 (6H), 2.92-2.69 (3H), 2.68-2.42 (1H), 1.99-2.05 (5H), 1.41-1.23 (33H) 0.90-0.65 (3H). ^{13}C NMR (93 MHz, $CDCl_3$) 175,7 175,6 175.3, 171.3, 100.0, 99.8, 99.7, 96.3, 49.9, 31.7, 31.6, 30.6, 29.9, 29.9, 29.6, 29.5, 29.4, 29.2, 29.1, 29.0, 22.5, 20.5, 17.5, 13.9, 11.9.

α –Man (4OH)(TAA)$_4$G (80)



Gly [Ala Ala Thr]$_4$

76 : R = Ac α-Man(4OAc)(TAA)$_4$G
80 : R = H α-Man(4OH)(TAA)$_4$G

After cleavage from the resin, compound **76** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **76** was obtained with a retention time of 11.44 min. LRMS (EI, NH_4^+) m/z: calcd for $C_{98}H_{145}N_{13}O_{54}$ $(M+Na)^+$, 2390.9; found, 2390.3.

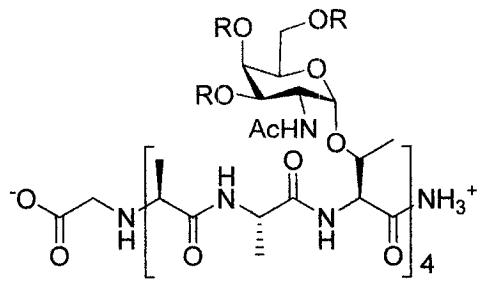
The compound was then readily debenzylated to afford compound **80** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **80** was obtained with a retention time of 9.98 min. LRMS (ES, NH_4^+) m/z: calcd for $C_{66}H_{113}N_{13}O_{38}$ $(M+H)^+$, 1697.7; found, 1697.9.

80 : IR (neat) cm^{-1}: 3334, 1640.

^1H NMR (300 MHz, MeOD) δ: 5.31 (1H), 5.10 (2H), 4.82 (2H), 4.62 (22H), 4.19-3.97 (5.3H), 3.82-3.65 (7.5H), 2.20 (2.5H), 3.17-3.03 (3H), 2.68-2.50 (1.4H), 1.46-2.3 (4H), 2.02-

1.91 (20H), 1.52-1.18 (9H). ^{13}C NMR (93 MHz, $CDCl_3$) 170.7, 170.3, 170.0, 163.4, 163.4, 158.4, 158.4, 132.5, 131.4, 118.3, 118,2, 115.9. 68.9, 68.9, 68.4, 44.9, 35.1, 33.5, 30.6, 29.9, 29.9, 29.9, 29.5, 29.4 , 29.2, 29.1, 29.0, 22.5, 20.5, 19.1, 19.0, 18.7, 17.5, 13.9, 11.9.

α –GalNAc (3OH)(TAA)$_4$G (81)



Gly [Ala Ala Thr]$_4$

77 : R = Ac α-GalNAc(3OAc)(TAA)$_4$G
81 : R = H α-GalNAC(3OH)(TAA)$_4$G

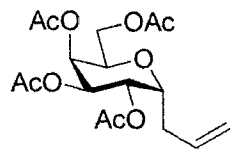
After cleavage from the resin, compound **77** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **77** was obtained with a retention time of 10.95 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{98}H_{149}N_{18}O_{50}$ $(M-H)^+$, 2362.9; found, 2363.7.

The compound was then readily debenzylated to afford compound **81** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **81** was obtained with a retention time of 9.70 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{74}H_{125}N_{17}O_{38}$ $(M-H)^+$, 1859.7; found, 1857.4.

81 : IR (neat) cm^{-1}: 3329, 1638.

^1H NMR (360MHz, $CDCl_3$) δ: 5.00-4.61 (20H), 4.44-4.03 (3H), 3.98-3.41(3.4H), 3.28 (14H), 2.37-1.95 (4H), 1.79-1.12 (38H), 1.00-0.71 (18H); ^{13}C NMR (67MHz, $CDCl_3$) : 169.5.169.4, 169.3, 123.6, 123.6, 117.6, 80.1, 79.9, 79.8, 37.5, 37.2, 37.1, 29.9, 29.7, 29.7, 29.7, 29.1, 28.5, 28.4, 22.5, 22.3, 22.0, 17.9.

3-(2,3,4,6-Tetra-*O*-acetyl-α-D-galactopyranosyl) propene (89)[11]

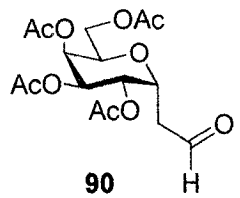


89

This compound was synthesized according to a slightly modified procedure of Roe *et al.*.[11] 2,3,4,6-Tetra-*O*-acetyl-α-D-bromogalactose **29** (1.24 g, 3.01 mmol) was added to a solution of allyl phenyl sulfone (1.42 mL, 7.52 mmol) and bis-tributyl tin (2.19 ml, 4.21 mmol) in benzene (15.4 mL). The solution was degassed with argon and sonicated for 30 minutes. The flask was then sealed and irradiated till completion under a 450W mercury arc lamp. Reaction time for completion fluctuated between 7h and 10h. The reaction required careful monitoring since degradation products appears rapidly after completion. The reaction mixture was directly filtered onto a silica gel column packed with hexane. The organostannanes were flushed with 3 void volumes of hexane, and the remaining sulfone with a hexane/EtOAc (5:1). The product was the eluted with 3:1 hexane:EtOAc and concentrated to afford the allylated derivative as a colorless oil. (873 mg, 2.44 mmol, 87%). Rf 0.25 (hexane:EtOAc 3:1).

^{1}H NMR (360 MHz, $CDCl_3$) δ: 5.78-5.74 (m, 1H), 5.42 (dd, J_1 = 2.4 Hz, J_2 = 3.1 Hz, 1H), 5.28 (dd, J_1 = 4.8 Hz, J_2 = 9.3 Hz, 1H), 5.22 (dd, J_1 = 3.1 Hz, J_2 = 9.3 Hz, 1H), 5.13 (ddd, J_1 = 1.6 Hz, J_2 = 3.0 Hz, J_3 = 16.6 Hz, 1H), 5.12 (ddd, J_1 = 1.4 Hz, J_2 = 3.0 Hz, J_3 = 10.0 Hz, 1H), 4.30 (1H, ddd, J=4.7 Hz, 5.2 Hz, 10.3 Hz), 4.21 (dd, J_1 = 8.9 Hz, J_2 = 12.5 Hz, 1H), 4.11-4.07 (m, 2H), 2.47-2.45 (m, 1H), 2.30-2.28 (m, 1H), 2.12, 2.07, 2.04, 2.03 (4s, 3H each); ^{13}C NMR (67MHz, $CDCl_3$) : 170.5, 170.1, 169.9, 169.8, 133.3, 117.6, 71.4, 68.2, 68.1, 67.9, 67.6, 61.4, 30.9, 20.8, 20.7, 20.7, 20.6; HRMS m/z calcd. for $C_{17}H_{25}O_9$ (M^++H), 373.1499; found, 373.1487.

D-Glycero-L-galacto-octose-3,7-anhydro-2-deoxy-,4,5,6,8 tetraacetate (90)[12]

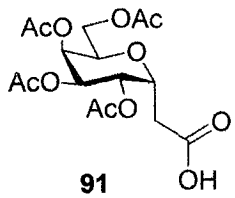


90

This compound was synthesized according to a slightly modified procedure of Eniade *et al.*.[12] Ozone gas was bubbled through a solution of **89** (450 mg, 1.21 mmol) in 10 mL of dry dichloromethane at –78°C until the solution turned blue. Nitrogen was then bubbled through the solution until it turned colorless. Triphenylphosphine (0.8 g, 2.9 mmol) was added, and the solution was allowed to warm up to room temperature and stirred overnight. Dichloromethane was removed under reduced pressure and the residue was purified by flash chromatography (toluene/acetone 4:1) to afford **90** (413 mg, 1.10 mmol, 91%). Rf 0.30 (toluene:acetone 3:1).

IR (film) (CH_2Cl_2) 1738 cm^{-1}; ^1H NMR (360 MHz, $CDCl_3$) δ: 9.64 (s, 1H), 5.15 (dd, J_1 = 5.8 Hz, J_2 = 2.8 Hz, 1H), 5.02 (dd, J_1 = 8.8 Hz, J_2 = 4.7 Hz, 1H), 4.96 (dd, J_1 =8.8 Hz, J_2 = 3.1 Hz, 1H), 4.63-4.61 (1H, m), 4.04 (dd, J_1 = 11.1 Hz, J_2 = 7.5 Hz, 1H), 3.91-3.89 (m, 1H), 3.84 (dd, J_1 = 11.2 Hz, J_2 = 4.6 Hz, 1H), 2.52-2.48 (m, 2H), 1.87 (s, 3H), 1.81 (s, 3H), 1.78 (s, 6H). ^{13}C NMR (67 MHz, $CDCl_3$) δ: 198.4, 169.6, 169.3, 169.1, 169.0, 69.8, 68.1, 68.0, 67.2, 66.9, 61.0, 42.0, 20.1, 20.1, 20.1, 19.9; LRMS (ES, NH_4^+) m/z calcd. for $C_{16}H_{23}O_{10}$ (M^++H), 375.1; found, 375.1.

D-Glycero-L-galacto-octonic acid,3,7-anhydro-2-deoxy-,4,5,6,8 tetraacetate (91)[12]

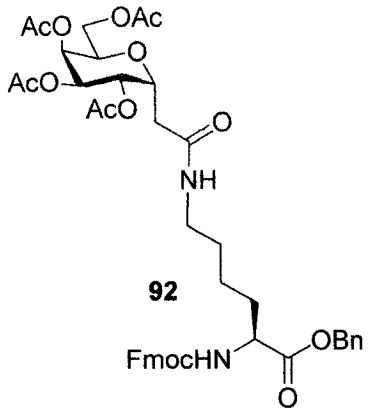


91

This compound was synthesized according to a slightly modified procedure of Eniade *et al.*.[12] To a solution of **90** (2.0 g, 5.4 mmol) in 1:1 (v/v) mixture of *tert*-butyl alcohol and 2-methyl-2-butene (20 mL) was added an aqueous solution of 3.7 g of potassium dihydrogen phosphate and 2.4 g of sodium chlorite in 10 mL of distilled water. The mixture was stirred vigorously at 0°C for 6 hours. The product was extracted with ethyl acetate and washed with 100 mL of water and 50 mL of brine solution. The organic layer was then dried over magnesium sulfate and concentrated. (1.82 g, 4.6 mmol, 86%). Rf 0.40 (CH_2Cl_2:MeOH 10:1).

IR (film) (CH_2Cl_2): 3706-2355, 1748 cm^{-1}. ^1H NMR (360 MHz, $CDCl_3$) δ: 5.38 (t, J = 2.7 Hz, 1H), 5.38-5.37 (m, 1H), 5.17 (dd, J_1 = 3.3, J_2 = 8.7 Hz, 1H), 4.66-4.64 (m, 1H), 4.18-4.08 (m, 3H), 2.71-2.61 (m, 2H), 2.07 (s, 3H), 2.01 (s, 3H), 1.99 (bs, 6H). ^{13}C NMR (360 MHz, $CDCl_3$) δ: 175.5, 170.6, 170.1, 169.9, 169.7, 69.4, 68.9, 67.9, 67.6, 67.1, 61.2, 33.1, 20.8, 20.8, 20.7, 20.7; LRMS (ES, NH_4^+) m/z calcd. for $C_{16}H_{23}O_{11}$ (M^++H), 391.1; found, 391.1.

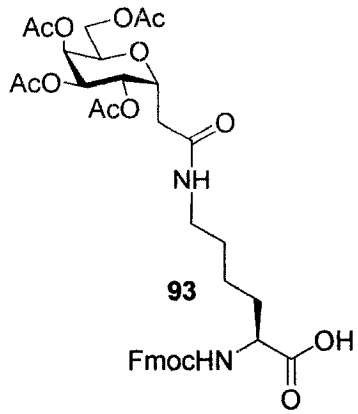
L-Lysine, *N*2-[(9H-fluoren-9-ylmethoxy) carbonyl]-*N*6-(4,5,6,8-tetra-*O*-acetyl-3,7-anhydro-2-deoxy-D-glycero-L-galacto-octonoyl)- benzyl ester (92)[12]



This compound was synthesized according to a slightly modified procedure of Eniade *et al.*.[12] Compound **91** (2.50 g, 6.41 mmol) was pre-activated with HBTU (1.18 g, 2.97 mmol) in dry CH_2Cl_2 (40 mL) at 25°C for 40 minutes. A solution of **AA4** (2.93 g, 6.38 mmol) in dry dichloromethane (40 mL) was then transferred with the activated carboxylic acid **91**. This was followed by the addition of diisopropylethylamine (3.6 mL, 20.0 mmol). The reaction mixture was allowed to stir overnight. After thi timethe compound was extracted with CH_2Cl_2 (50 mL), washed successively with saturated ammonium chloride, water and brine. The organic layer was then dried over magnesium sulphate and concentrated. The crude product was purified with silica gel flash chromatography eluting with CH_2Cl_2/methanol (50:1) to afford compound **92** (2.19 g, 80 %) as a white powder after Et_2O crystallization. Despite these purification steps we were unable to remove the urea derivative (tetramethyl urea) from the reaction. Consequently, compound **92** was carried out to the following step where it was completely purified. Rf 0.30 (CH_2Cl_2-MeOH 50:1).
IR (CH_2Cl_2) cm^{-1}: 3313, 1749; ^1H NMR (300 MHz, $CDCl_3$) δ: 7.77-7.72 (m, 2H), 7.61-7.55 (m, 2H), 7.45-7.23 (m, 9H), 6.13 (t, J = 5.4 Hz, 1H), 5.49 (d, J = 7.7 Hz, 1H), 5.35 (s, 1H), 5.27-5.23 (m, 1H), 5.17-5.10 (m, 3H), 4.69-4.62 (m, 1H), 4.44-4.32 (m, 3H), 4.25-4.21 (m, 2H), 4.11 (s, 2H), 3.26-3.13 (m, 2H), 2.56-2.46 (m, 1H), 2.39-2.33 (m, 1H), 2.09 (s, 3H), 2.02 (s, 6H), 2.0 (s, 3H), 1.92-1.81 (m, 1H), 1.72-1.66 (m, 1H), 1.53-1.43 (m, 2H), 1.32-1.24 (m, 2H). ^{13}C NMR (300 MHz, $CDCl_3$) δ: 172.1, 170.5, 169.9, 169.7, 169.5, 169.2, 155.9, 143.7, 143.6, 141.2, 135.1, 128.5, 128.4, 128.2, 127.6, 127.0, 125.0, 125.0, 119.9, 69.2, 68.9,

67.7, 67.7, 67.1, 66.9, 61.1, 53.6, 47.0, 38.5, 34.2, 32.0, 28.9, 22.4, 20.6, 20.5; HRMS m/z calcd. for $C_{44}H_{50}NO_{14}$ (M^+), 330.3262, found 330.3264.

L-Lysine, *N*2-[(9H-fluoren-9-ylmethoxy) carbonyl]-*N*6-(4,5,6,8-tetra-*O*-acetyl-3,7-anhydro-2-deoxy-D-glycero-L-galacto-octonoyl (93)[12]

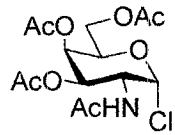


This compound was synthesized according to a slightly modified procedure of Eniade *et al.*.[12] Compound **92** (1.5 g, 1.80 mmol) was dissolved in ethanol (20 mL) and catalytic palladium on charcoal (10 % w/w) was added. The reaction mixture was stirred under hydrogen atmosphere for 24 hours. The catalyst was filtered through a Celite pad and washed with methanol. Solvents were removed under reduced pressure and the residue was eluted from a column of silica gel using a gradient of CH_2Cl_2/methanol (50:1→10:1) to afford a white powder **93** (1.28 g, 96 %) after crystallization in Et_2O. Rf 0.31 (CH_2Cl_2:MeOH 20:1).

IR (film) (CH_2Cl_2) cm^{-1}: 3367, 1748; ^1H NMR (300 MHz, $CDCl_3$) δ: 7.76 (d, J = 7.2 Hz, 2H), 7.60 (d, J = 6Hz, 2H), 7.42-7.37 (m, 2H), 7.33-7.28 (m, 2H), 6.25 (bs, 1H), 5.6 (d, J = 7.5 Hz, 1H), 5.39 (t, J = 3.3, 1H), 5.27-5.23 (m, 1H), 5.16-5.12 (m, 1H), 4.68-4.63 (m, 1H), 4.39 (d, J = 6.6 Hz, 2H), 4.33-4.27 (m, 1H), 4.24-4.19 (m, 1H), 4.1 (m, 2H), 3.28 (m, 2H), 2.58-2.38 (m, 2H), 2.11(bs, 3H), 2.04 (bs, 9H), 1.85 (m, 2H), 1.56 (m, 2H), 1.42 (m, 3H). ^{13}C NMR (300 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. HRMS m/z calcd. for $C_{44}H_{50}NO_{14}$ (M^+), 740.2793, found 740.2795.

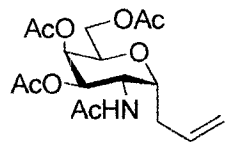
2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl chloride (96)[13]



96

Several literature procedures have been carried out in order to obtain compound **96**. The compound failed to precipitate in Et_2O and is highly unstable during silica gel chromatography purifications. Therefore, as most literature procedures intermediate **96** was not isolated prior to conversion to **97**.

3-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl)-1-propene (97)[11,13b]



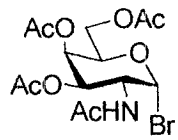
97

This compound was synthesized according to a slightly modified procedure of Cui *et al.*.[11,13b] To a solution of 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl chloride **96** and oxazoline (refer to Scheme 8 Chapter 5) in dry toluene (7 mL) were added allyltributyltin (3.27 mL, 10.7 mmol) and AIBN (87.4 mg, 0.53 mmol). The solution was degassed and then heated for 7 hr at 85°C under argon. The crude products were chromatographed on a column of silica gel with 4:1 EtOAc/hexane to afford a mixture of product **97** and oxazoline **98**. To a solution of the mixture in acetone (20 mL) was added a 1% aqueous solution of HCl (1 mL) to convert the oxazoline into a more polar compound,

1,3,4,6-tetra-*O*-acetyl-2-amino-2-deoxy-α-D-galactopyranose hydrochloride, to facilitate separation. The solvent was then evaporated, and the residue was dissolved in CH_2Cl_2. The solution was washed successively with saturated $NaHCO_3$ and brine, dried (Na_2SO_4) and evaporated. The syrup obtained, which was chromatographed on a column of silica gel with 4:1 EtOAc/hexane to afford **97** as a white solid (293.5 mg, 14% over 2 steps). Rf 0.27 (hexane:EtOAc 4:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 5.67 (m, 1H), 5.44 (d, J = 8.4 Hz, 1H), 5.40 (t, J = 3.2 Hz, 1H), 5.16 (dd, J_1 = 9.4 Hz, J_2 = 3.2 Hz, 1H), 5.00 (dq, J_1 = 17.0 Hz, J_2 = 1.0 Hz, 1H), 4.98 (dq, J_1 = 10.2 Hz, J_2 = 1.0 Hz, 1H), 4.76 (ddd, J_1 = 9.4 Hz, J_2 = 8.4 Hz, J_3 = 4.8 Hz, 1H), 4.37 (m, 1H), 4.36 (dd, J_1 = 11.5 Hz, J_2 = 7.1 Hz, 1H), 4.21 (dd, J_1 = 11.5 Hz, J_2 = 5.4 Hz, 1H), 3.94 (m, 1H), 2.19 (m, 1H), 2.00 (m, 1H), 1.67 (s, 3H), 1.66 (s, 3H), 1.63 (s, 3H), 1.50 (s, 3H). ^{13}C NMR (125 MHz, $CDCl_3$) δ: 170.6, 170.4, 170.0, 169.9, 133.4, 117.4, 71.1, 68.8, 68.2, 66.8, 61.2, 23.0, 20.7, 20.6, 20.5. LRMS (ES, NH_4^+) m/z calcd. for $C_{17}H_{25}NO_8$ (M^++H), 372.2; found, 372.1.

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-galactopyranosyl bromide (99)

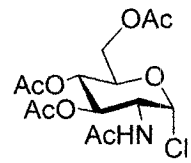


99

To a stirred solution of **94** (100 mg, 257 μmol) cooled at 0°C and under dry atmosphere was added dropwise a of 20% HBr in acetic acid solution (3 mL). The resulting solution was stirred at 10°C for approximately one hour till disappearance of the starting material. The reaction mixture was then diluted with cold dichloromethane (4 mL) and washed with cold water until the pH of the organic layer was 7. The organic layer was then dried over magnesium sulfate, filtered and concentrated under reduced pressure. The residue obtained was purified by flash chromatography (hexane/EtOAc 6:1) to afford after coevaporation with diethyl ether a white solid (35 mg, 87 μmol, 34%). This compound was used without any further purification. Rf 0.30 (hexane:EtOAc 6:1).

LRMS (ES, NH_4^+) m/z calcd. for $C_{14}H_{20}BrNO_8$ (M^++H), 409.0, 411.1; found, 409.0, 411.1.

2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride (100) (α:β=3:1)[13e]



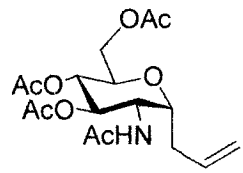
100

This compound was synthesized according to a slightly modified procedure of Horton.[13e] 2-Amino-2-deoxy-α-D-glucose hydrochloride (20 g, 93 mmol) was added to a solution of sodium methoxide (2 g Na, 20 mL MeOH) and stirred for 5 hours. The solution was then filtered using a Büchner funnel and the filtrate washed twice with 20 mL of methanol. The combined filtrates were immediately treated with 12 mL of acetic anhydride and stirred overnight at ~ 25°C. The reaction was refrigerated for several hours at 4°C to complete the crystallization and filtered using a Büchner funnel. The solid was washed with 20 mL of methanol and ether (3 * 20 mL) and dried under vacuum over phosphorus pentoxide. The resulting 2-acetamido-2-deoxy-α-D-glucopyranose was dissolved in a 500 mL flask containing acetyl chloride (20 mL) saturated with HCl. The flask was equipped with a condenser and warmed for 16 hours at 30°C. Dichloromethane (80 mL) was added and the resulting solution vigorously stirred in a large beaker with ice water. The organic layer was extracted and washed again with a cold saturated solution of sodium bicarbonate. The organic layer was dried over magnesium sulfate, filtered and evaporated until 15 mL of solvent remained. Dry ether was added (100 mL) and the solution was allowed to stand for 8 hrs at ~ 25°C and then 5 hrs at 0°C. The resulting precipitate was filtered, washed twice with ether. The resulting white powder (26.4 g, 72.3 mmol, 78% yield) was dried under high vacuum overnight over sodium hydroxide and phosphorus pentoxide.

^{1}H NMR (300 MHz, $CDCl_3$) δ: 6.16 (d, J = 3.6 Hz, 1H), 5.84 (d, J = 8.7 Hz, 1H), 5.29 (t, J = 10.1 Hz, 1H), 5.18 (t, J = 9.1 Hz, 1H), 4.54-5.15 (m, 1H), 4.25-4.23 (m, 2H), 4.10 (d, J = 10.5 Hz, 1H), 2.07 (s, 3H), 2.02 (bs, 6H), 1.96 (s, 3H). ^{13}C NMR (56 MHz, $CDCl_3$) δ:

171.6, 170.8, 170.3, 169.3, 93.8, 71.1, 70.3, 67.1, 61.3, 53.7, 23.3, 20.9, 20.8, 20.7. LRMS (ES, NH_4^+) m/z calcd. for $C_{14}H_{20}ClNO_8$ (M^+-Cl), 330.1; found, 330.1.

3-(2-Acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl) propene (101)[13]

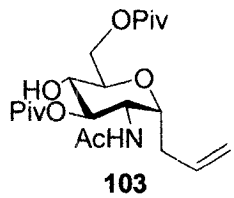


101

This compound was synthesized a combionation of several different procedures.[13] A suspension of 2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl chloride **100** (5.4 g, 16.4 mmol), allyltributyltin (25 mL, 78 mmol) and AIBN (800 mg, 4.6 mmol) in distilled THF (40 mL) was degassed for 15 min. The mixture was then stirred for 2 hours at 55°C and 1 hour at 68°C under dry argon gas. The THF was removed *in vacuo* and the remaining residue was partitioned between 100 mL of CH_3CN and 500 mL of hexane. The CH_3CN layer was extracted with additional hexane (5 * 500 mL) to remove the remaining organostannanes and concentrated to dryness. The desired compound was first purified by silica gel chromatography eluting with 2:1 hexane/EtOAc to obtain a waxy solid that crystallized from a solution of 3:1 hexane/EtOAc to give the product **97** (α:β = 12:1) as a white solid (4.4 g, 11.9 mmol, 73% yield). Rf 0.30 (hexane:EtOAc 2:1).

α anomer : ^{1}H NMR (500 MHz, $CDCl_3$) δ: 6.63 (d, J = 8.5 Hz, 1H), 5.82-5.74 (m, 1H), 5.16-5.10 (m, 3H), 4.94 (t, J = 7.5 Hz, 1H), 4.13 (dt, J_t = 10.0 Hz, J_d = 5.0 Hz, 1H), 4.08 (dd, J_1 = 12 Hz, J_2 = 7.5 Hz, 1H), 4.01 (dt, J_t = 3.1, J_d = 7.4 Hz, 1H), 3.93-3.89 (m, 1H), 3.72 (dt, J_t = 7.5 Hz, J_d = 3.5 Hz, 1H), 2.31-2.25 (m, 1H), 2.16-2.11 (m, 1H), 1.87 (bs, 9H), 1.78 (s, 3H); ^{13}C NMR (125 MHz, $CDCl_3$) δ: 170.8, 170.7, 170.0, 169.2, 133.4, 117.6, 71.2, 70.3, 70.0, 68.2, 61.8, 50.4, 31.8, 23.1, 20.8, 20.8, 20.7. LRMS (ES, NH_4^+) m/z: 372.1 (M^++H), 394.1 (M^++Na), 765.1 ($2M^+$+Na). HRMS m/z calcd. for $C_{14}H_{20}NO_5$ (M^+-allyl), 330.1189; found, 330.1188.

3-(2-Acetamido-2-deoxy-6,3-di-pivaloyl-α-D-glucopyranosyl) propene (103)[14]



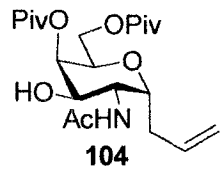
103

This compound was synthesized according to a slightly modified procedure of Cipolla *et al.*.[14] After being dried overnight under reduce pressure over phosphorus pentoxide, triol **102** (1.0 g, 4.07 mmol) was dissolved in a 3:1 mixture of distilled CH_2Cl_2/pyridine. The solution was cooled at -20°C and pivaloyl chloride (1 mL, 8.15 mmol) was added dropwise over 30 min. The reaction was stirred for 6 hours. After this time, the mixture was then allowed to warm up to 25°C and stirred for an additional 10 hours. The resulting solution was washed successively with 5% HCl and brine. The organic layer was dried over $MgSO_4$, filtered and evaporated. The crude residue was purified by flash chromatography (hexane/EtOAc, 1:1). Compound **103** was obtained as an inseparable mixture (α:β = 12:1) (1.54 g, 3.74, 92%). Rf 0.27 (hexane:EtOAc 1:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 6.18 (d, J = 8.4 Hz, 1H), 5.81-5.68 (m, 1H), 5.12-4.92 (m, 3H), 4.52 (dd, J_1 = 6.3 Hz, J_2 = 12 Hz, 1H), 4.23-4.11 (m, 3H), 3.76 (dt, J_t = 6,9 J_d = 2.7, 1H), 3.50 (t, J = 7.5 Hz, 1H), 2.47-2.37 (m, 1H), 2.29-2.22 (m, 1H), 1.91 (bs, 3H), 1.19 (bs, 18H); ^{13}C NMR (75 MHz, $CDCl_3$) δ: 180.1, 179.6, 170.4, 134.1, 117.8, 72.9, 72.5, 72.1, 68.8, 63.2, 51.4, 39.4, 32.0, 27.6, 27.4, 23.6, 23.6 LRMS (ES, H^+) m/z: 414.2 (M^++H).

.LRMS (ES, NH_4^+) m/z calcd. for $C_{21}H_{35}NO_7$ (M^++H), 414.2; found, 414.2.

3-(2-Acetamido-2-deoxy-6,4-di-pivaloyl-α-D-galactopyranosyl) propene (104)[14]

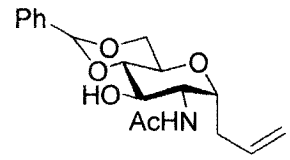


104

This compound was synthesized according to the procedure of Cipolla *et al.*.[14] Glucoside **103** (93 mg, 0.22 mmol) was dissolved in a 2:1 pyridine/CH_2Cl_2 mixture (2 mL) and the solution was cooled to 0°C. Triflic anhydride (93 μL, 0.53 mmol) was added portionwise until complete consumption of the starting material. Water (222 μL, 12.3 mmol) was then added, and the reaction mixture, allowed to stir overnight at room temperature. The reaction was diluted with CH_2Cl_2, the organic layer washed sequentially with 5% aqueous HCl and water, dried over Na_2SO_4, filtered and evaporated under reduced pressure. Purification by flash chromatography (1:1) hexane:EtOAc + 0.2% EtOH afforded 78 mg of **104** as a yellow oil (84% yield). The characterization was consistent with that reported in reference 14. Rf 0.23 (hexane:EtOAc 1:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 5.86 (d, J = 7.5 Hz, 1H), 5.81-5.68 (m, 1H), 5.13-5.03 (m, 3H), 4.59 (t, J = 10.2 Hz, 1H), 4.38-4.32 (m, 1H), 4.18-4.08 (m, 2H), 4.01 (dd, J_1 = 3.6 Hz, J_2 = 6.3 Hz, 1H), 3.98 (d, J = 3.9 Hz, 1H), 2.35-2.25 (m, 1H), 2.21-2.12 (m, 1H), 1.98 (bs, 3H), 1.21 (bs, 9H) 1.16 (bs, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 178.3, 178.2, 171.3, 134.1, 117.3, 70.7, 68.7, 67.3, 61.0, 60.5, 52.2, 39.1, 38.7, 27.2, 27.2, 23.0, 23.0. .LRMS (ES, NH_4^+) m/z calcd. for $C_{21}H_{35}NO_7$ (M^++H), 414.2; found, 414.2, 436.1 (M^++Na), 849.2 ($2M^+$+Na).

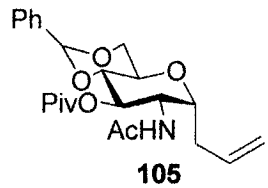
3-(2-Acetamido-4,6-benzylidene-2-deoxy-α-D-glucopyranosyl) propene[15]



This compound was synthesized according to a slightly modified procedure of Barroca *et al.*.[15] A solution of sodium methoxide (5 mL, pH = 10) was slowly added to 230 mg of 3-(2-acetamido-3,4,6-tri-*O*-acetyl-2-deoxy-α-D-glucopyranosyl)propene **101**, (620 μmol) and stirred for 10 hours. The solution was then neutralized using acidic resin IR-120, concentrated to dryness and dried overnight under reduce pressure over phosphorus pentoxide. Without further purification 2-acetamido-(2-deoxy-α-D-glucopyranosyl)propene **102** (151 mg, 616 μmol) was dissolved in distilled benzaldehyde (1.9 mL, 18.5 mmol) and degassed for 30 min under dry argon. Anhydrous TFA (100 μL, 925 μmol) was added dropwise at 0°C to the solution and the mixture stirred for 3 hours at 25°C. When no starting material remained as observed by TLC, the reaction was concentrated to dryness. The viscous residue obtained was triturated with hexane (3 * 20 mL) and then H_2O (2 * 20 mL) and filtered. After recrystallization from methanol and drying under reduce pressure over phosphorus pentoxide overnight, 172mg, (518 μmol, 84% yield) of the desired compound was obtained as a white powder (α:β = 12:1).

^{1}H NMR (500 MHz, CD_3OD) δ: 7.60-7.33 (m, 5H), 5.82-5.74 (m, 1H), 5.60 (s, 1H), 5.15 (dd, J_1 = 1.5Hz, J_2 = 17Hz, 1H), 5.07 (d, J = 10 Hz, 1H), 4.18-4.09 (m, 3H), 3.87 (t, J = 8.0 Hz, 3H), 3.71 (t, J = 10 Hz, 1H), 3.62-3.58 (m, 1H), 3.52 (t, J = 8.0 Hz, 1H), 2.61-2.55 (m, 1H), 2.32-2.27 (m, 1H), 1.98 (bs, 3H),. ^{13}C NMR (125 MHz, CD_3OD) δ: 172.1, 137.6, 132.5, 129.2, 128.4, 127.9, 127.5, 126.0, 115.7, 101.6, 83.0, 74.2, 68.7, 67.7, 63.4, 54.2, 30.3, 20.1. LRMS (CI, Isobutylene) m/z: 334.2 (M^++H), 316 (M^+-OH). HRMS m/z calcd. for $C_{15}H_{18}NO_8$ (M^+-allyl) 292.1185; found, 292.1185.

3-(2-Acetamido-4,6-benzylidene-2-deoxy-3-pivaloyl-α-D-glucopyranosyl) propene (105)[15]

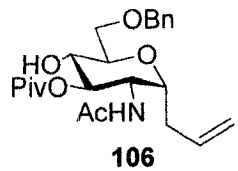


105

This compound was synthesized according to a slightly modified procedure of Barroca *et al.*.[15] To a solution of 3-(2-Acetamido-4,6-benzylidene-2-deoxy-α-D-glucopyranosyl)propene (124 mg, 372 μmol) in CH_2Cl_2 (1.5 mL), freshly distilled pyridine (500 μL), freshly distilled pivaloyl chloride (139 μL, 1.12 mmol) and a catalytic amount of DMAP were added at 0°C under dry argon. The reaction was stirred for 20 hours at 25°C and quenched by adding methanol (1 mL). The resulting solution was washed successively with H_2O, saturated $NaHCO_3$ and brine, dried ($MgSO_4$), and concentrated to dryness. The resulting residue was purified by silica gel chromatography (hexane/EtOAc, 2:1) and a mixture of diastereoisomers (α:β = 12:1) was obtained as a white powder (148 mg, 354 μmol, 95%). Rf 0.30 (hexane/EtOAc 2:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.41-7.31 (m, 5H), 5.97 (d, J = 7.5 Hz, 1H), 5.75-5.61 (m, 1H), 5.54 (s, 1H), 5.19 (t, J = 10.2Hz, 1H), 5.05 (d, J = 7.2 Hz, 1H), 5.01 (s, 1H), 4.42-4.34 (m, 1H), 4.31-4.20 (m, 2H), 3.76-3.66 (m, 2H), 3.60 (dd, J_1 = 4.2, J_2 = 9.6 Hz, 1H), 2.53-2.41 (m, 1H), 2.34-2.25 (m, 1H), 1.89 (bs, 3H), 1.17 (bs, 9H). ^{13}C NMR (56 MHz, $CDCl_3$) δ: 180.4, 170.3, 137.2,133.6, 129.1, 128.4, 126.0, 117.8, 101.3, 79.9, 74.4, 70.1, 69.4, 64.1, 53.4, 39.3, 30.9, 27.2, 23.2 LRMS (CI, H^+) m/z: 418.0 (M^++H), (M^++K) 458.0. HRMS m/z calcd. for $C_{23}H_{31}NO_6$ (M^+) 417.2151; found, 417.2151.

3-(2-Acetamido-6-*O*-benzyl-2-deoxy-3-pivaloyl-α-D-glucopyranosyl) propene (106)[15]

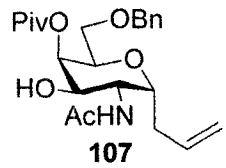


106

This compound was synthesized according to a slightly modified procedure of Barroca *et al.*.[15] Compound **105** (43 mg, 103 μmol), 4Å powdered molecular sieves and triethylsilane (96 μL, 598 μmol) in distilled CH_2Cl_2 were stirred for 1 hour at 25°C under dry argon. The reaction was cooled to 0°C and boron trifluoride diethyl etherate (26 μL, 206 μmol) was added. The reaction was then allowed to warm up to 25°C was stirred for 2 hours. After this time, triethylamine (66 μL, 473 μmol) was added and the reaction was filtered and concentrated to dryness. The resulting residue was dissolved in CH_2Cl_2 (5 mL) and washed successively with water, saturated $NaHCO_3$ and water, dried ($MgSO_4$) and concentrated. The residue was purified by silica gel chromatography (hexane/EtOAc, 1:1) and the two anomers were separated to afford the α-anomer as white powder (33 mg, 77% yield). After purification by silica gel chromatography, the purity of the α anomer was verified by NMR to be higher than 98/2. Rf 0.32 (hexane/EtOAc 1:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.31-7.21 (m, 5H), 6.28 (d, J = 8.7 Hz, 1H), 5.81-5.67 (m, 1H), 5.09 (d, J = 17.1 Hz, 1H), 5.05 (d, J = 9.3 Hz, 1H), 4.99 (dd, J = 7.5Hz, J = 9.6 Hz, 1H), 4.55 (d, J = 12 Hz, 1H), 4.49 (d, J = 12Hz, 1H), 4.24-4.16 (m, 1H), 4.09-4.04 (m, 1H), 3.75-3.68 (m, 3H), 3.60-3.56 (m, 1H), 3.05 (bs, 1H), 2.53-2.41 (m, 1H), 2.34-2.19 (m, 1H), 1.87 (bs, 3H), 1.17 (bs, 9H); ^{13}C NMR (56 MHz, $CDCl_3$) δ: 180.1, 170.1, 137.8, 134.0, 128.6, 128.0, 127.9, 117.5, 73.9, 72.9, 72.5, 71.8, 70.6, 70.5, 51.6, 39.2, 31.4, 27.2, 23.3 LRMS (CI, Isobutylene) m/z: 420 (M^++H), 401 (M^+-18). HRMS m/z calcd. for $C_{20}H_{28}NO_6$ (M^+-allyl), 378.1916, found 378.1916.

3-(2-Acetamido-2-deoxy-6,4-di-pivaloyl-α-D-galactopyranosyl) propene (107)[15]

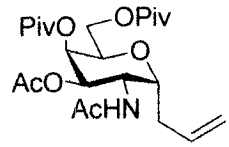


107

This compound was synthesized according to a slightly modified procedure of Barroca *et al.*.[15] Triflic anhydride (1.8 mL) was added dropwise at -20°C under dry argon to a solution of **106** (2.8 g, 6.68 mmol) in distilled 1,2-dichloroethane (100 mL) and distilled pyridine (10 mL). The reaction mixture was stirred at 0°C for 1 hour. After this time 20 mL of water were added and the mixture was stirred under reflux for 1.5 hours. The solution was then washed sequentially with 5% HCl, brine, dried ($MgSO_4$) and concentrated. The residue was eluted from a column of silica gel using a gradient Hexane/EtOAc (1:0→1:1) to afford a white powder (2.32 g, 5.54 mmol, 83%). Rf 0.30 (hexane:EtOAc 1:1).

^{1}H NMR (500 MHz, $CDCl_3$) δ: 7.32-7.27 (m, 5H), 6.10 (d, J = 7.5 Hz, 1H), 5.80-5.72 (m, 1H), 5.18 (t, J = 3.6 Hz, 1H), 5.08 (d, J = 13.2 Hz, 1H), 5.05 (d, J = 7.8Hz, 1H), 4.53 (d, J = 12 Hz, 1H) 4.47 (d, J = 12Hz, 1H), 4.35-4.31 (m, 1H), 4.17-4.13 (m, 1H), 4.08-4.03 (m, 1H), 3.95 (dd, J_1 = 3.3 Hz, J_2 = 7.2 Hz, 1H), 3.71 (dd, J_1 = 6.8 Hz, J_2 = 10.4 Hz, 1H), 3.51 (dd, J_1 = 4.2 Hz, J_2 = 10.5 Hz, 1H), 3.42 (bs, 1H), 2.36-2.30 (m, 1H), 2.18 (dt, J_t = 6.3 Hz, J_d = 15.3 Hz, 1H), 1.99 (bs, 3H), 1.16 (bs, 9H); ^{13}C NMR (125 MHz, $CDCl_3$) δ: 178.4, 171.3, 137.8, 134.2, 128.7, 128.1, 128.1, 117.7, 73.7, 73.7, 71.4, 69.0, 67.9, 67.8, 52.3, 39.3, 33.2, 27.3, 23.4 LRMS (CI, Isobutylene) m/z: 420 (M^++H), 401 (M^+-18). HRMS m/z calcd. for $C_{20}H_{28}NO_6$ (M^+-allyl), 378.1916; found, 378.1916.

3-(2-Acetamido-2-deoxy-4,6-di-pivaloyl-α-D-galactopyranosyl) propene (109)

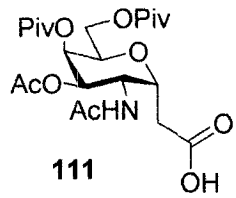


109

Under inert atmosphere at 25°C, compound **108** (200 mg, 484 μmol) was dissolved in an 1:2 pyridine/Ac_2O solution (15 mL) and a catalytic amount of DMAP (20 mg) was added. The reaction was stirred for 10 hours at 25°C. After this time, the misture was diluted in CH_2Cl_2 (25 mL) and then respectively washed with satd solution of $CuSO_4$ (3x30 mL), H_2O (30 mL), dried ($MgSO_4$) and concentrated. The residue was eluted from a column of silica gel using a gradient of hexane/EtOAc (1:0→1:2) to afford a white powder (216 mg, 474 μmol, 98% yield). Rf 0.45 (hexane:EtOAc 1:2)

^1H NMR (300 MHz, $CDCl_3$) δ: 5.80 (d, J = 7.8 Hz, 1H), 5.77-5.65 (m, 1H), 5;28 (t, J = 3 Hz, 1H), 5.15-5.02 (m, 3H), 4.45-4.37 (m, 1H), 4.33-4.21 (m, 2H), 4.08-4.03 (m, 1H), 3.94 (dd, J_1 = 11.1 Hz, J_2 = 5.4 Hz, 1H), 2.42-2.31 (m, 1H), 2.25-2.19 (m, 1H), 1.98 (bs, 3H), 1.92 (bs, 3H), 1.18 (bs, 9H) 1.13 (bs, 9H); ^{13}C NMR (75 MHz, $CDCl_3$) δ: 178.2, 177.6, 170.2, 166.4, 133.7, 117.7, 77.4, 68.9, 68.7, 68.4, 66.6, 61.2, 49.1, 39.3, 38.9, 27.2, 27.2, 23.4, 22.3, 20.9. HRMS m/z calcd. for $C_{23}H_{37}NO_8$ (M^++H) 455.2519; found, 455.2514.

D-Glycero-L-galacto-octonic acid, 4-acetamido-5-acetate-3,7-anhydro-2,4-deoxy-6,8-dipivaloyl (111)[16]

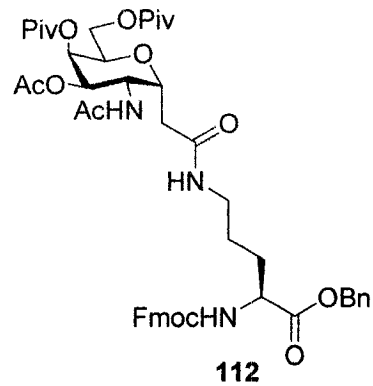


111

This compound was synthesized according to a slightly modified procedure of Buskas *et al.*.[16] Allylated pyranose **109** (250 mg, 529 μmol) was dissolved in a 1:1 solution of dichloromethane and acetonitrile. Sodium periodate (452 mg, 2.12 mmol) was added followed by a catalytic amount of ruthenium trichloride trihydrate. The reaction was allowed to stir at room temperature overnight. The solution was then filtered through Celite, transferred to a separatory funnel with dichloromethane, and acidified to pH 2 by dropwise addition of 20% hydrochloric acid to pH 2. The organic layer was washed successively with water and brine, dried over magnesium sulfate, filtered and concentrated to afford the carboxylic acid derivative **111** in 84 % yield (210 mg, 444 μmol). Like for most of our acid building block the ^{1}H NMR spectrum of **111** gave broad signals which were difficult to interpret. However the observation of a single spot by TLC, a ^{1}H signal around 12 ppm and the expected HRMS let us conclude that the compound obtained was teh target compound . Thus compound **111** was converted without further purification to **112**.

LRMS (ES, H^+) m/z: 474.0 (M^++H), 496.2 (M^++Na), 968.6 ($2M^+$+Na). HRMS m/z calcd. for $C_{22}H_{35}NO_{10}$ (M^++H), 473.2261; found, 473.2260.

L-Ornithine, N^{α}-[(9H-fluoren-9-ylmethoxy) carbonyl]-*N*6-(4-acetamido-5-acetate-3,7-anhydro-2,4-deoxy-6,8-dipivaloyl) D-glycero-L-galacto-octonoyl-benzyl ester (112)[12]



112

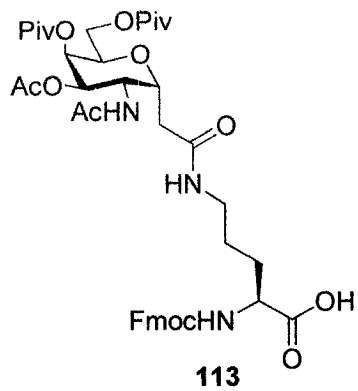
This compound was synthesized according to a slightly modified procedure of Eniade *et al.*.[12] Compound **111** (255 mg, 539 μmol) was pre-activated with HBTU (215 mg, 539 μmol) in dry CH_2Cl_2 (10 mL) at 25°C for 40 minutes. A solution of α-*N*-Fmoc ornithine benzyl ester (496 mg, 1.08 mmol) in dry dichloromethane (10 mL) was then transferred with the activated carboxylic acid, followed by the addition of diisopropylethylamine (600 μL). The reaction mixture was allowed to stir overnight. After this time, the mixture was extracted with CH_2Cl_2 (10 mL), washed successively with saturated ammonium chloride, water and brine. The organic layer was dried over magnesium sulphate and concentrated. The crude product was purified with silica gel flash chromatography eluting with (50:1) CH_2Cl_2/MeOH to afford compound **112** (412 mg, 458 μmol, 85 %) as a white powder after Et_2O crystallization. Despite these purification steps we were unable to remove the urea derivative (tetramethyl urea) from the mixture. Consequently, compound **112** was carried out to the following step where it was completely purify. Rf 0.22 (CH_2Cl_2:MeOH 50:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.75 (d, J = 7.1 Hz, 2H), 7.55 (d, J = 7.2 Hz, 2H), 7.37-7.22 (m, 9H), 6.85 (d , J = 7.5 Hz, 1H), 6.76 (s, 1H), 5.92 (d , J = 8.1 Hz, 1H), 5.32 (s, 1H), 5.13-5.09 (m, 3H), 4.69-4.71 (m, 1H), 6.85 (dd , J_1 = 9.9 Hz, J_2 = 3.0 Hz, 1H), 4.40-3.96 (m, 7H), 3.21-3.18 (m, 2H), 2.62-2.41 (m, 2H), 1.95 (s, 3H), 1.89 (s, 3H), 1.22 (s, 9H), 1.14 (s, 9H).

^{13}C NMR (75 MHz, $CDCl_3$) δ: 178.5, 177.9, 172.5, 171.3, 170.9, 170.5, 156.7, 144.2, 144.0, 141.6, 135.6, 129.0, 128.9, 128.6, 128.1, 127.5, 125.5, 120.4, 77.4, 77.369.1, 68.5, 67.6, 67.5, 66.4, 61.2, 54.2, 48.6, 47.4, 39.5, 39.4, 39.1, 39.0, 31.4, 30.2, 30.1, 27.5, 27.4, 26.0,

23.5, 21;2 LRMS (ES, H^+) m/z: 900.4 (M^++H), 923.6 (M^++Na), 939.2 (M^++K), 1823.9 ($2M^+$+Na), 2723 ($3M^+$+Na). HRMS m/z calcd. for $C_{49}H_{61}N_3O_{13}$ (M^++H), 899.4204; found, 899.4207.

L-Ornithine, N^{α}-[(9H-fluoren-9-ylmethoxy) carbonyl]-*N*6-(4-acetamido-5-acetate-3,7-anhydro-2,4-deoxy-6,8-dipivaloyl) D-glycero-L-galacto-octonoyl (113)[12]

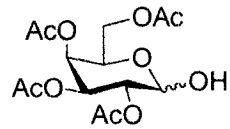


113

This compound was synthesized according to a slightly modified procedure of Eniade *et al.*.[12] Compound **112** (400 mg, 445 μmol) was dissolved in ethanol (7 mL) and a catalytic amount of palladium on charcoal (10 % w/w) was added. The reaction mixture was stirred under a hydrogen atmosphere for 24 hours. The catalyst was filtered through a Celite pad and washed with methanol. Solvents were removed under reduce pressure and the residue was eluted from a column of silica gel using a gradient CH_2Cl_2/methanol (50:1→10:1) to afford a white powder **113** (349 mg, 432 μmol, 97 %) after crystallization in Et_2O. Rf 0.30 (CH_2Cl_2:MeOH 30:1)

HRMS m/z calcd. for $C_{44}H_{50}NO_{14}$ (M^+), 809.3735; found, 809.3735.

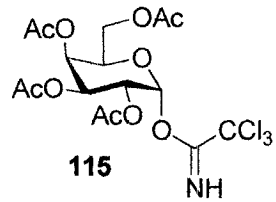
2,3,4,6-tetra-*O*-acetyl-α-D-galactopyranose (114)[17]



114

This compound was synthesized according to the procedure of de la Fuente *et al.*.[17] D-galactose pentaacetate (5.0 g, 12.8 mmol) was dissolved in DMF (50 mL), hydrazine acetate (1.42 g, 15.4 mmol) was added and the reaction was stirred at 50°C for 1 hour. The reaction was then diluted with EtOAc (50 mL) and washed with brine. The organic layer was dried over magnesium sulphate, filtered and concentrated. The residue was eluted from a column of silica gel using a gradient of EtOAc/hexane (1:4→1:1) to afford **114** as a white solid (4 g, 90%). The characterizations were consistent with that reported in reference 17.
LRMS (ES, H^+) m/z: 349.3 (M^++H), 330.2 (M^+-H_2O).

Trichloroacetimidate 2,3,4,6-tetra-*O*-acetyl-Dgalactopyranoside (115)

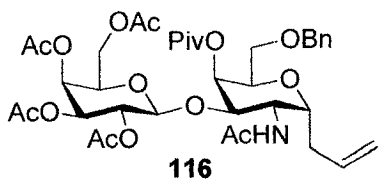


115

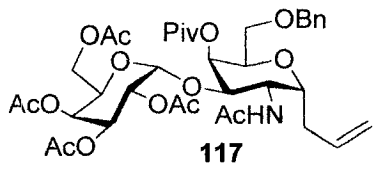
Compound **114** (520 mg, 1.49 mmol) was dissolved in dry dichloromethane (100 mL) under an argon atmosphere at -10°C. The solution was degassed and sonicated for 10 minutes with 3Å molecular sieves. The mixture was then treated with Cl_3CCN (1.5 mL, 14.9 mmol, 10 equiv.) and DBU (275 μL, 1.79 mmol) and the mixture was stirred at -10°C for 15 minutes or until completion as observed by TLC (EtOAc:hexane, 1:1) the mixture was filtered over Celite and silica gel (toluene/EtOAc 6:1), and the filtrate was concentrated to give **115** as an amorphous solid (736 mg, 1.49 mmol, 100%). The compound was verified by NMR and only the α-anomer was present. Without any further purification compound **115** was coupled with **107** and **103** to afford compounds **116**, **117** and **118**. Rf 0.33 (hexane:EtOAc 1:1).

2,3,4,6-Tetra-*O*-acetyl-β-D-galactopyranosyl-(1→3)-3-(2-acetamido-6-*O*-benzyl-2-deoxy-4-pivaloyl-α-D-galactopyranosyl) propene (116)

2,3,4,6-Tetra-*O*-acetyl-α-D-galactopyranosyl-(1→3)-3-(2-acetamido-6-*O*-benzyl-2-deoxy-4-pivaloyl-α-D-galactopyranosyl) propene (117)



116



117

Both compounds (**116** and **117**) can be crystallized by slow diffusion process by dissolving the disaccharide in Et_2O and slowly diffusing hexane over 3 days at 0°C. Using this technique colorless needle shape crystals were obtained.

Method using the glycoside bromide 29 as glycoside donor

To a solution of compound **107** (20 mg, 48 μmol) and **29** (100 mg, 0.24 mmol) in dry dichloromethane (5 mL) was added K_2CO_3 (41 mg, 0.29 mmol) and AgOTf (50 mg, 192 μmol), under an argon atmosphere at -42°C. The solution was degassed and sonicated for 10 minutes with 4Å molecular sieves at -42°C and let stirred 1 hour at 0°C and 48h at 25°C. The mixture was filtered over Celite, washed respectively with $NaHCO_3$ (15 mL) and H_2O (15 mL), dried ($MgSO_4$) and concentrated. The residue was eluted from a column of silica gel using a gradient of Et_2O/CH_2Cl_2 (8:1→4:1) to afford **116** and **117** (α/β 1.5/1) as a white powder with 51% yield. **116** Rf 0.28; **117** Rf 0.26 (Et_2O:CH_2Cl_2 6:1).

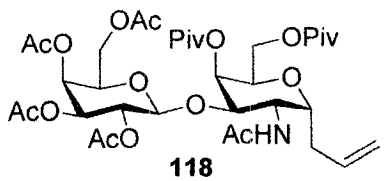
Method using the glycoside trichloroimidate 115 as glycoside donor

To a solution of compound **107** (500 mg, 1.20 mmol), TMSOTf (21 μL, 120 μmol) and 4Å molecular sieves in dry dichloromethane (40 mL) under an argon atmosphere and at 0°C was added **115** (1.3 g, 2.6 mmol) dropwise. The solution was stirred for 30 minutes at 0°C and 6h at 25°C. The mixture was filtered over Celite and concentrated. The residue was eluted from a column of silica gel using a gradient from Et_2O/CH_2Cl_2 (8:1→4:1) to afford **116** and **117** (α/β 1/3) as a white powderwith 73% yield.

116: ^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.31-7.19 (m, 5H), 6.26 (d, J = 7.2 Hz, 1H), 6.09 (d, J = 4.1 Hz, 1H), 5.73-5.64 (m, 1H), 5.41-5.33 (m, 1H), 5.12-5.00 (m, 2H), 4.95-4.88 (m, 2H), 4.51 (d, J = 12.5 Hz, 1H), 4.49 (d, J = 12.5 Hz, 1H), 4.26 (t, J = 6.2 Hz, 1H), 4.22 (t, J = 6.1 Hz, 1H), 4.12 (t, J = 6.3 Hz, 1H), 4.10-3.93 (m, 6H), 3.43 (d, J = 10.9 Hz, 1H), 2.32-2.24 (m, 1H), 2.19-1.85 (m, 16H), 1.07 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 177.6, 171.2, 170.5, 170.4, 170.1, 170.0, 138.1, 133.6, 128.4, 128.0, 127.8, 122.2, 117.7, 97.4, 75.6, 73.8, 73.2, 72.1, 69.3, 66.9, 66.4, 66.2, 65.2, 61.3, 60.4, 51.3, 38.8, 34.7, 27.1, 26.3, 23.0, 21.0, 20.7.LRMS (ES, H^+) m/z: 777.3 (M^++H). HRMS m/z calcd. for $C_{38}H_{51}NO_{16}$ (M^++H), 777.3208; found, 777.3205.

117: ^{1}H NMR (300 MHz, $CDCl_3$) δ: 7.33-7.17 (m, 5H), 6.11 (d, J = 3.8 Hz, 1H), 5.87 (d, J = 7.2 Hz, 1H), 5.78-5.58 (m, 1H), 5.41-5.33 (m, 1H), 5.25-5.15 (m, 1H), 5.11-4.88 (m, 4H), 4.63-4.60 (m, 1H), 4.58-4.40 (m, 2H), 4.26-4.40 (m, 1H), 4.18-3.92 (m, 7H), 3.43 (d, J = 11.1 Hz, 1H), 2.32-2.21 (m, 1H), 2.17-1.84 (m, 16H) 1.15-1.02 (m, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 177.9, 170.7, 170.7, 170.4, 169.8, 169.8, 138.8, 133.8, 128.6, 128.2, 128.0, 125.8, 117.9, 104.0, 86.4, 80.7, 79.8, 76.3, 75.4, 73.4, 69.7, 67.1, 66.4, 62.9, 61.9, 51.4, 38.9, 34.7, 27.2, 25.9, 23.4, 21.1, 20.9. LRMS (ES, H^+) m/z: 777.3 (M^++H). HRMS m/z calcd. for $C_{38}H_{51}NO_{16}$ (M^++H), 777.3208; found, 777.3205.

2,3,4,6-tetra-*O*-acetyl-β-D-galactopyranosyl-(1→3)-3-(2-acetamido-2-deoxy-4,6-dipivaloyl-α-D-galactopyranosyl) propene (118)

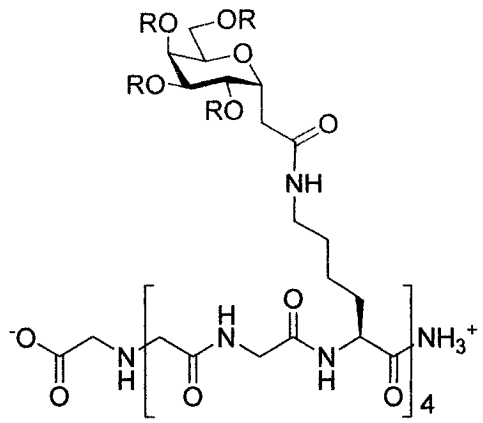


118

The same trichloroimidate procedure than for the formation of **116** was performed on compound **103**. We obtained **118** with 73% and an α/β ratio of 1/3. Rf 0.25 (Et_2O/CH_2Cl_2 6:1).

^{1}H NMR (300 MHz, $CDCl_3$) δ: 5.92 (d, J = 8.3 Hz, 1H), 5.74-5.60 (m, 1H), 5.39-5.16 (m, 2H), 5.09-4.94 (m, 4H), 4.82 (dd, J_1 = 10.4 Hz, J_2 = 13.0, 1H), 4.59 (d, J = 8.2 Hz, 1H), 4.16-3.82 (m, 8H), 2.30-2.13 (m, 1H), 2.12-1.89 (m, 16H), 1.17 (s, 9H), 1.14 (s, 9H). ^{13}C NMR (75 MHz, $CDCl_3$) δ: 177.9, 177.6, 170.5, 170.3, 170.3, 170.2, 169.6, 133.5, 118.0, 99.2, 73.0, 72.7, 71.0, 70.9, 69.0, 67.0, 65.8, 61.1, 61.1, 59.1, 49.6, 39.0, 38.8, 35.6, 27.3, 27,2, 23.4, 20.9, 20.8, 20.7.LRMS (ES, H^+) m/z: 744.4 (M^++H). HRMS m/z calcd. for $C_{35}H_{53}NO_{16}$ (M^++H) 743.3364, found 743.3366.

[Gal(OH)$_4$](LGG)4)G



119 : R_1 = Ac [Gal(OAc)$_4$](LGG)$_4$G
123 : R_1 = OH [Gal(OH)$_4$](LGG)$_4$G

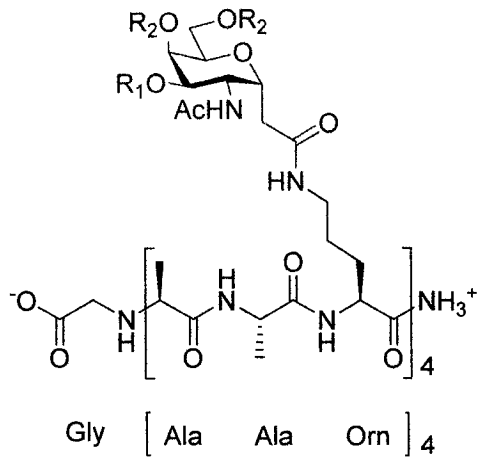
After cleavage from the resin, compound **119** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **119** was obtained with a retention time of 10.05 min.

The compound was then readily debenzylated to afford compound **123** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **123** was obtained with a retention time of 7.88 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{104}H_{157}N_{17}O_{54}$ (M-H) 2531.1; found, 1530.8.

123 : IR (neat) cm^{-1}: 3339, 1629.

^1H NMR (500 MHz, CD_3OD) δ: 4.77 (3H), 4.72-4.65 (10H), 4.35 (1H), 4.18 (1H), 3.92-3.80 (3.8H), 3.72-3.47 (4H), 3.09-3.0 (2.5H), 2.08-2.42 (1H), 2.23-2.05 (11H), 1.75-1.59 (2H), 1.49-1.17 (3.5H). ^{13}C NMR (93 MHz, CD_3OD) δ: 174.4, 174.4, 170.5, 170.1, 169.8, 156.2, 127.7, 127.1, 125.1, 11.9, 69.2, 69.0, 67.8, 67.7, 67.0, 61.2, 53.5, 50.7, 47.1, 39.1, 28.7, 22.1, 20.7.

[GalNAc(OH)$_3$](OAA)$_4$G



122 : R_1 = Ac; R_2 = Piv [GalNAc(OAc)(OPiv)$_2$](OAA)$_4$G
126 : R_1, R_2 = OH [GalNAc(OH)$_3$](OAA)$_4$G

After cleavage from the resin, compound **122** was purified by C-18 reverse phase HPLC[5] (isocrat, 40 to 70% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **122** was obtained with a retention time of 10.24 min.

The compound was then readily debenzylated to afford compound **126** which was also purified by C-18 reverse phase HPLC[5] (isocrat, 5 to 20% B in A). A: 0.1% aqueous trifluoroacetic acid and B: 0.1% trifluoroacetic acid in acetonitrile. Compound **126** was obtained with a retention time of 8.03 min. (MALDI, MeOH, negative ion) m/z: calcd for $C_{87}H_{145}N_{21}O_{38}$ (M-H), 2091.0; found, 2090.2.

126 : IR (neat) cm^{-1}: 3330, 1638.

^{1}H NMR (500 MHz, CD_3OD) δ: 4.77 (17H), 4.24-3.70 (2H), 3.58-3.44 (2.6H), 2.92 (1.1H), 2.44-2.14 (2.7H), 1.43-2.21 (6H). ^{13}C NMR (93 MHz, CD_3OD) δ: 172.4, 172.1, 172.2, 171.1, 170.5, 158.5, 158.4, 158.2, 118.7, 118.1, 114.9, 71.3, 71.1, 69.4, 69.4, 68.9, 67.8, 67.6, 63.0, 43.1, 39.9, 34.2, 30.1, 30.0, 24.2, 22.1, 20.7, 16.9.

[1] Istvan, S; Lajos, K. *Synthesis* **1986**, 4, 303.

Liu, M.; Young, Jr., V.G.; Lohani, S.; Live, D.; Barany G. *Carbohydr. Res.* **2005**, 340, 1273.

[2] Chang, C.D.; Waki, M.; Meienhofer, F.; Lundell, E.O.; Haug, J.D. *Int. J. Pept. Protein Res.* **1980**, 15, 59.

[3] Frische, K.; Meldal, M.; Werdelin, O.; Mouritsen, S.; Jensen, T.; Galli-Stampino, L.; Bock, K. *J. Pept. Sci.*, **1996**, 2, 212.

[4] (a) Varon, D.; Lioy, E.; Patarroyo, M.E.; Schratt, X.; Unverzagt, C. *Aust. J. Chem.*, **2002**, 55(1 & 2), 161.

(b) Salvador, L.A.; Elofsson, Mikael; Kihlberg, Jan. *Tetrahedron* **1995**, 51, 5643.

[5] Polaris U5 C-18 reverse phase HPLC 250 mm x 4.6mm (analytical), 250mm x 21.2mm (semiprep) flow rate of 0.6mL/min (analytical), 12.5 mL/min (semiprep).

[6] Motawia, M. S.; Olsen, C. E.; Denyer, K.; Smith, A. M.; Moller, B. L. *Carbohydr. Res.* **2001**, 330, 309.

[7] Lacombe, J.M.; Pavia A.A. *J.Org. Chem.* **1983**, 48, 2557.

[8] The experimental procedure and yields were improved compare to the original publication: Lemieux, R.U.; Ratcliffe, R.M. *Can. J. Chem.* **1979**, 57, 1244.

[9] Both α and β anomer were separated and characterized:

llen, J. R.; Harris, C. R.; Danishefsky, S. J. *J. Am. Chem. Soc.* **2001**, 123, 1890.

[10] Kunz, H.; Birnbach, S. *Angew. Chem. Int. Ed.* **1986**, 98, 354.

Kunz, H.; Birnbach, S.; Wernig, P. *Carbohydr. Res.* **1990**, 202, 207.

[11] The reaction conditions were slightly modified from the procedure: Roe, B. A.; Boojamra, C. G.; Griggs J. L.; Bertozzi, C. *J. Org. Chem.* **1996**, 61, 6442.

[12] Compound **90** and **91** have been previously published but were never fully characterized. Eniade, A.; Murphy, A.V.; Landreau, G.; Ben, R.N. *Bioconjugate Chemistry* **2001**, 12, 817.

[13] (a) Keck, G. E.; Yates, J. B. *J. Am. Chem. Soc.* **1982**, 104, 5829.

(b) Cui, J.; Horton, D. *Carbohydr. Res.* **1998**, 309, 319.

(c) Tarasiejska, Z.; Jeanloz, R. W. *J. Am. Chem. Soc.*; **1958**, 80, 6325.

(d) Heidlas, J. E.; Lees, W. J.; Pale, P.; Whitesides, G. M. *J. Org. Chem.* **1992**, 57, 146.

(e) Horton D. *Methods Carbohydr. Chem.* **1972**, 6, 282.

(f) Baker, B. R.; Joseph, J.; Schaub, R. E.; Williams, J. H. *J. Org. Chem.* **1954**, 19, 1786.

[14] Cipolla L.; La Ferla, B.; Lay, L.; Peri, F.; Nicotra, F. *Tetrahedron: Asymmetry* **2000**, 11, 295.

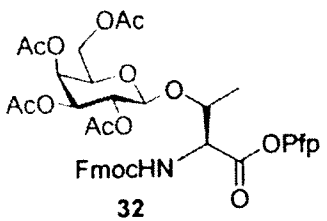
[15] Barroca, N.; Jacquinet, J.-C. *Carbohydr. Res.* **2002**, 337, 673.

[16] Buskas, T.; Soderberg, E.; Konradsson, P.; Fraser-Reid, B. *J. Org. Chem.* **2000**, 65, 958.

[17] de la Fuente, J.M.; Penadés, S. *Tetrahedron. Asym.* **2002**, 13, 1879.

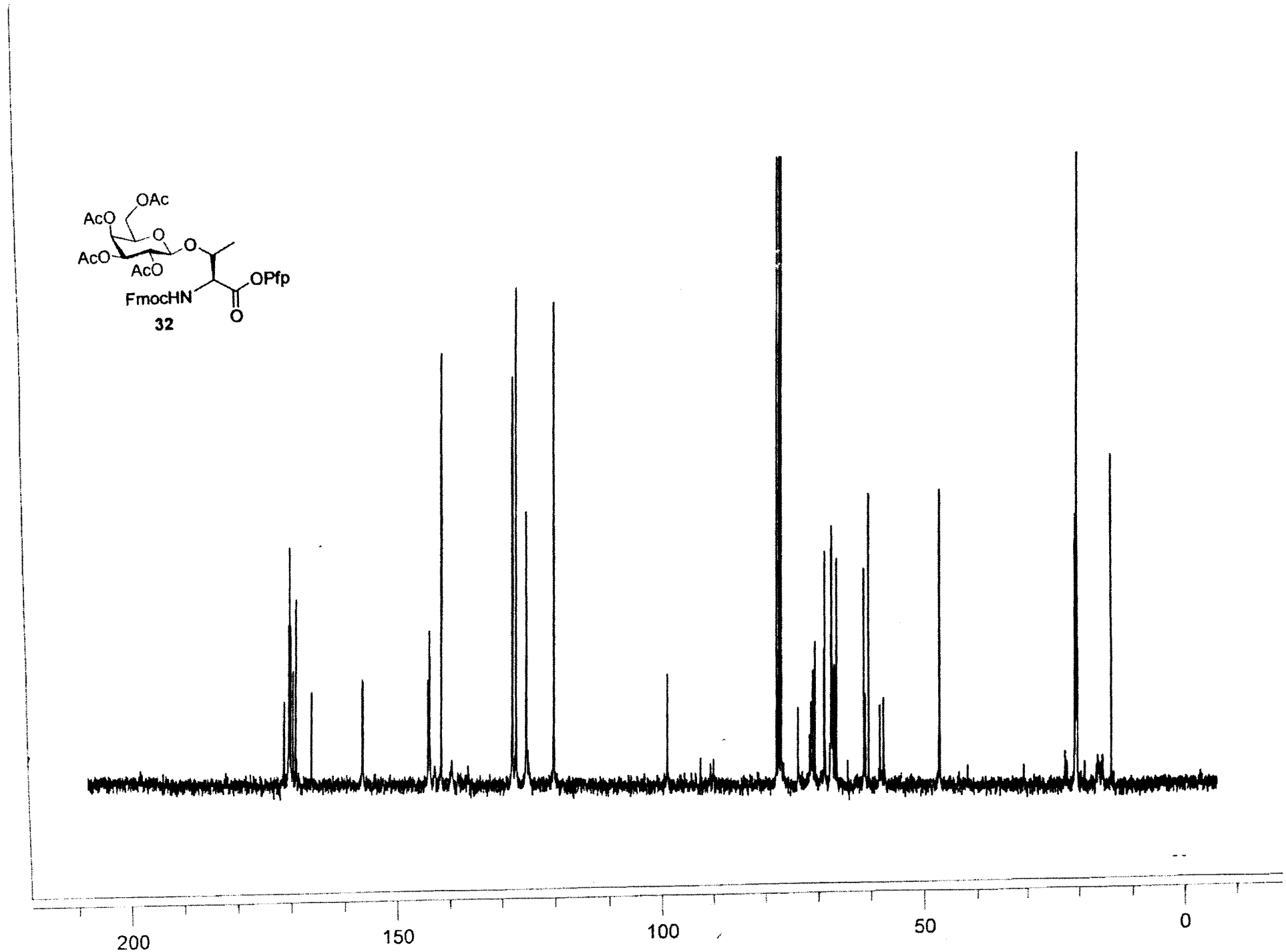
APPENDIX

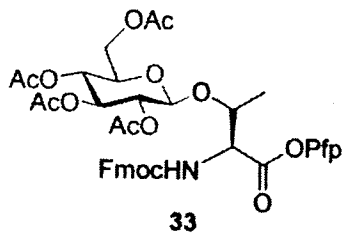
^{1}H AND ^{13}C SPECTRA


OAc
AcO
O
AcO
O
AcO
OPfp
FmocHN
O

32

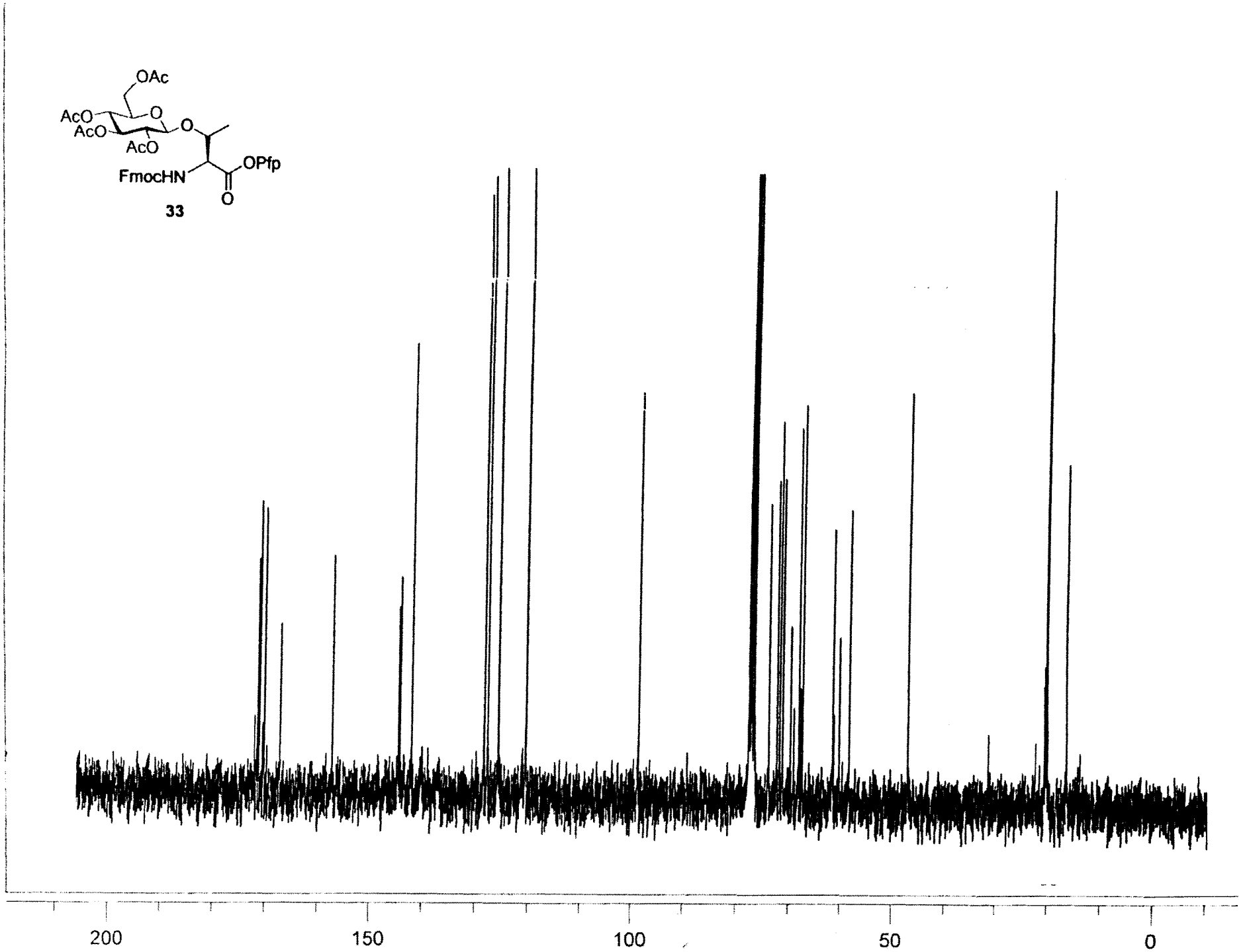
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

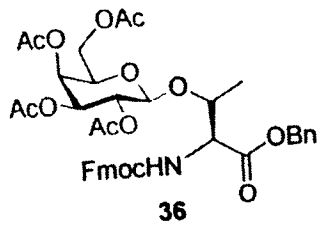

OAc
AcO
AcO
AcO
O
O
OPfp
FmocHN
O
32
200
150
100
50
0


OAc
AcO
AcO
AcO
O
O
FmocHN
OPfp
O

33

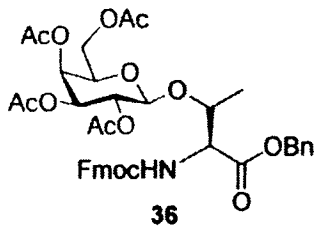
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)


OAc
AcO
AcO
AcO
O
O
FmocHN
OPfp
O
33
200
150
100
50
0

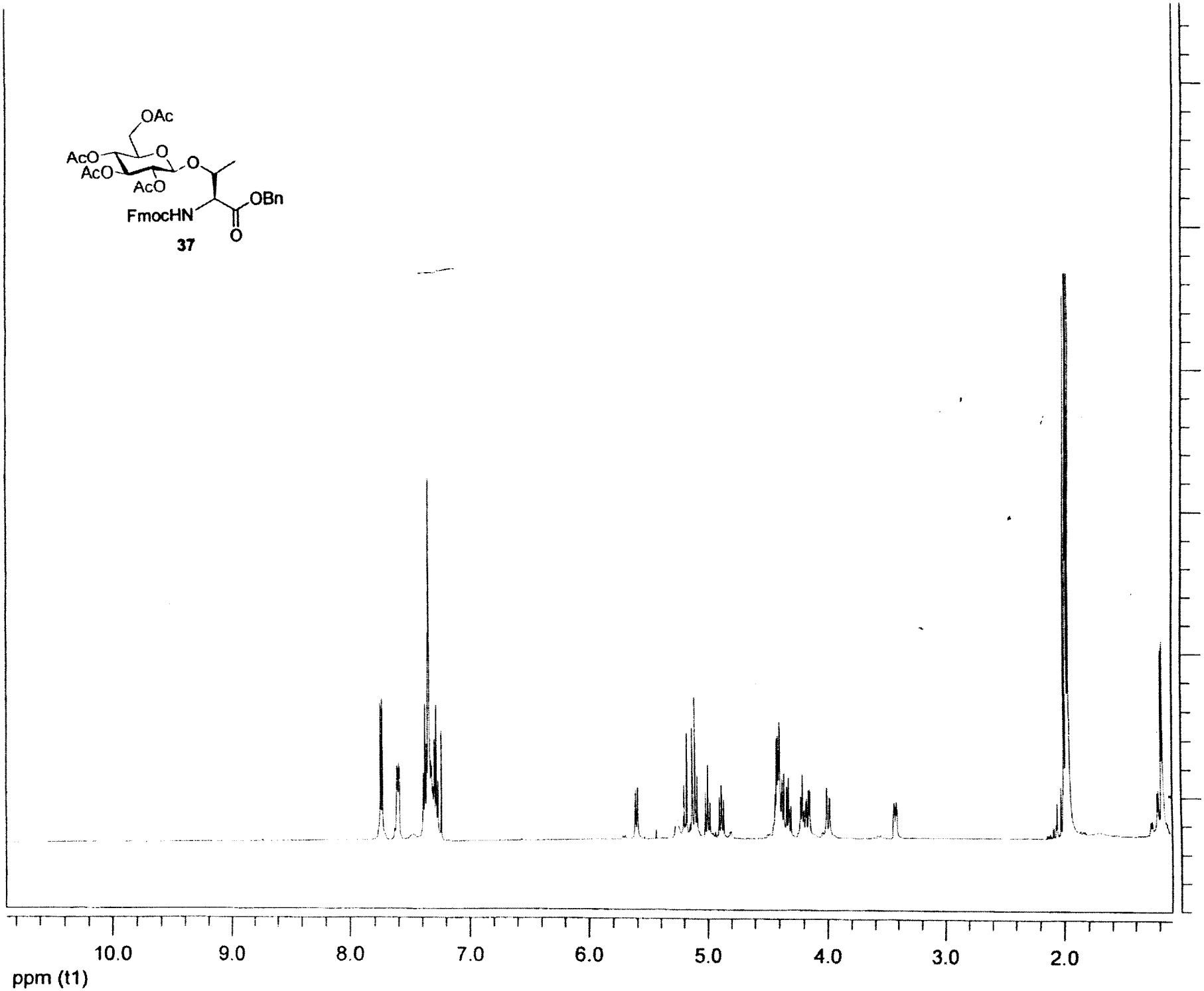

AcO
OAc
AcO
O
O
AcO
FmocHN
OBn
O

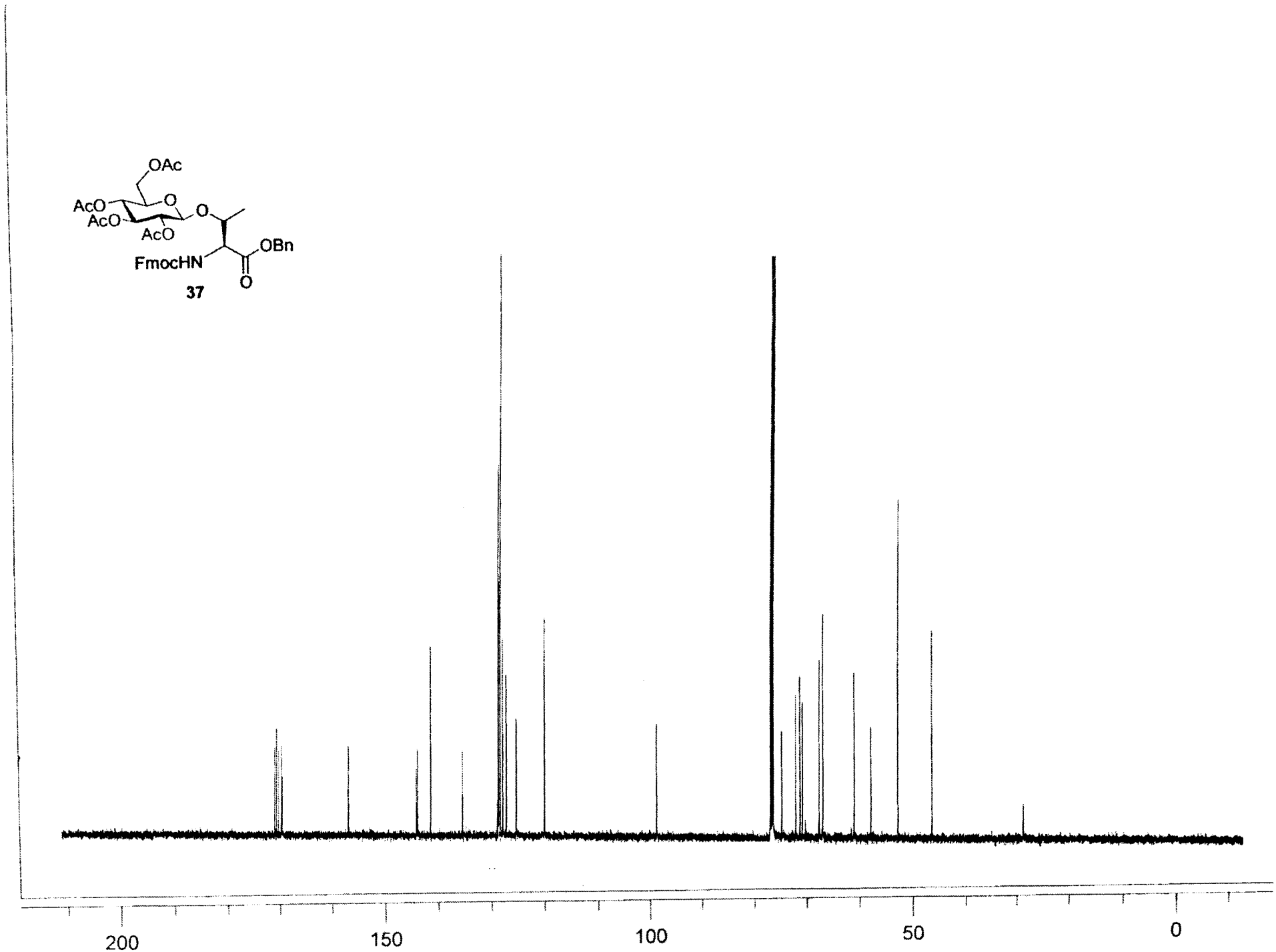
36

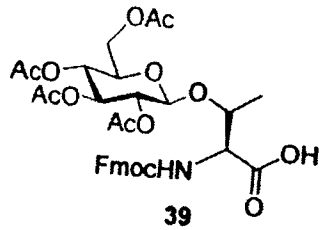
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)



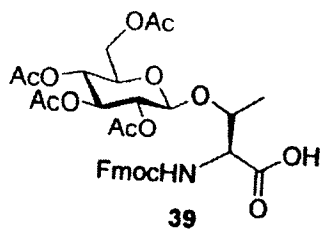
36


OAc
AcO
AcO
AcO
O
O
OBn
FmocHN
O
37
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
ppm (t1)

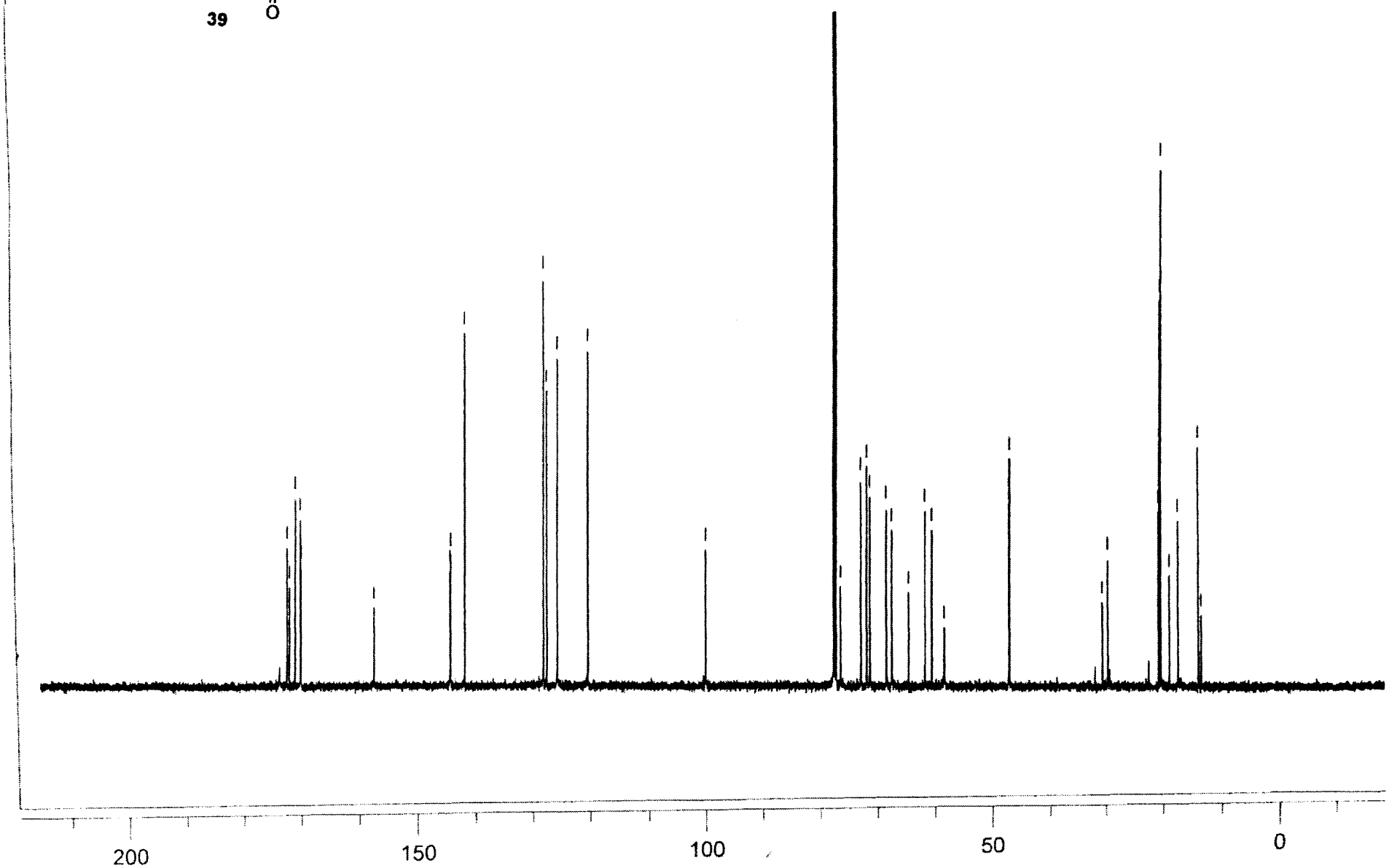

OAc
AcO
AcO
AcO
O
O
OBn
FmocHN
O
37
200
150
100
50
0

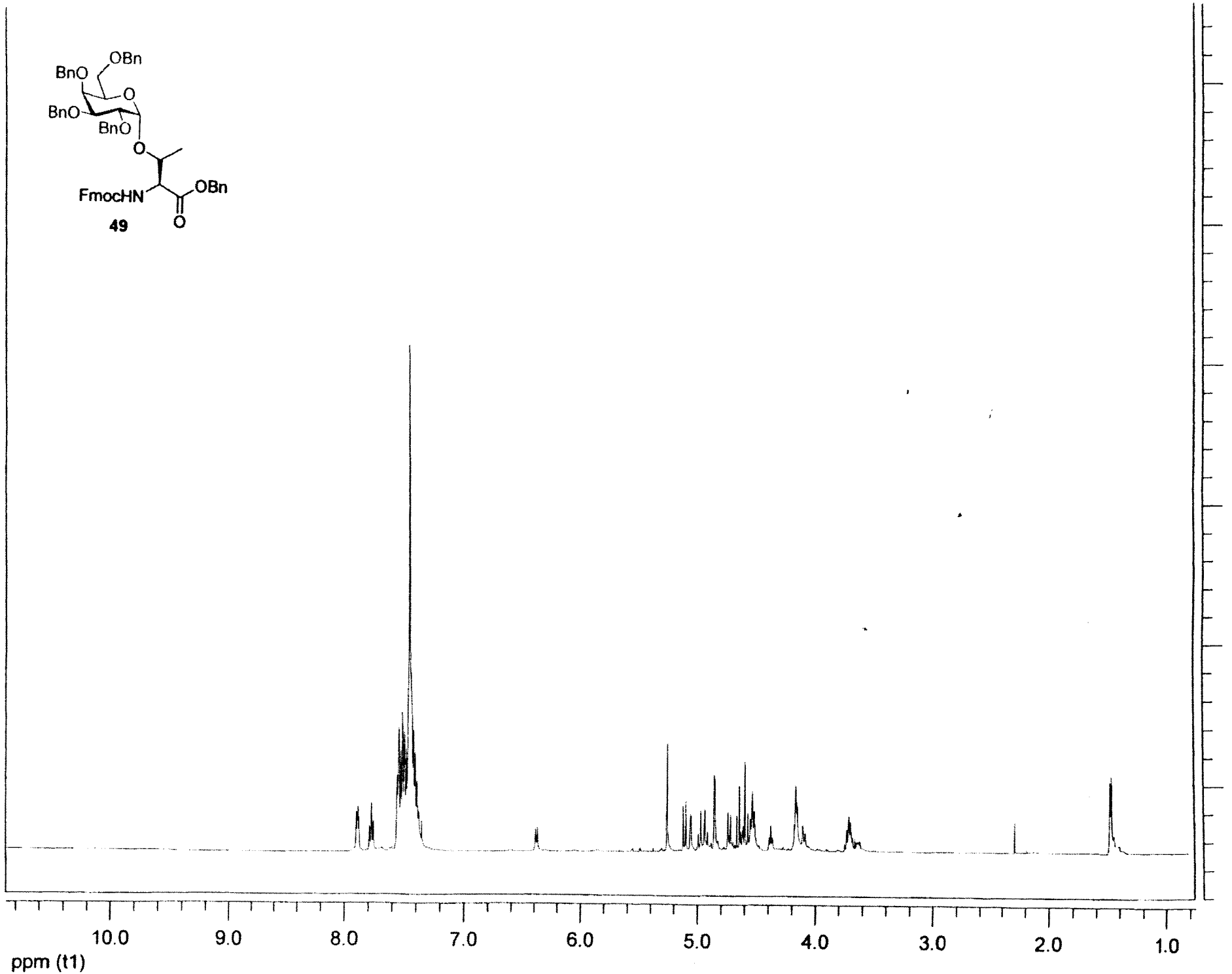


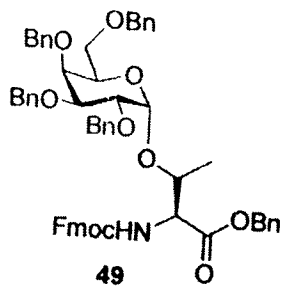
39



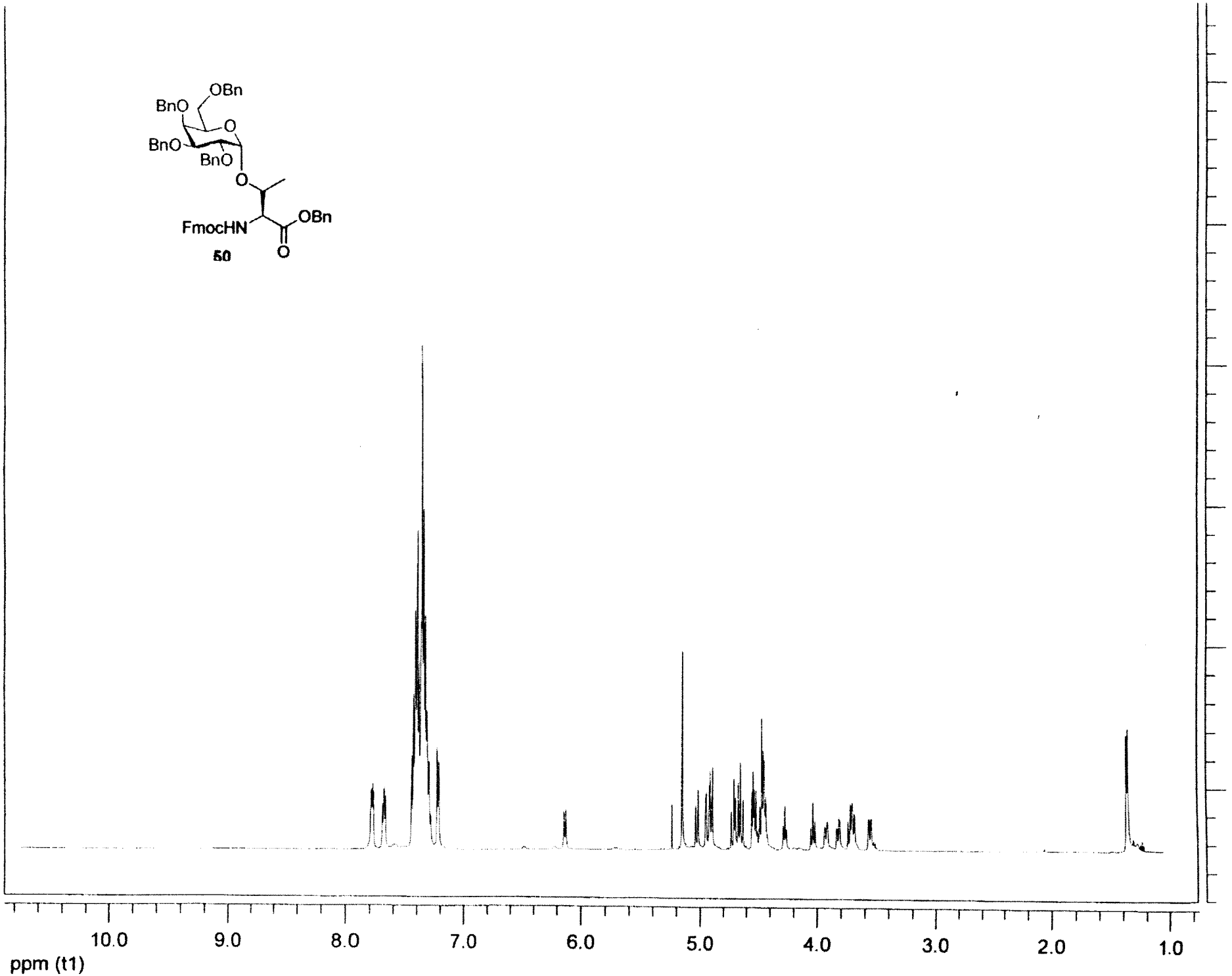
39

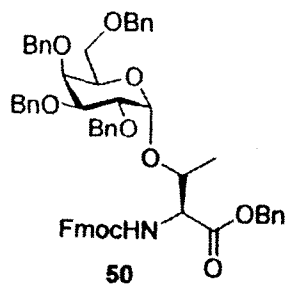



OBn
BnO
O
BnO
BnO
O
FmocHN
OBn
O
49
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

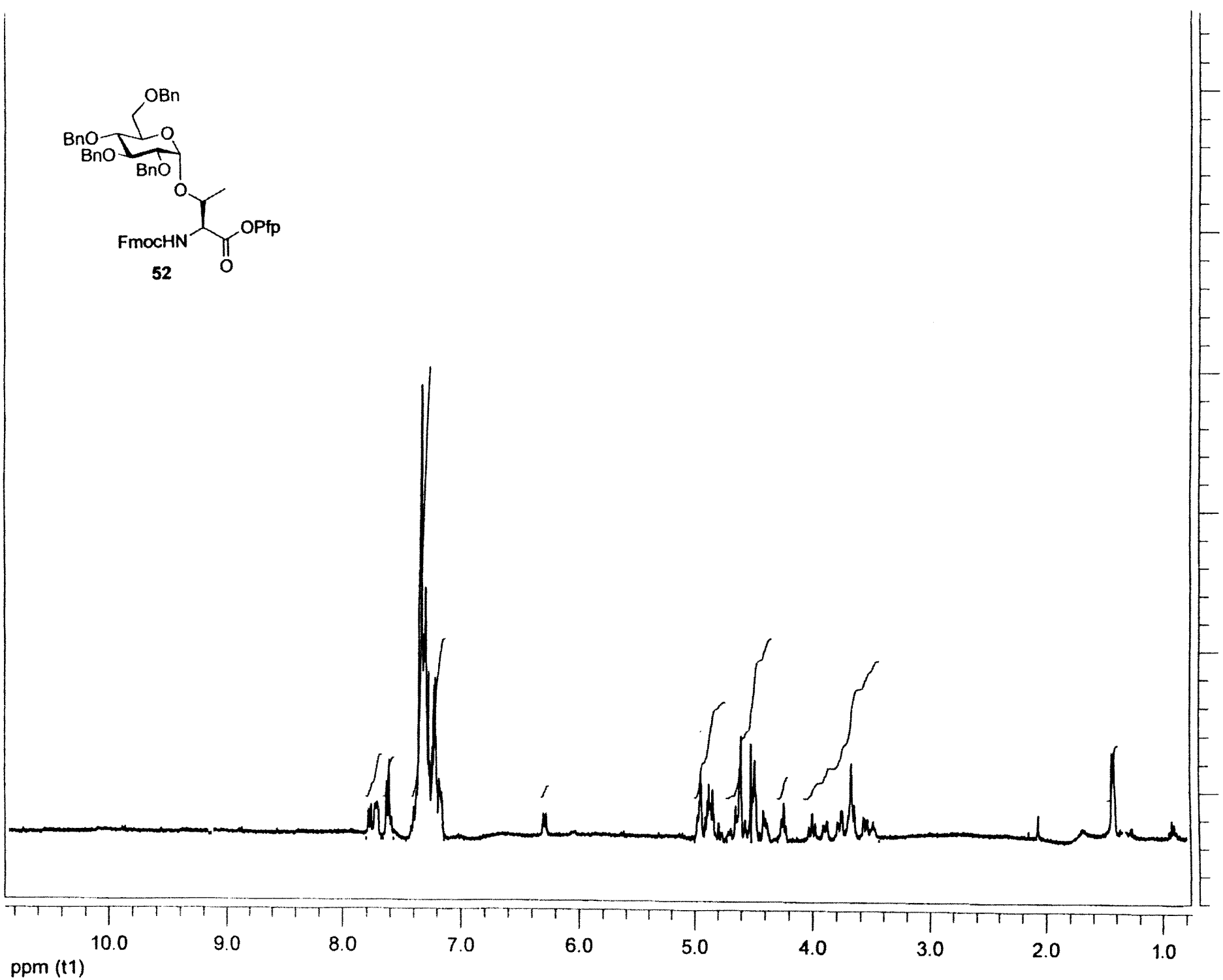


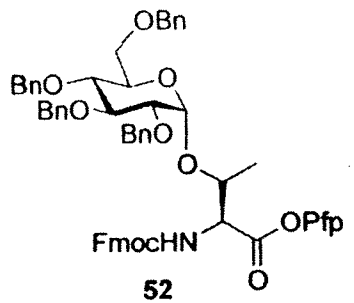
49


OBn
BnO
O
BnO
BnO
O
FmocHN
OBn
O
50
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)



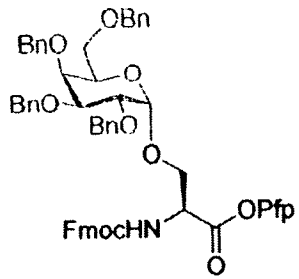
50


OBn
BnO
BnO
BnO
O
O
FmocHN
OPfp
O
52
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)


OBn
BnO
BnO
BnO
O
FmocHN
OPfp
O

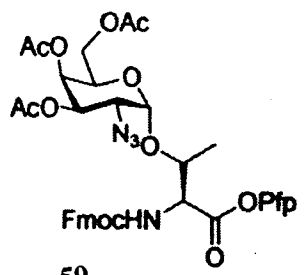
52

200
180
160
140
120
100
80
60
40
20
0
PPM


OBn
BnO
O
BnO
BnO
O
FmocHN
OPfp
O

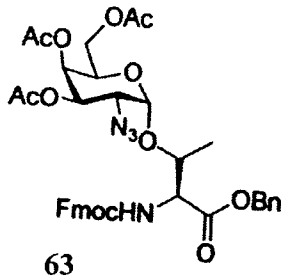
53

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

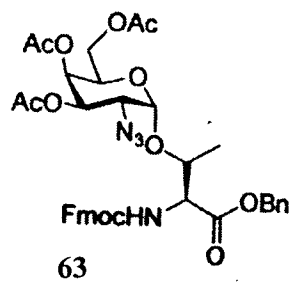

OAc
AcO
O
AcO
N_3
O
FmocHN
OPfp
O

59

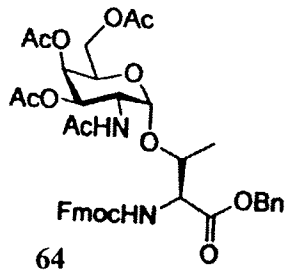
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)



63

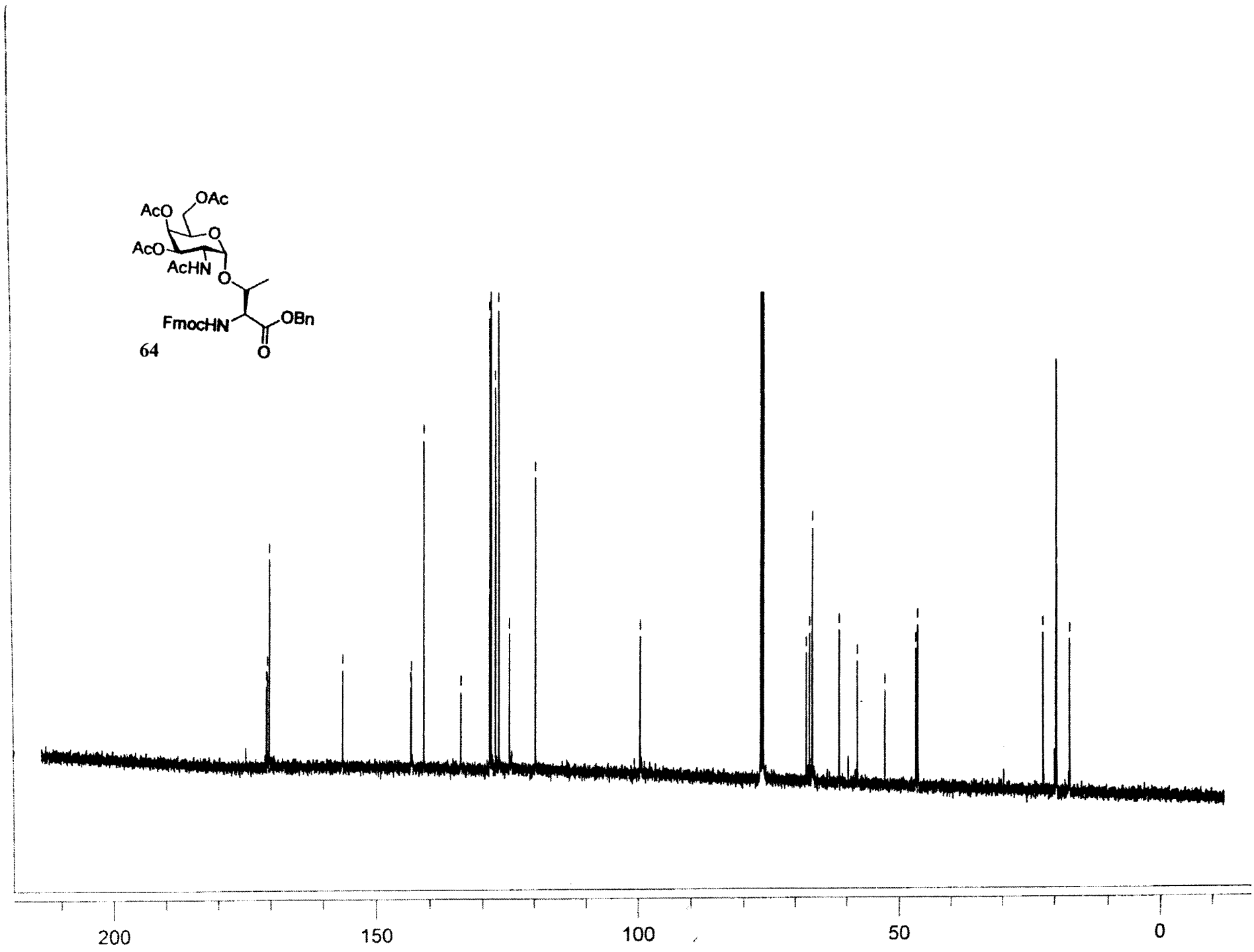


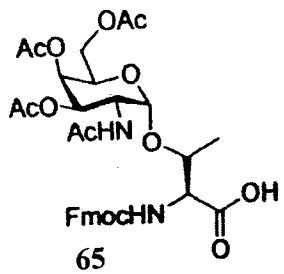
63


AcO
OAc
AcO
O
AcHN
O
FmocHN
OBn
O

64

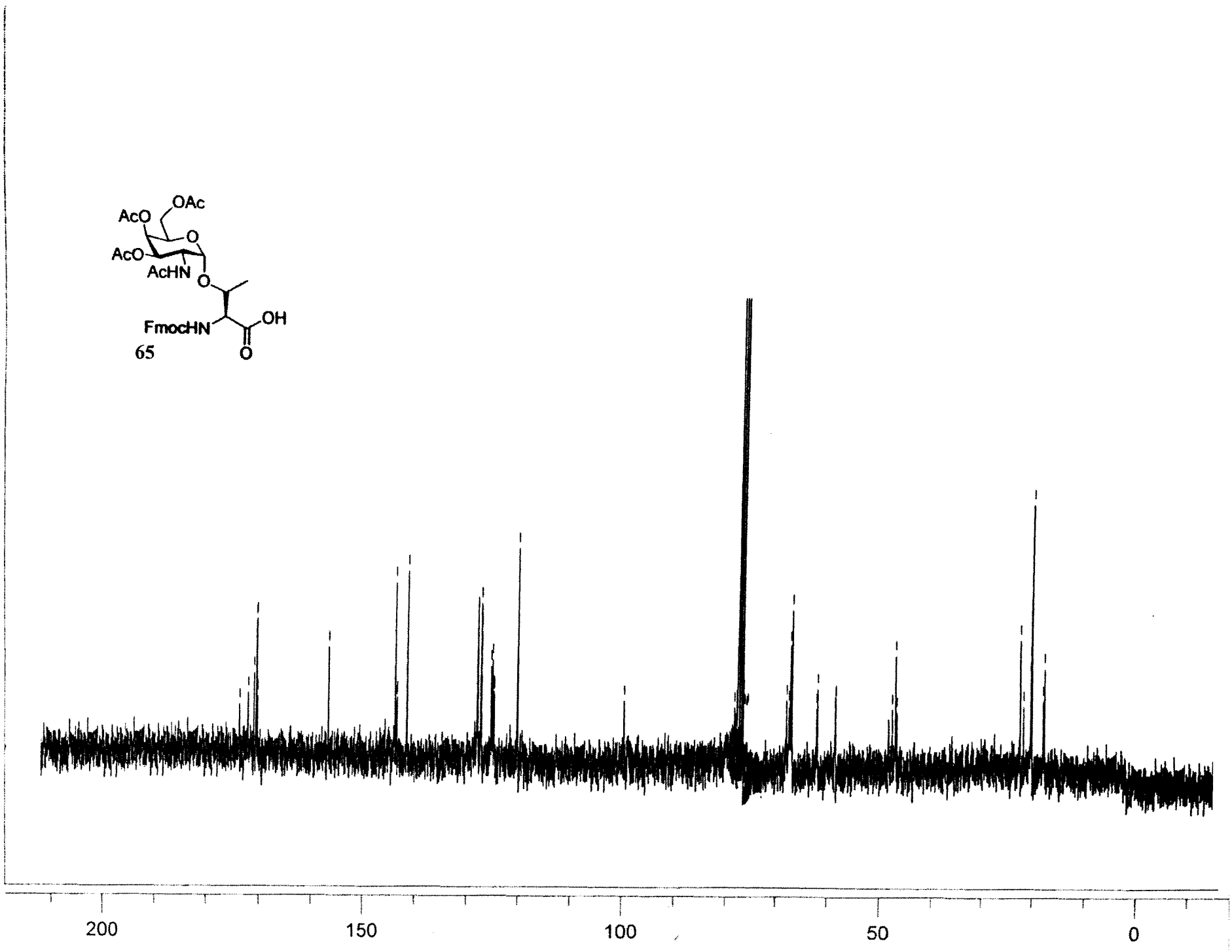
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

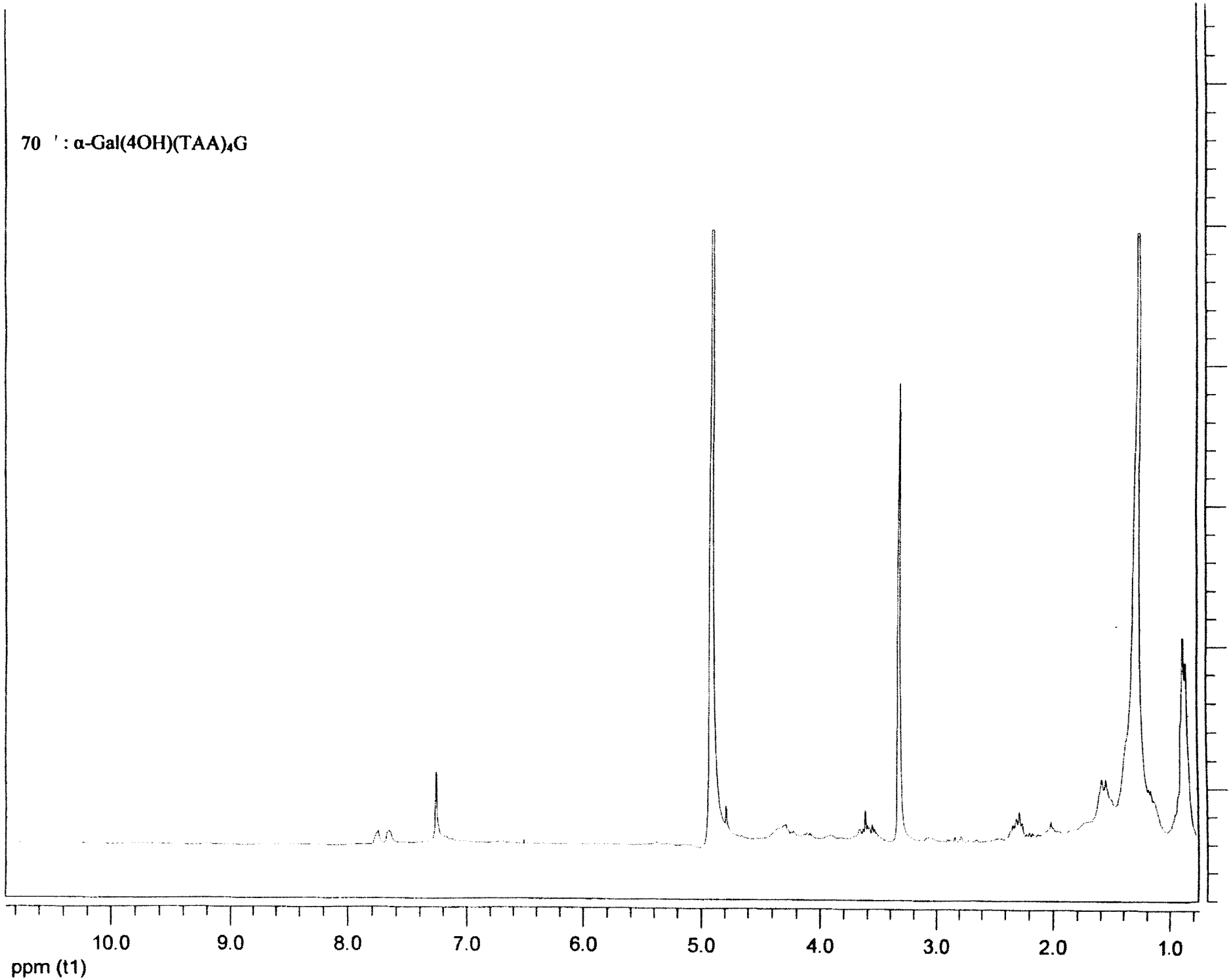

AcO
OAc
O
AcO
AcHN
O
FmocHN
OBn
O
64
200
150
100
50
0


OAc
AcO
O
AcO
AcHN
O
OH
FmocHN
O

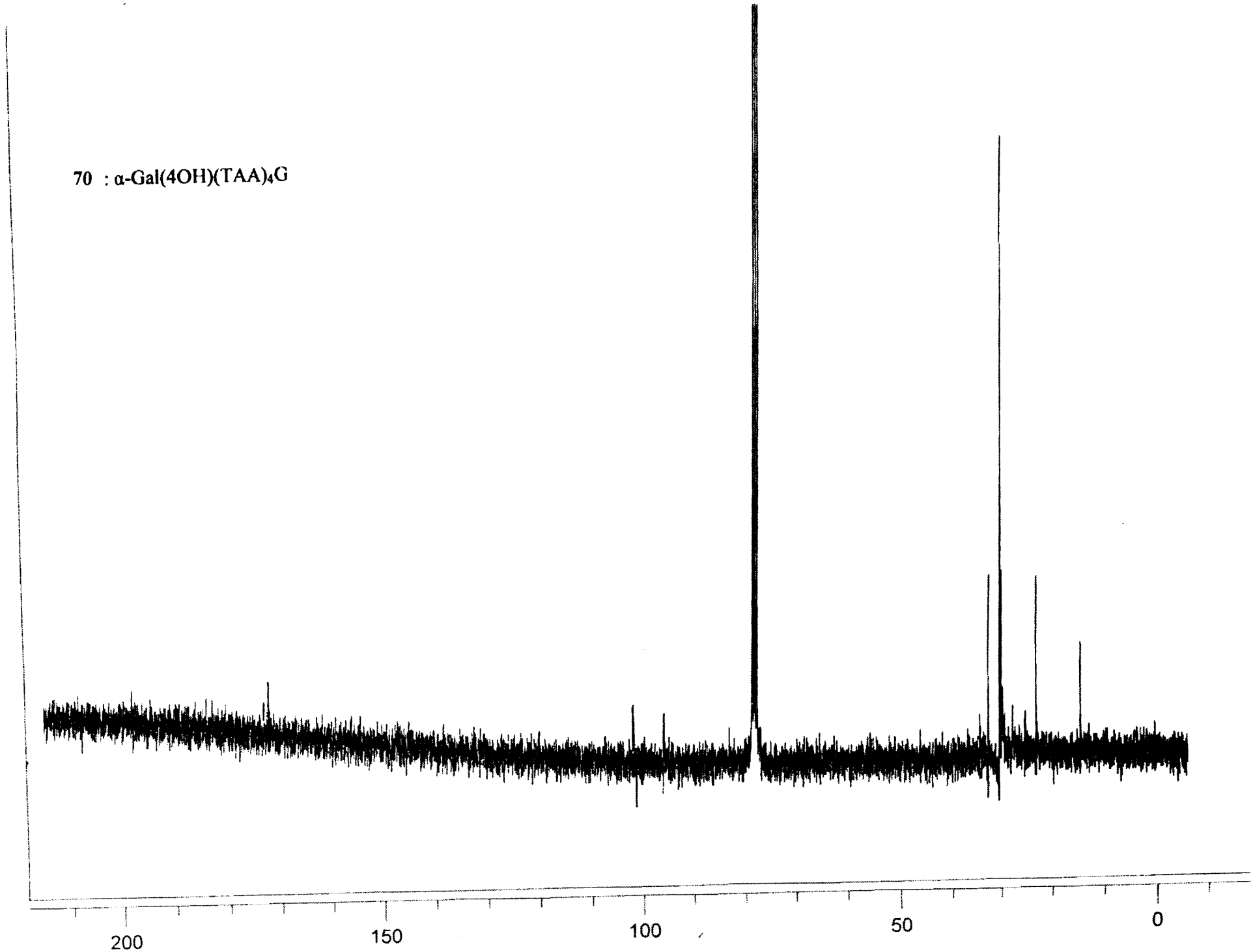
65

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

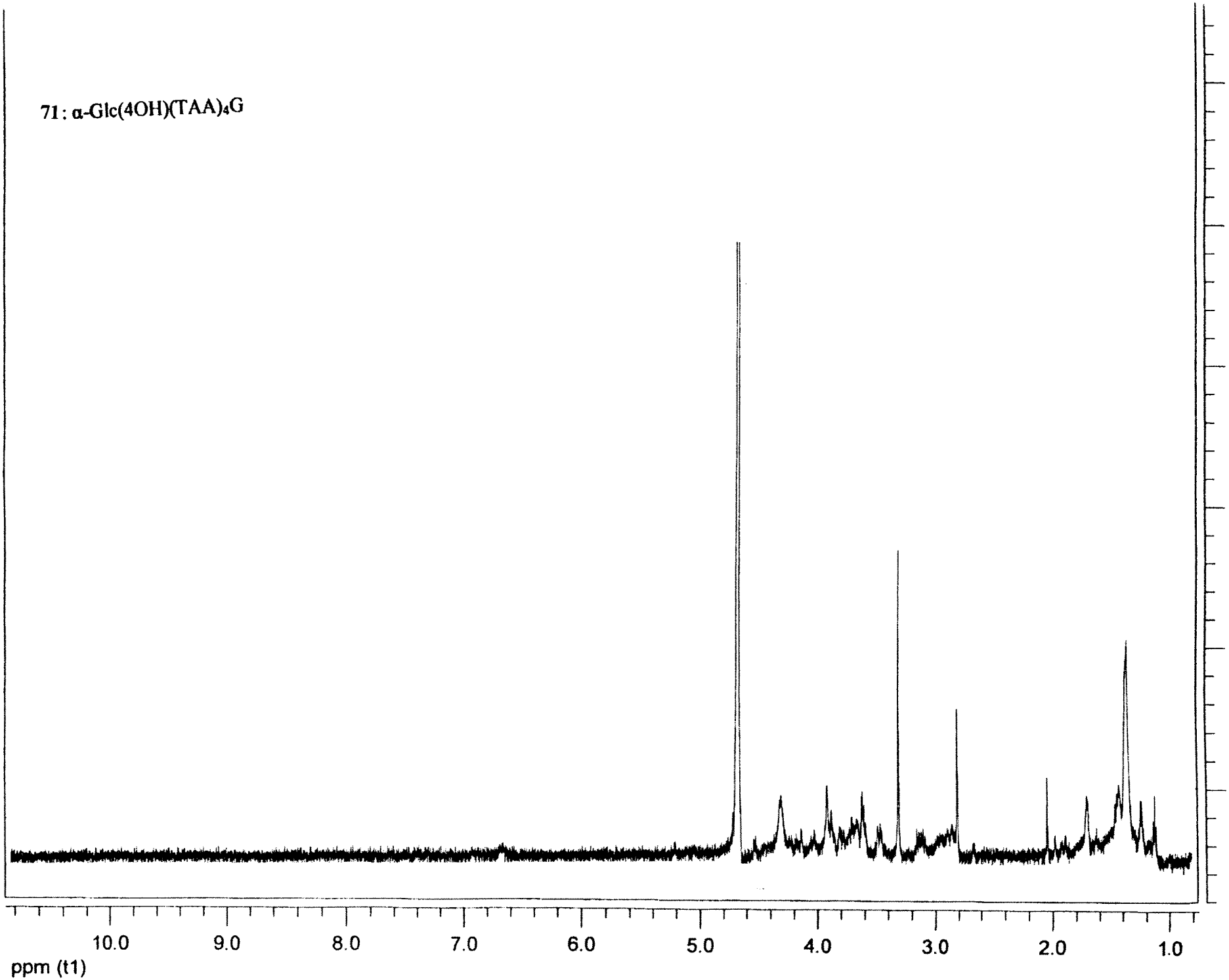

AcO
OAc
O
AcO
AcHN
O
FmocHN
OH
O
65
200
150
100
50
0


70 ' : α-Gal(4OH)(TAA)4G
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

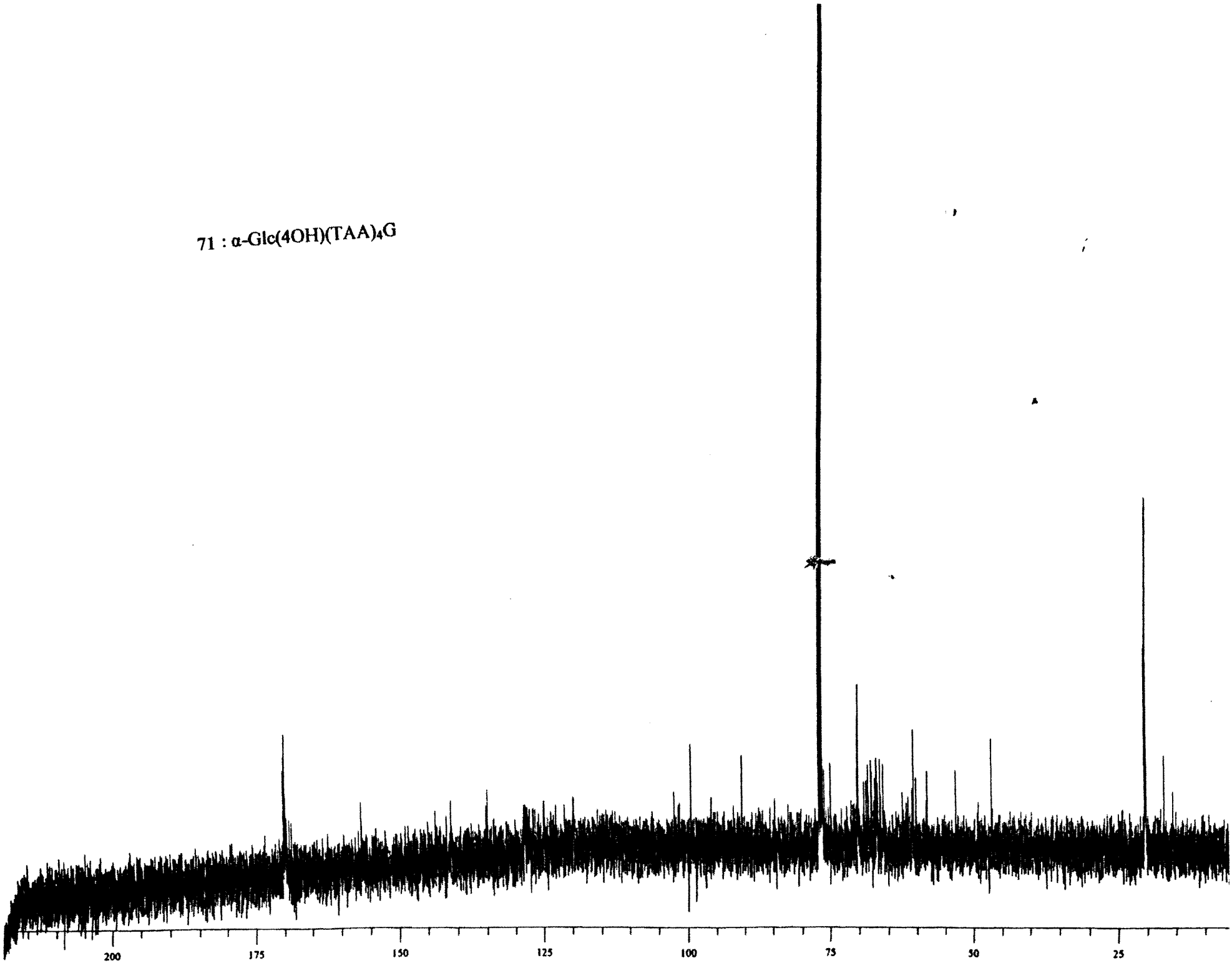
70 : α-Gal(4OH)$(TAA)_4$G



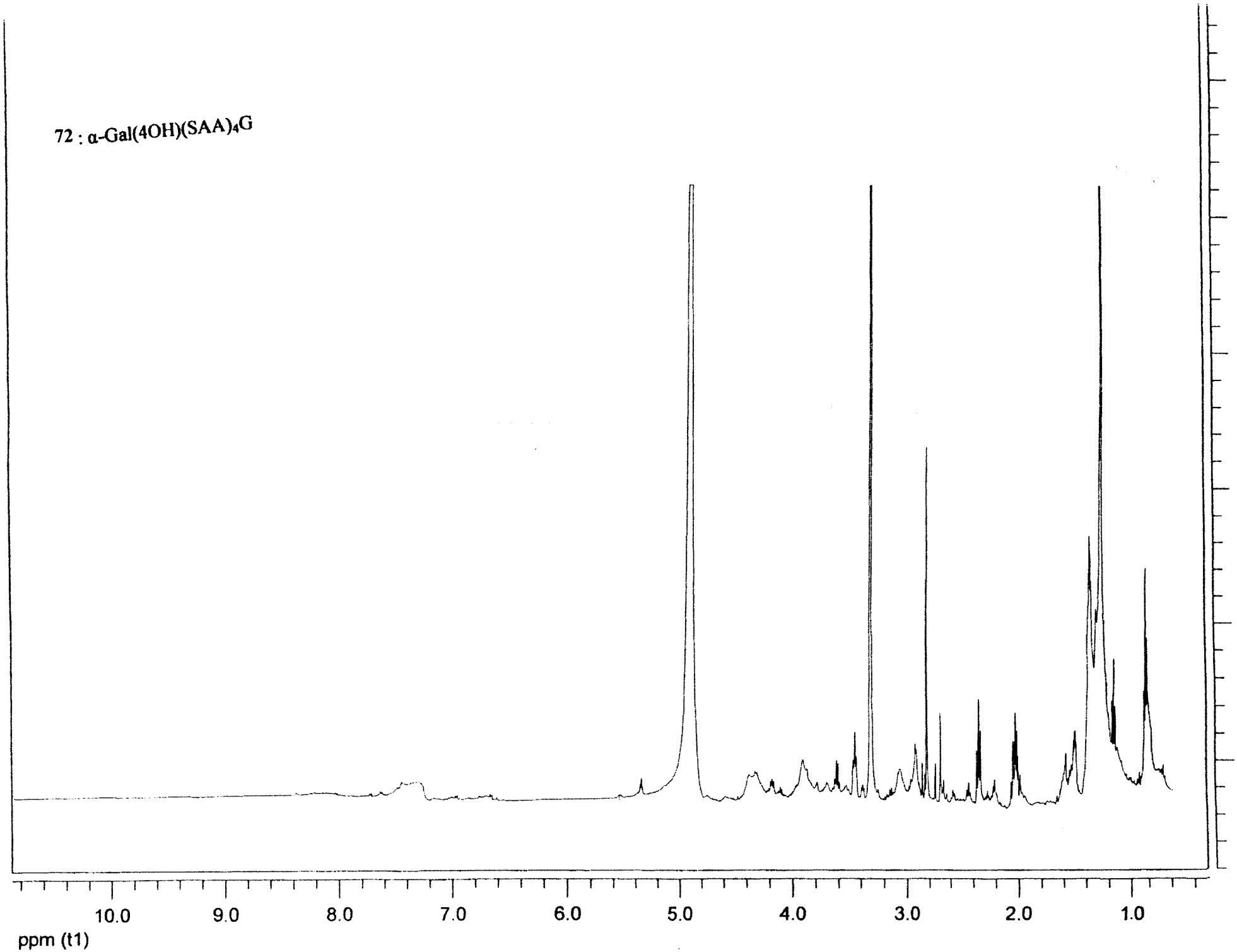
71 : α-Glc(4OH)(TAA)$_4$G



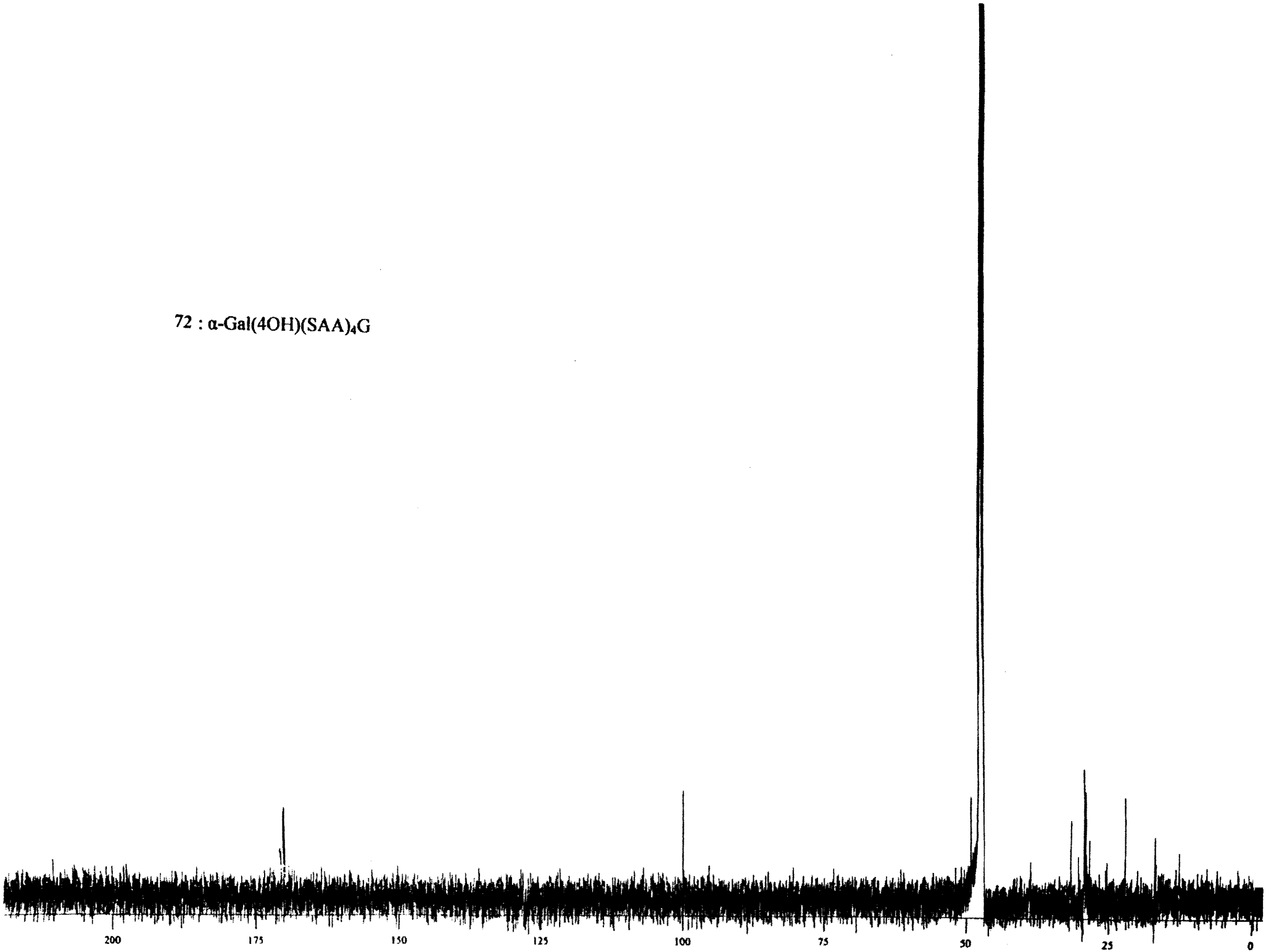
71 : α-Glc(4OH)$(TAA)_4$G



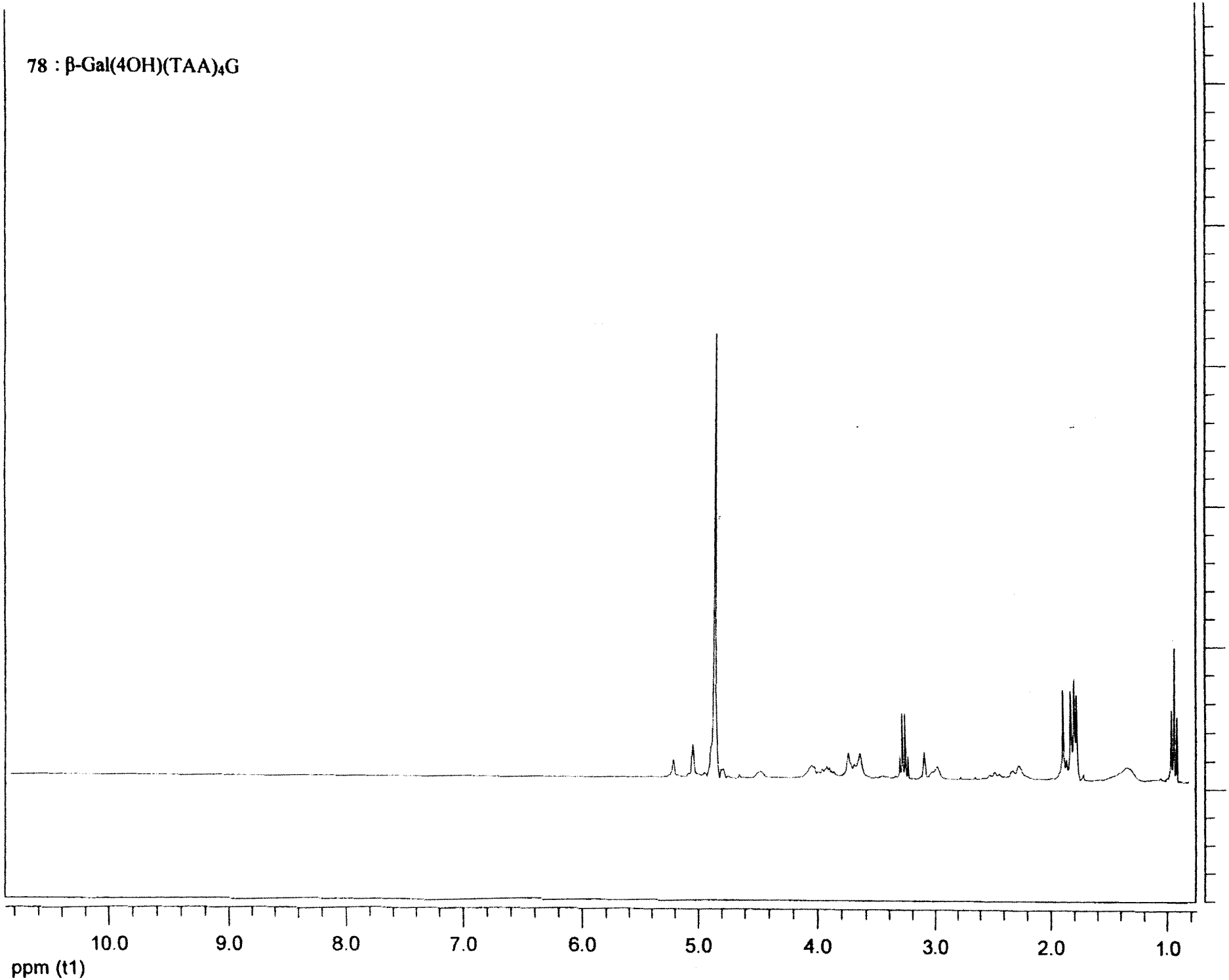
72 : α-Gal(4OH)(SAA)$_4$G



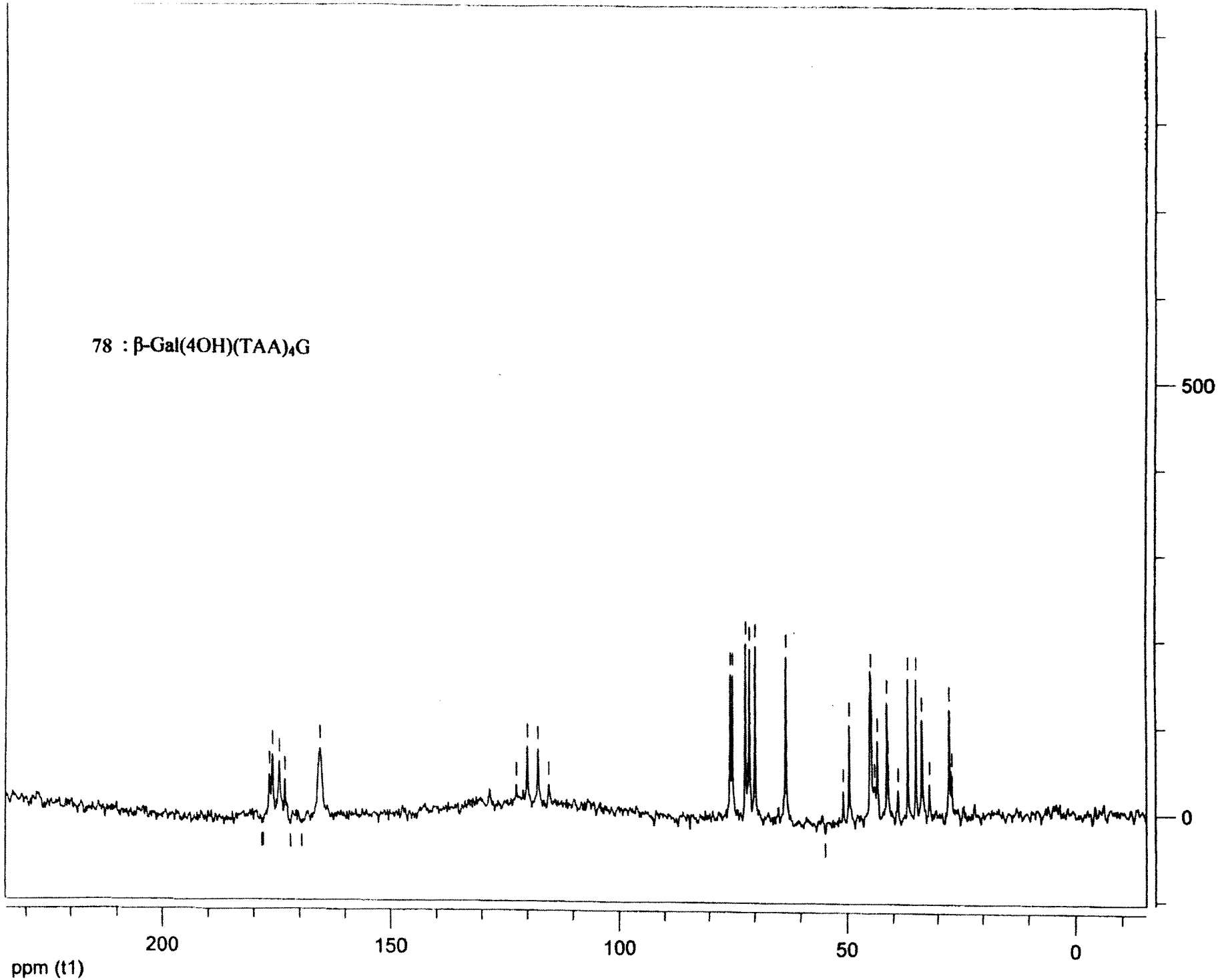
72 : α-Gal(4OH)$(SAA)_4$G



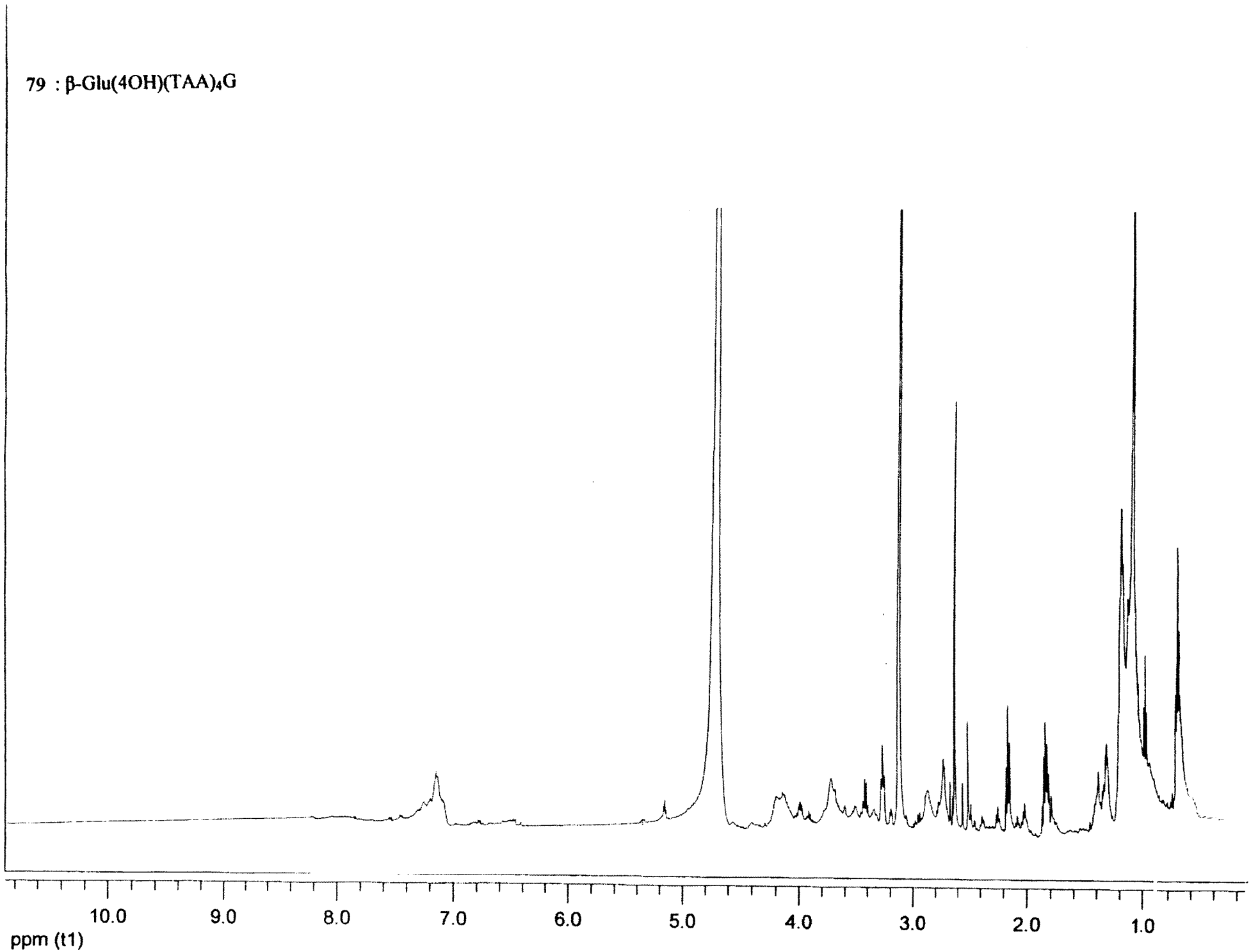
78 : β-Gal(4OH)$(TAA)_4$G



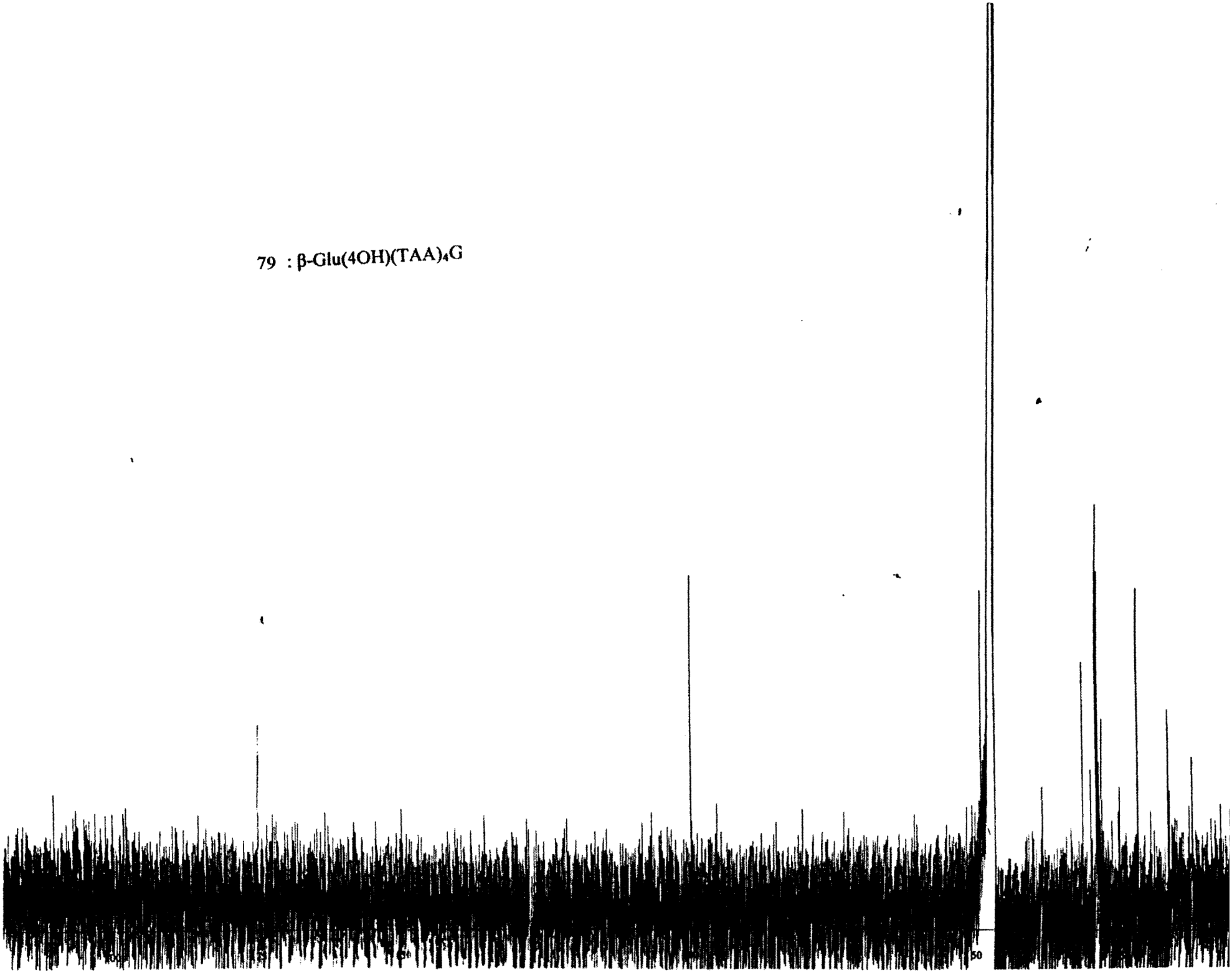
78 : β-Gal(4OH)$(TAA)_4$G



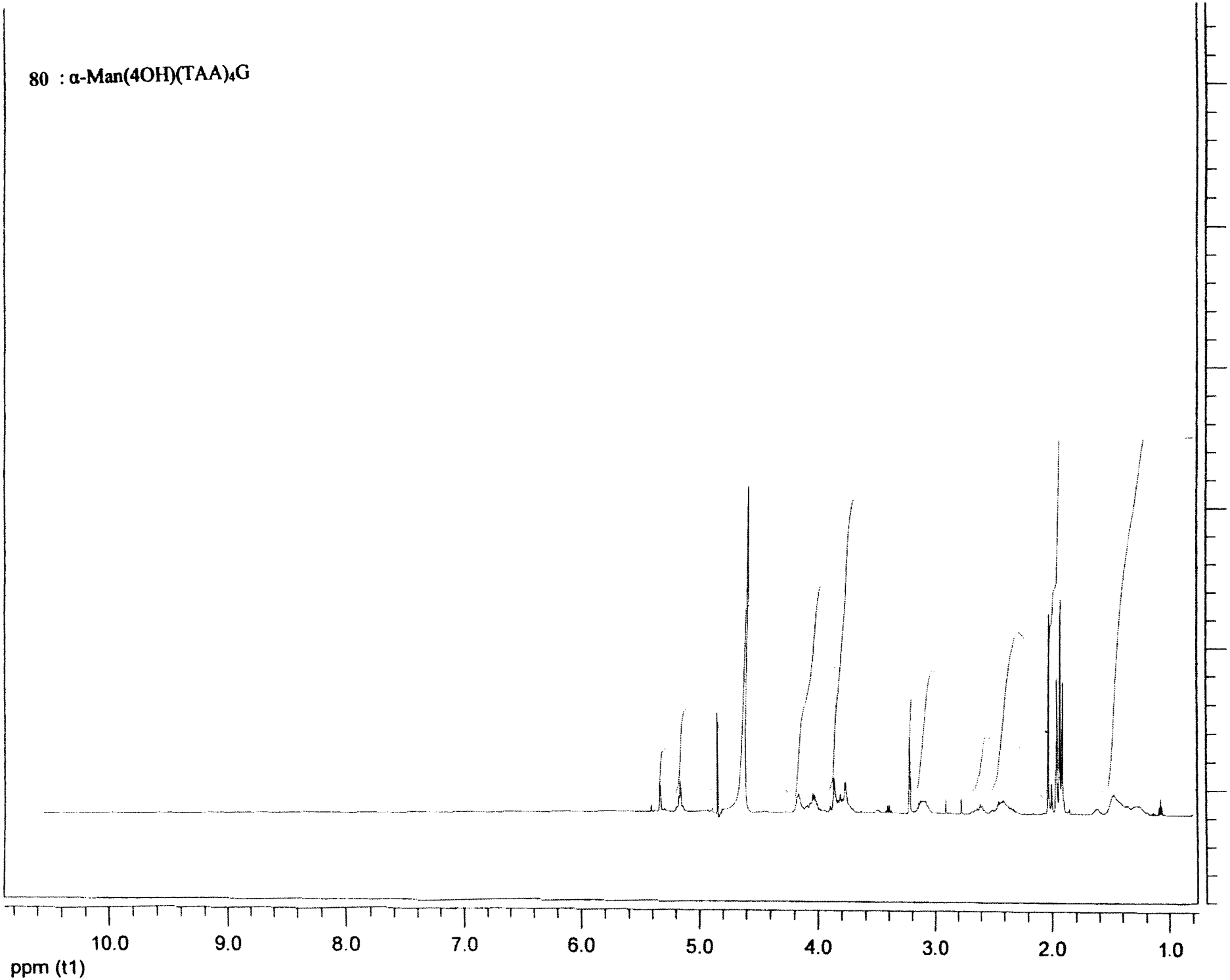
79 : β-Glu(4OH)$(TAA)_4$G



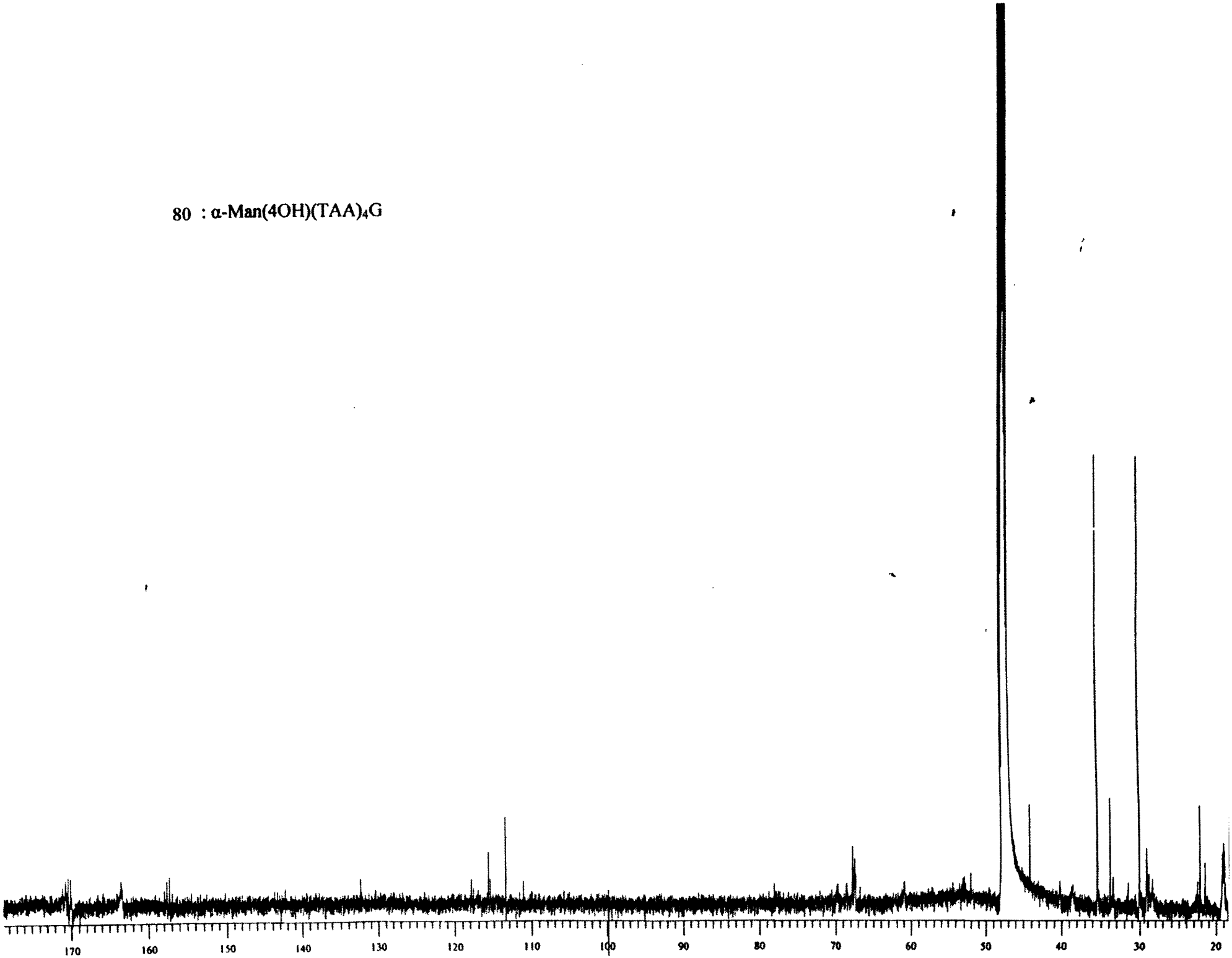
79 : β-Glu(4OH)$(TAA)_4$G



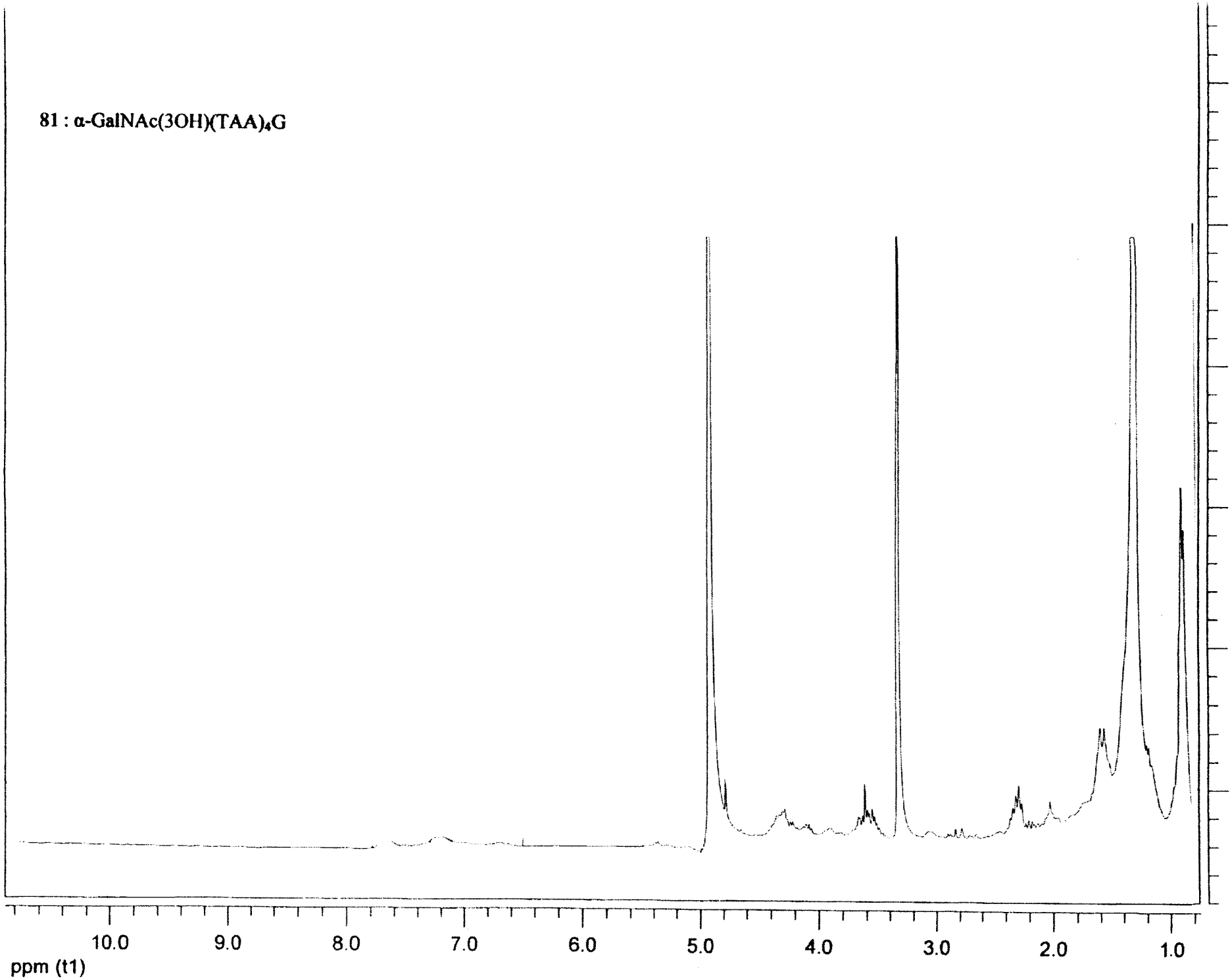
80 : α-Man(4OH)$(TAA)_4$G



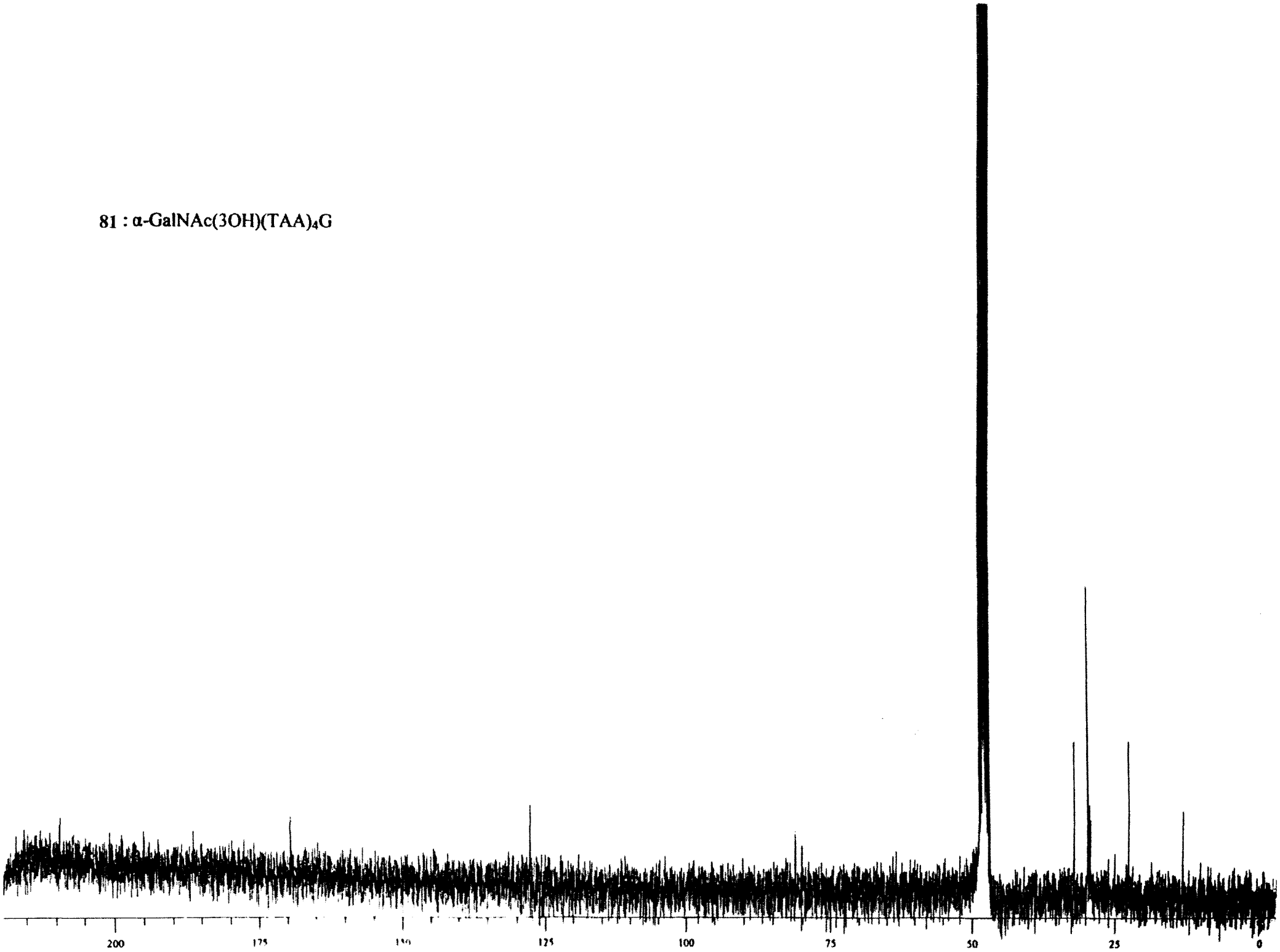
80 : α-Man(4OH)$(TAA)_4$G

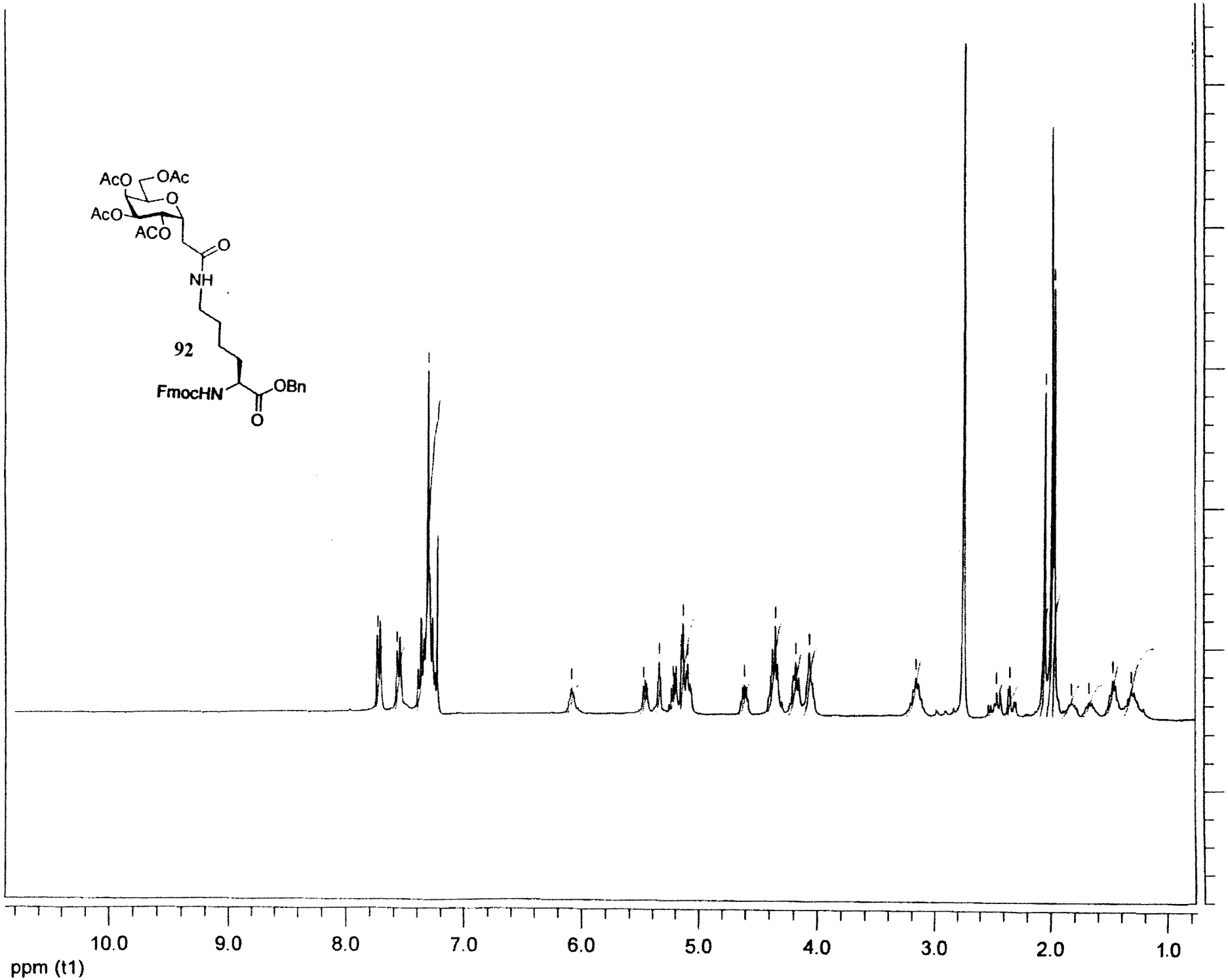


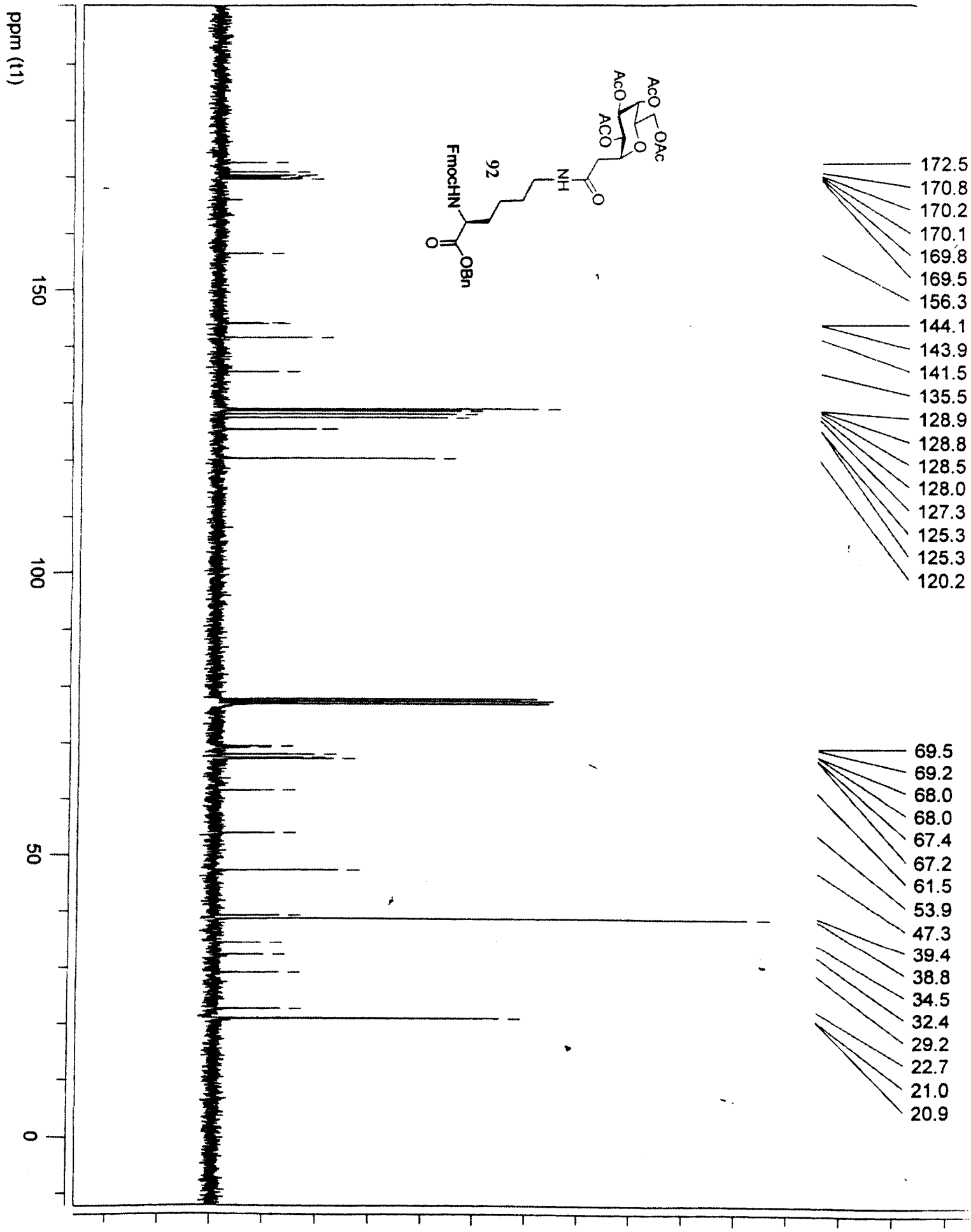
81 : α-GalNAc(3OH)(TAA)$_4$G

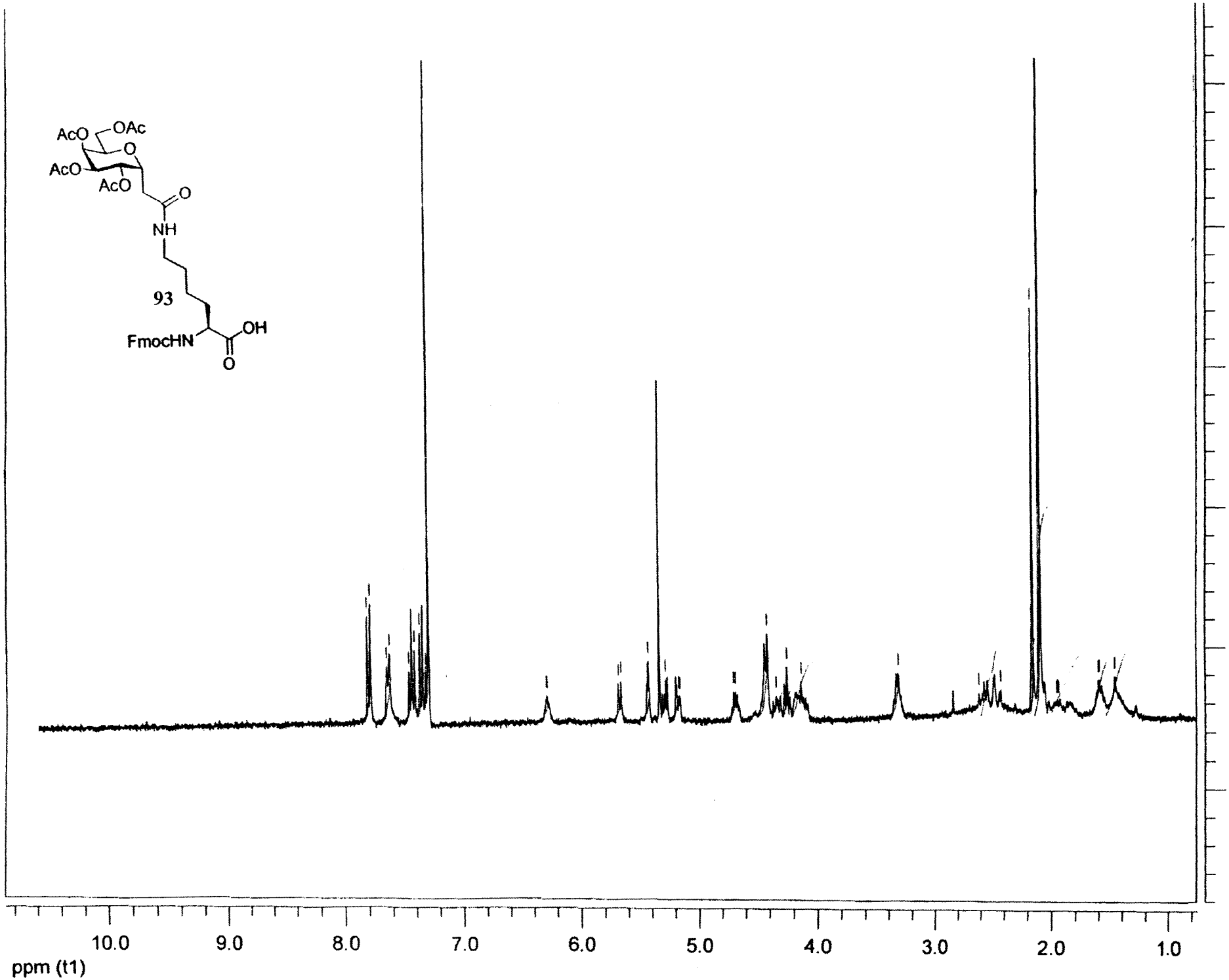


81 : α-GalNAc(3OH)$(TAA)_4$G

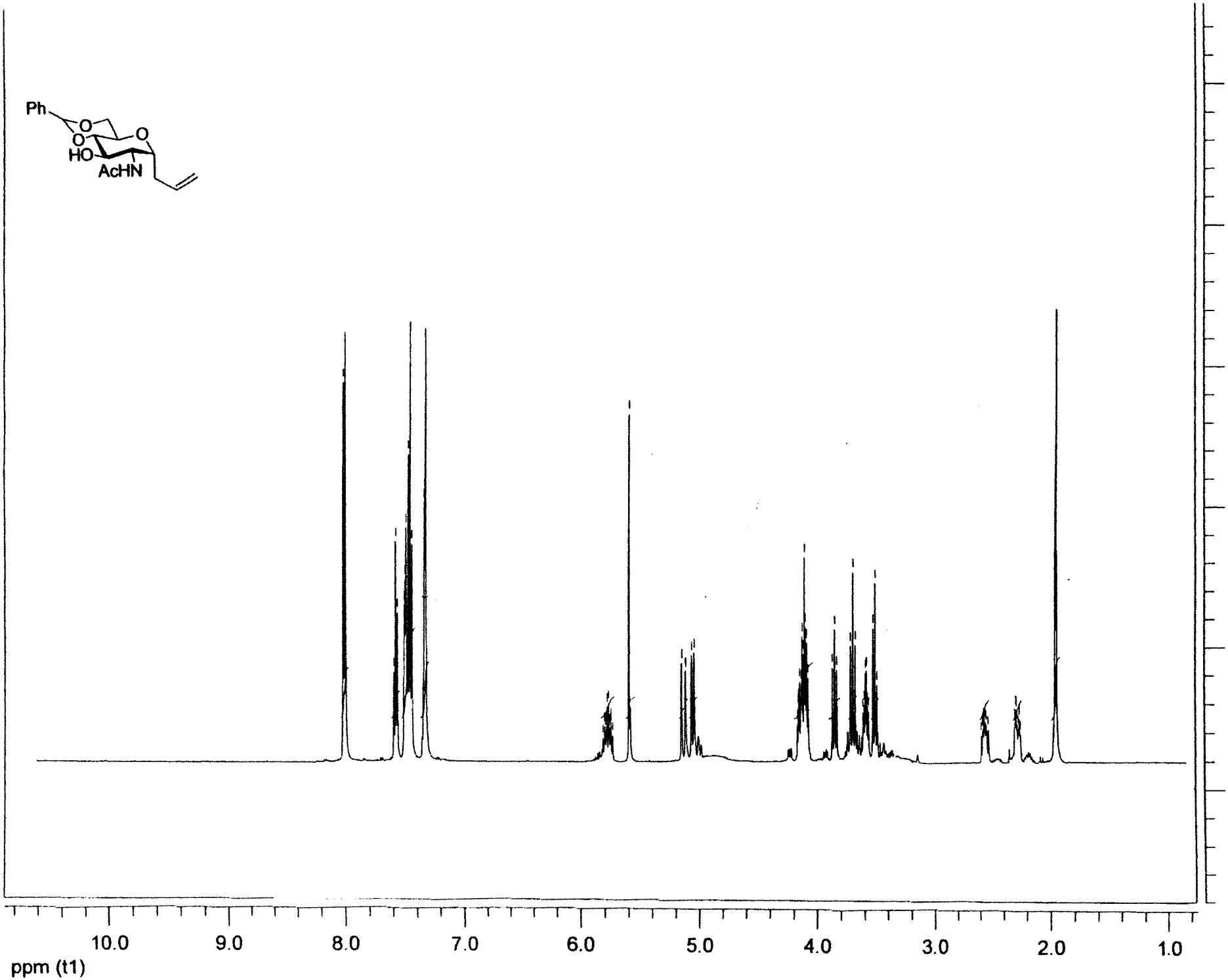



AcO
OAc
O
AcO
ACO
O
NH
92
FmocHN
OBn
O
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

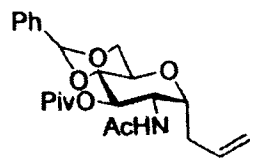

ppm (t1)
0
50
100
150
92
FmocHN
OBn
NH
AcO
AcO
AcO
OAc
O
172.5
170.8
170.2
170.1
169.8
169.5
156.3
144.1
143.9
141.5
135.5
128.9
128.8
128.5
128.0
127.3
125.3
125.3
120.2
69.5
69.2
68.0
68.0
67.4
67.2
61.5
53.9
47.3
39.4
38.8
34.5
32.4
29.2
22.7
21.0
20.9


AcO
OAc
O
AcO
AcO
O
NH
93
FmocHN
OH
O
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

ppm (t1)
200
150
100
50
0
0
100
200
300
400
93
FmocHN
NH
O
OH
AcO
AcO
AcO
OAc
O
174.8
172.8
171.3
170.8
170.4
170.1
156.5
144.0
141.6
128.1
127.4
125.4
120.3
69.5
69.3
68.2
68.1
67.3
61.6
53.8
51.0
47.4
39.5
34.5
32.1
29.0
22.5
21.1

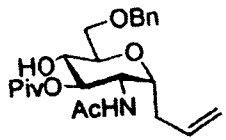

Ph
O
O
O
HO
AcHN
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

200
150
100
50
0
Ph
O
HO
O
AcHN
O

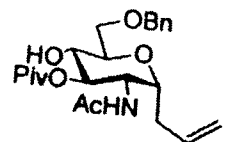


105

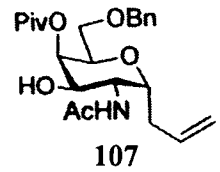
200
150
100
50
0
Ph
PivO
AcHN
105



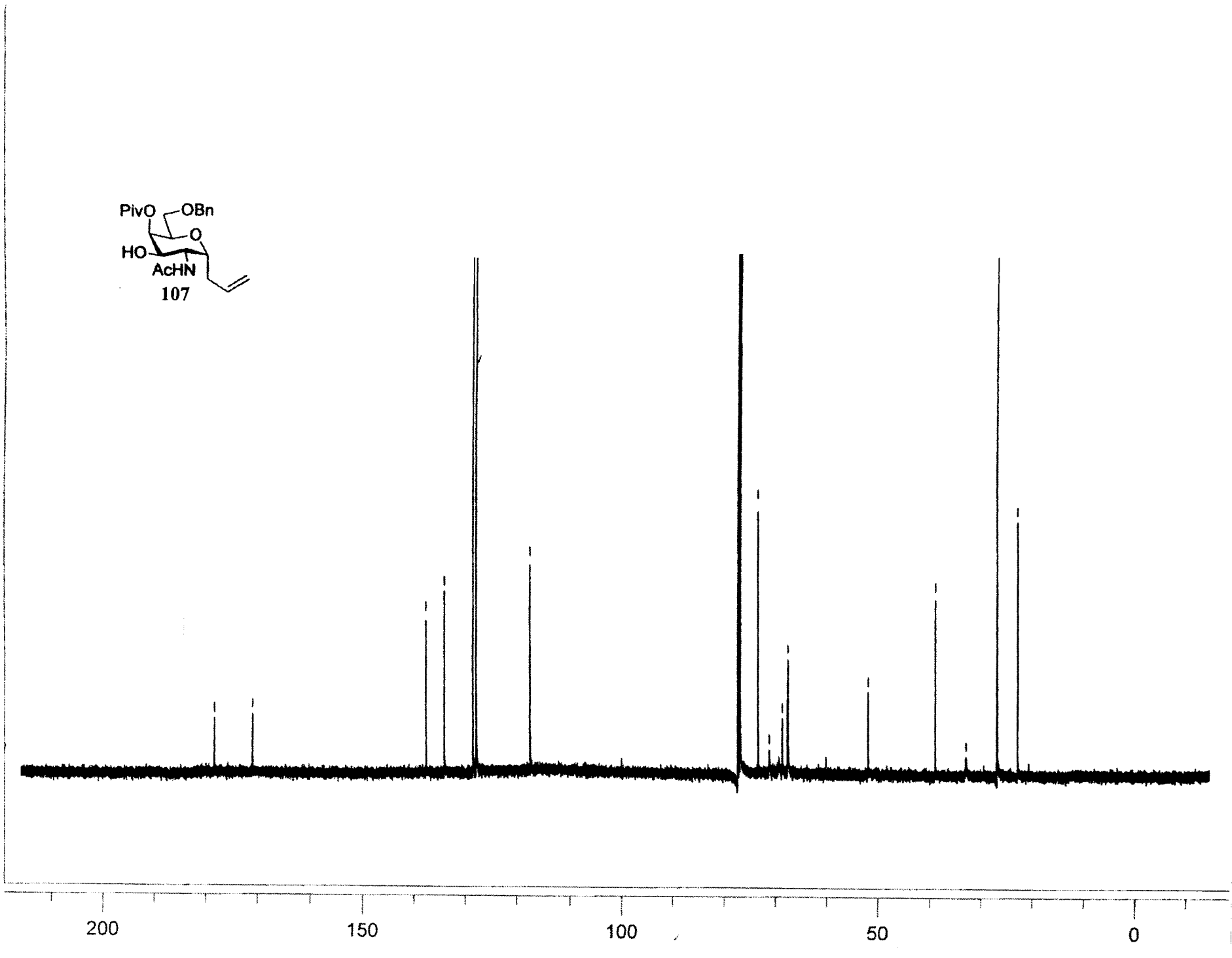
106

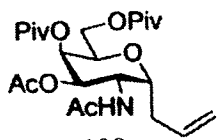


106

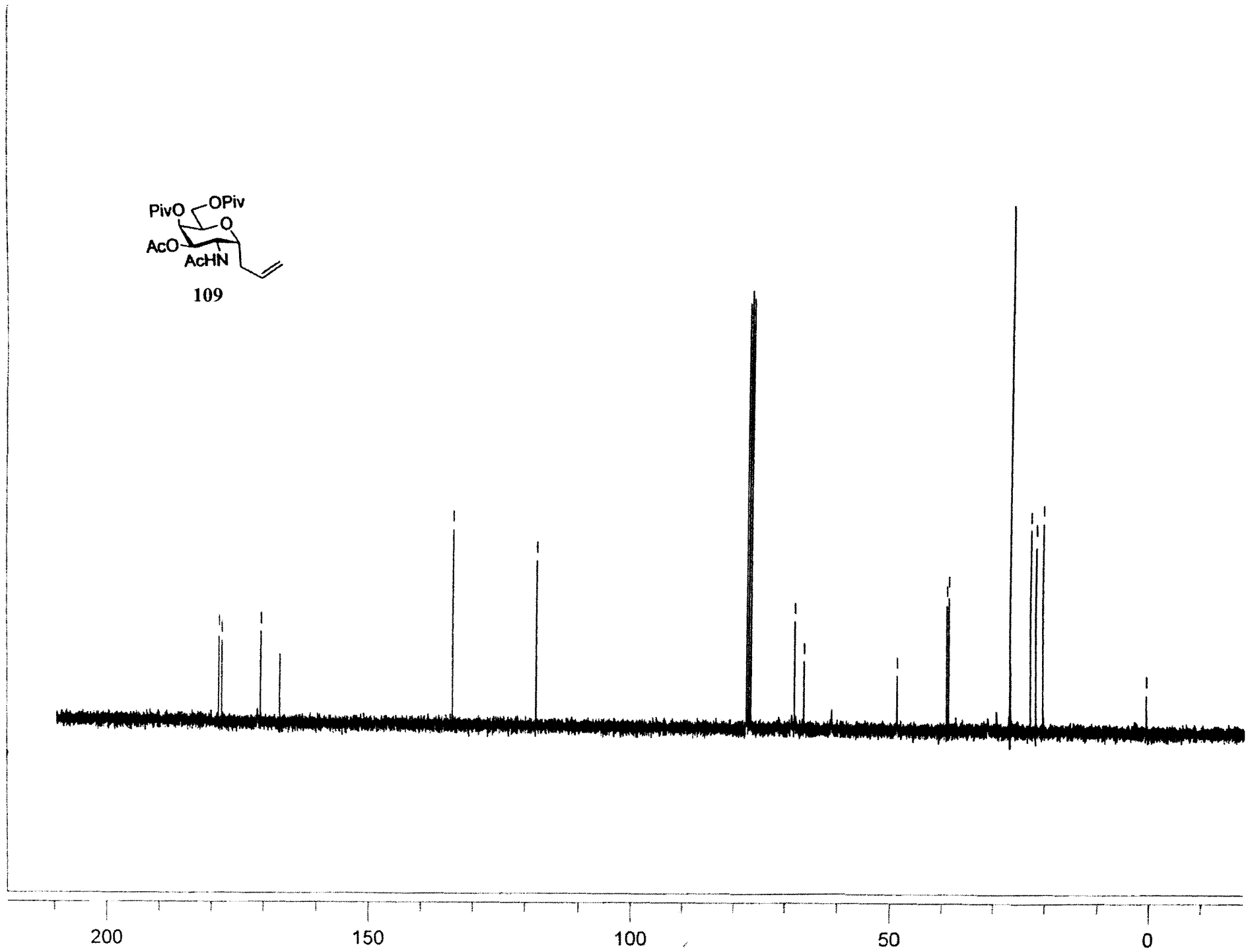


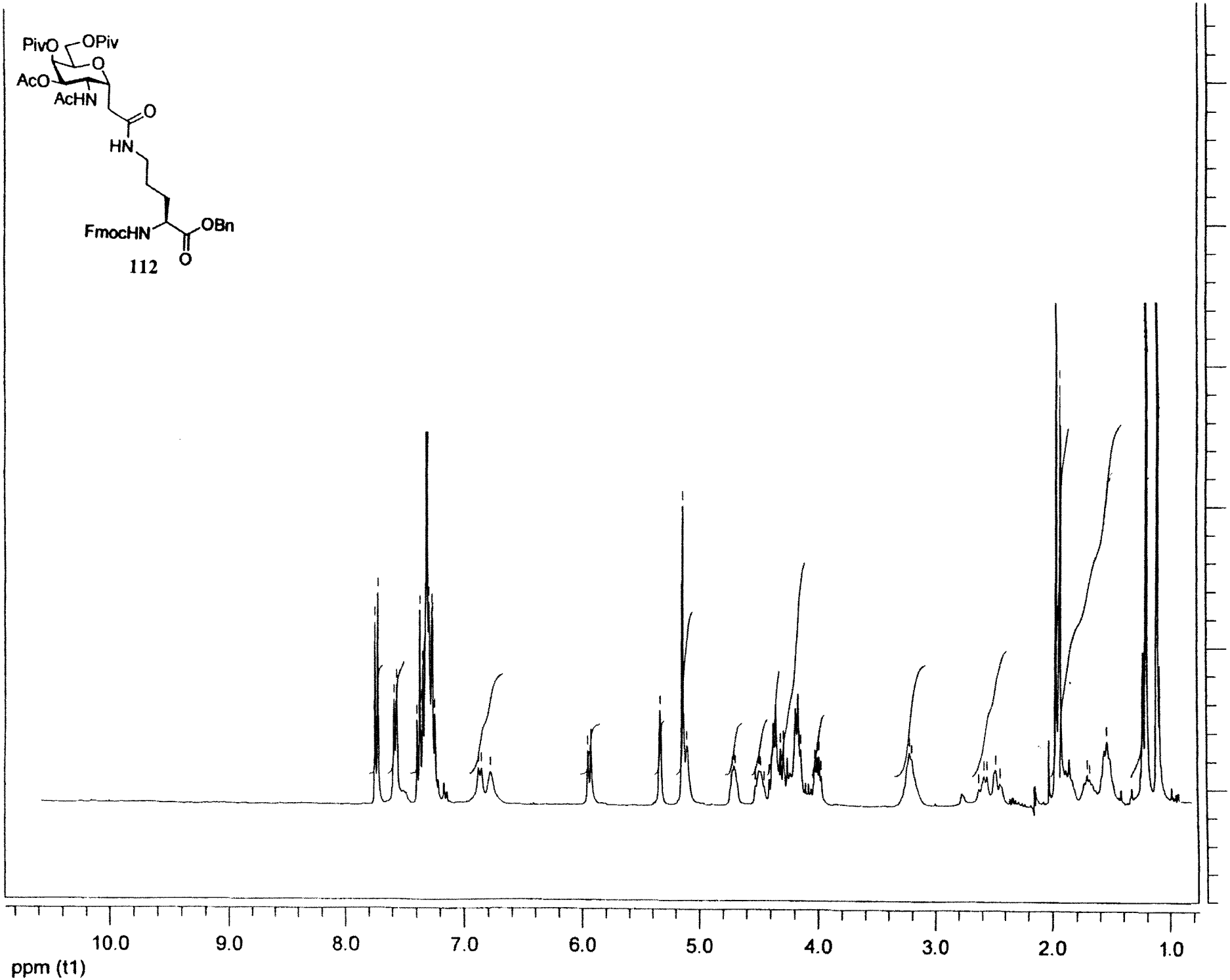
107

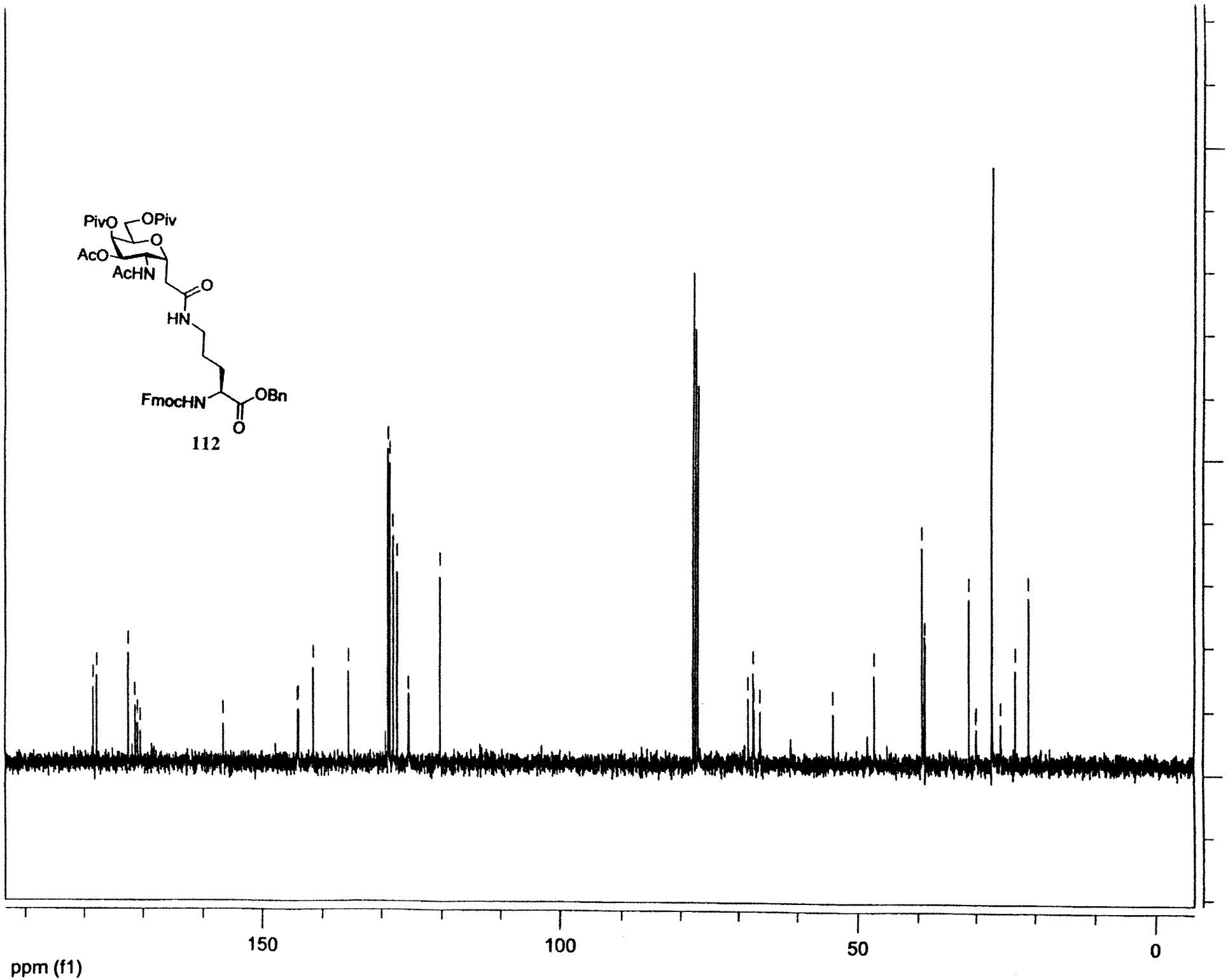

PivO
OBn
O
HO
AcHN
107
200
150
100
50
0

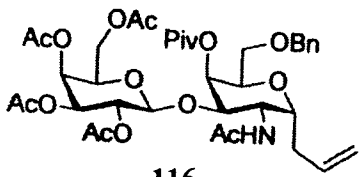


109


PivO
OPiv
O
AcO
AcHN
109
200
150
100
50
0

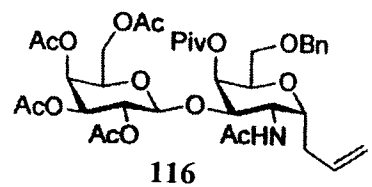

PivO
OPiv
O
AcO
AcHN
O
HN
FmocHN
OBn
O
112
10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)


PivO
OPiv
O
AcO
AcHN
O
HN
FmocHN
OBn
O
112
150
100
50
0
ppm (f1)

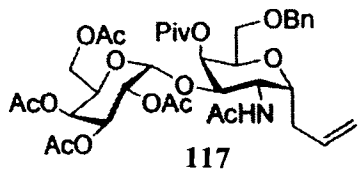

AcO
OAc
PivO
OBn
O
AcO
O
AcO
AcHN

116

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

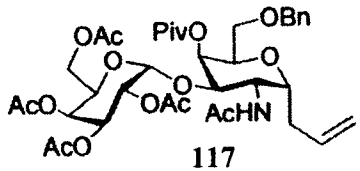


116

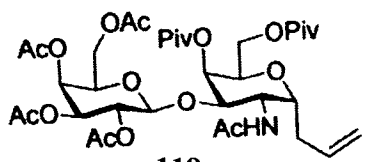

OAc
PivO
OBn
O
O
O
AcO
OAc
AcHN
AcO

117

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

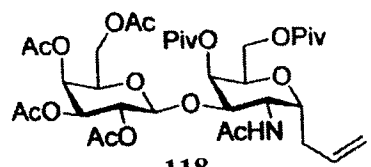


117


AcO
OAc
PivO
OPiv
AcO
AcO
AcHN

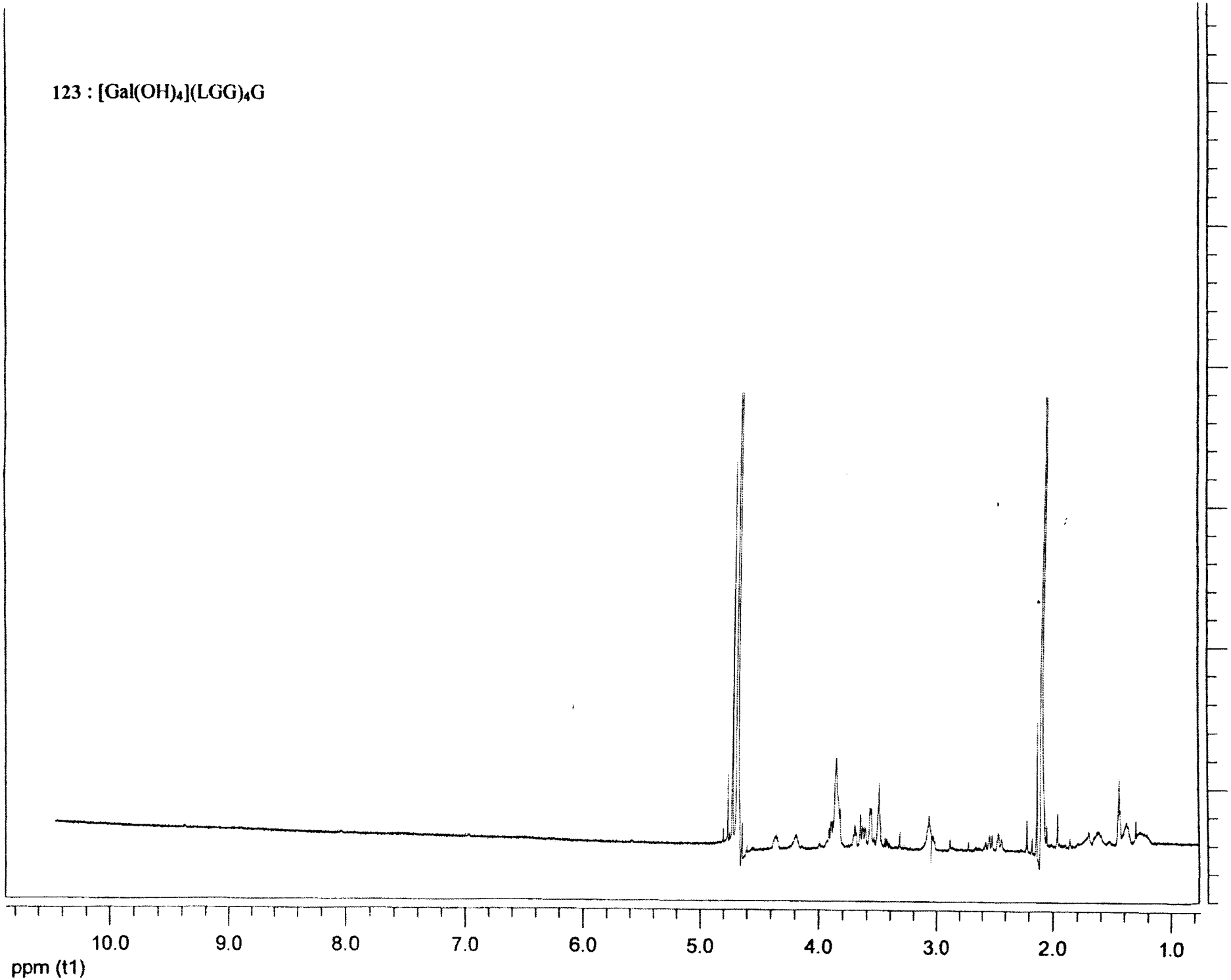
118

10.0
9.0
8.0
7.0
6.0
5.0
4.0
3.0
2.0
1.0
ppm (t1)

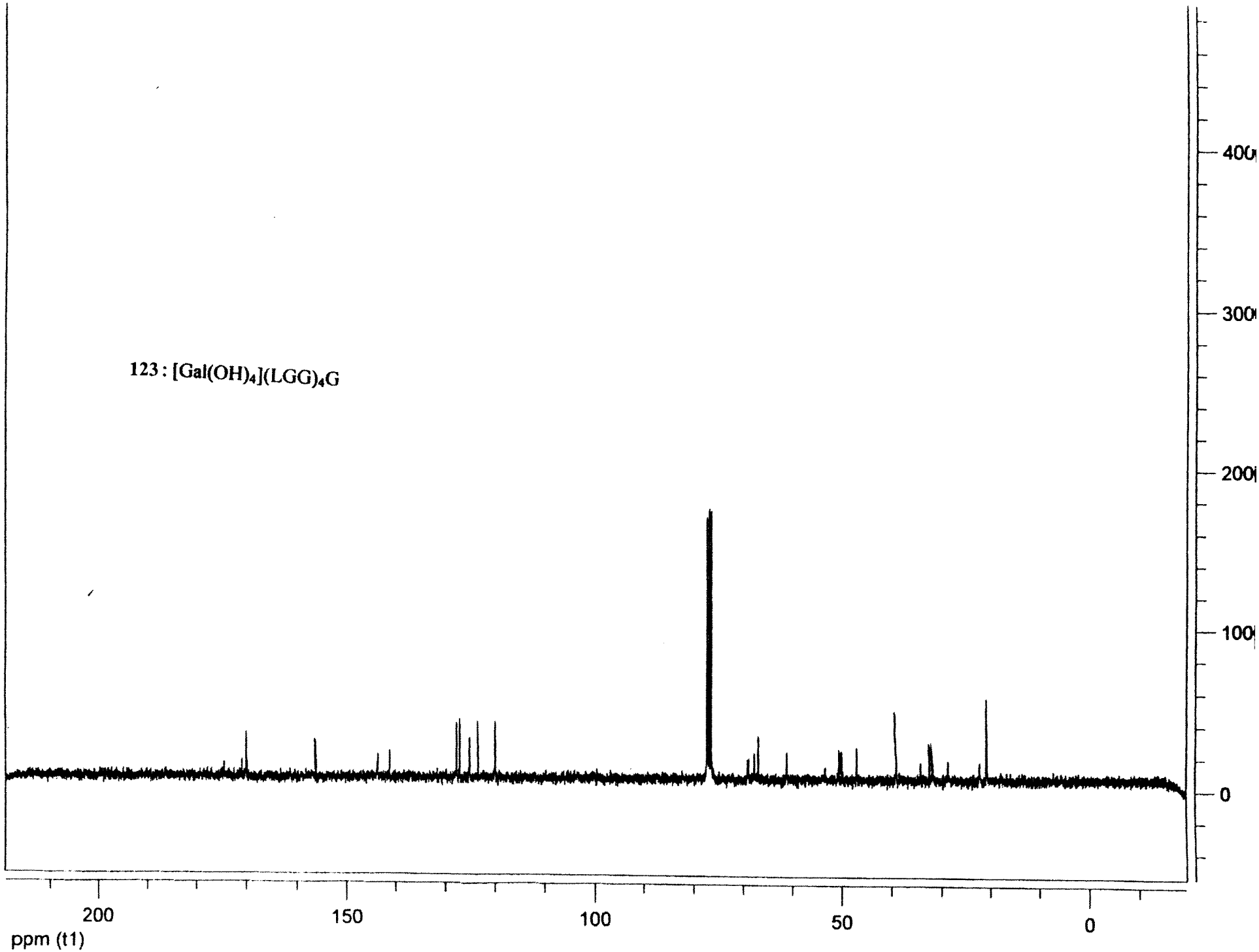


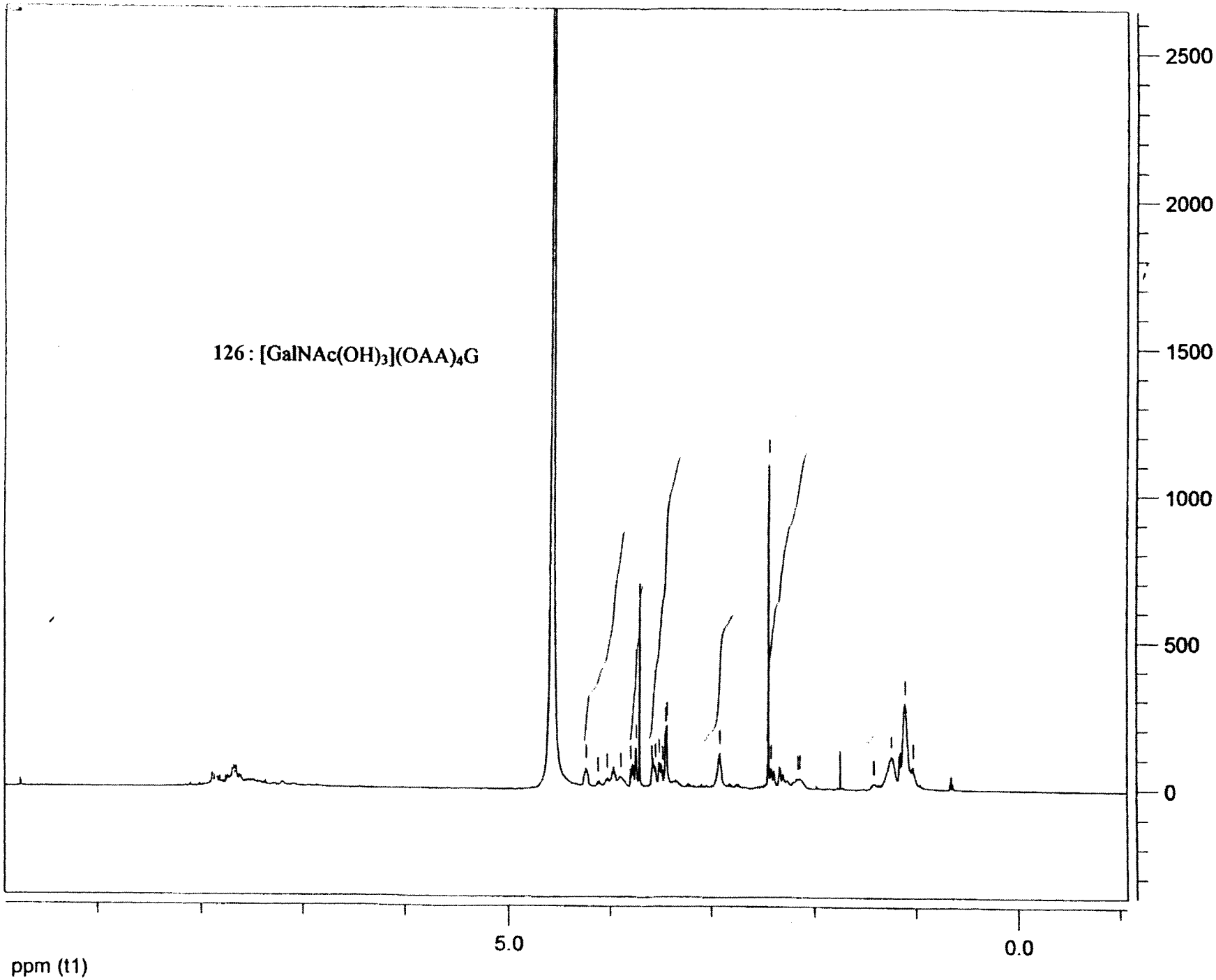
118

123 : $[Gal(OH)_4](LGG)_4G$



123 : $[Gal(OH)_4](LGG)_4G$




126 : [GalNAc(OH)3](OAA)4G
2500
2000
1500
1000
500
0
5.0
0.0
ppm (t1)

126 : [GalNAc(OH)$_3$](OAA)$_4$G